JÁNOSSY : VEKTORSZÁMÍTÁS I.

Jánossy Lajos
akadémikus,
tanszékvezető egyetemi tanár

Tasnádi Péter
egyetemi adjunktus

Vektorszámítás

I. kötet

Vektor- és tenzoralgebra

Tankönyvkiadó, Budapest, 1980

Egyetemi tankönyv
Készült az oktatási miniszter rendeletére

Bírálók:

KOVÁCS ISTVÁN
a fizikai tudományok doktora,
tanszékvezető egyetemi tanár

PÁL LÉNÁRD
akadémikus,
a Központi Fizikai Kutató Intézet főigazgatója

ISBN 963 17 3414 5
ISBN 963 17 3415 3

ELŐSZÓ

E tankönyv előzménye egyetemi előadássorozat, amelyet az egyik szerző több éven keresztül, kezdő fizikusok számára tartott.

Az előadás célja az volt, hogy a kezdő fizikusok megfelelő időben és színvonalasan ismerkedjenek meg azokkal a matematikai módszerekkel, amelyek a fizikában szükségesek, elkerülve ezzel, hogy a matematikai tételeket akkor ismerjék meg alaposan, amikor e tételeket a fizikában már „elfogadott" szabályként régen használták.

A matematika egy részének ilyenfajta előrehozása bizonyos áldozatokat kíván. Igyekeztünk azonban ezt az alkalmazásra szabott matematikát logikailag precízen megfogalmazni.

Az előadások kifejlesztése során felmerült, hogy hasznos lenne a tárgykört a kidolgozott módszerek alkalmazásával kibővíteni. A tervet megvalósítandó, az előadások anyagát tartalmazó jegyzetet nagymértékben kibővítve, kézikönyvet dolgoztunk ki. Jelen mű a kézikönyv első kötete.

A könyv — bizonyos részek kihagyásával — elsőéves hallgatók számára is hasznosnak bizonyulhat, de olvasóként magasabb évfolyamú vagy már gyakorló fizikusokat, mérnököket, kémikusokat, általában tehát oktatókat és kutatókat képzeltünk el.

A mű nem matematikusok számára készült, azonban a tárgyalásmódban matematikusok is találhatnak érdekességet. Ez az érdekesség egyrészt az egyes tételek levezetésekor alkalmazott sajátos fogásokban, másrészt talán az anyag szokásostól eltérő felépítésében rejlik.

A mű sajátossága, hogy noha matematikailag precíz kíván lenni, nem matematikai tételekből és ezek bizonyításából indul ki, hanem fizikai meggondolások alapján bizonyos fogalmak bevezetésének célszerűségét mutatja ki, és csak további lépésként adja meg az így bevezetett fogalmak és módszerek általános matematikai megfogalmazását.

Példaként említjük, hogy vektorokat kezdetben mint sajátságos mennyiségeket vezettünk be, amelyekkel fizikai mennyiségeket, pl. sebességet, erőt stb. írhatunk le. A vektorösszeadást és a vektorok különböző szorzatait is fizikai példák kapcsán vezetjük be, s csak ezután térünk ki a vektorösszeadás és -szorzás koordinátareprezentációban kifejezhető formáira. A gondolatmenet világosabb érzékeltetésére különleges jelölést vezettünk be a fizikai mennyiségekre, $\underline{a}$-val jelöljük pl. magát a vektormennyiséget és

$$\mathcal{K}(\underline{a})=\mathbf{a}$$

-val azt az $\mathbf{a}=a_1, a_2, a_3$ számhármast, amely az $\underline{a}$ vektort adott $\mathcal{K}$ koordinátarendszerben reprezentálja.

Az $\underline{a}$ és $\mathbf{a}$ közötti különbséget lényegesebbnek tartjuk, mint pl. a vektor és tenzor közötti különbséget.

Módszerünk sajátosságait még egy példával érzékeltetjük. A vektorok általános tulajdonságainak felhasználásával meghatározzuk, hogy egy szilárd test pontjaihoz vezető helyzetvektorok hogyan változnak, ha a testet elforgatjuk, illetve több egymás utáni elforgatásnak vetjük alá. A kapott körülményes matematikai formalizmust *általánosítjuk*, és e meggondolások alapján vezetjük be a mátrix fogalmát, megmutatva a mátrixokra vonatkozó műveleti szabályok célszerűségét is.

A tárgyalt anyag feldolgozása során mindvégig kerüljük az egyes mennyiségek tisztán matematikai úton történő posztulálását. A matematikai szabályok ismertetése előtt mindig láttatni akarjuk a problémát, amely az adott formalizmus kidolgozását szükségessé teszi.

Reméljük, hogy e mű mind az egyetemi hallgatók, mind a kutatók számára segítséget nyújt egyes problémák kidolgozásában, de reméljük azt is, hogy a szemlélet, amellyel megmutatjuk a fizikai valóságtól a formalizmushoz vezető utat, ugyancsak hasznosnak bizonyul mindazok számára, akik a matematika segítségével új fizikai jelenségeket igyekeznek leírni.

Budapest, 1977. június 30.

A szerzők

I. SKALÁR- ÉS VEKTORMENNYISÉGEK

1. Skaláris mennyiségek

1.1. Fizikai mennyiségek és mérőszámok

A fizikai jelenségek mennyiségi leírására számokat használunk. Legegyszerűbbek az olyan mennyiségek, amelyek egyetlen számadattal jellemezhetők. Ezeket a mennyiségeket *skaláris* mennyiségeknek nevezzük. (Skaláris mennyiség például a térfogat, tömeg, elektromos töltés stb.)

Amikor egy fizikai mennyiséget mérőszámmal jellemzünk, tulajdonképpen azt mondjuk meg, hogy a fizikai mennyiség egy előre megválasztott mértékegységnek hányszorosa. Az, hogy egy rúd hossza három méter, annyit jelent, hogy a rúd egyik végétől a másik végéig a méterrudat háromszor helyezhetjük egymás mellé. Ugyanezt fejezi ki az is, hogy a rúd hossza 300 cm, csak ekkor a mértékegységül választott hossz rövidebb. A rudat tehát egyformán jól jellemezhetjük a 3-as, illetve a 300-as mérőszámmal is.

Lényeges, hogy a rúd hosszát, mint *fizikai mennyiséget*, ne tévesszük össze a *mérőszámmal*. A rúd egy objektum, amelynek tulajdonságai számokkal bizonyos mértékben, de nem kimerítően jellemezhetők.

A fizikai mennyiségek közötti összefüggéseket — és ezen összefüggések törvényszerűségeit — sok esetben a mérőszámok közötti függvénykapcsolatokkal adhatjuk meg. Ezek a függvénykapcsolatok gyakran algebrai formát öltenek. A következőkben röviden összefoglaljuk az algebrai műveletekre vonatkozó szabályokat.

1.2. Algebrai szabályok

Először összefoglaljuk a nem negatív számokra vonatkozó algebrai törvényeket, majd röviden foglalkozunk a negatív számok bevezetésével.

Összeadás:

$$a+b=b+a \quad \text{kommutatív törvény} \tag{1.2-1}$$

$$a+0=a \quad \text{zérus definíciója} \tag{1.2-2}$$

$$a+(b+c)=(a+b)+c \quad \text{asszociatív törvény} \tag{1.2-3}$$

Érvényes továbbá az

$$a+b>a \tag{1.2-4}$$

egyenlőtlenség, ha

$$b>0.$$

Szorzás:

$$ab=ba \quad \text{kommutatív törvény} \tag{1.2-5}$$

$$1\cdot a=a \quad \text{az egység definíciója} \tag{1.2-6}$$

$$0\cdot a=0 \quad \text{a zérusra vonatkozó szorzási törvény} \tag{1.2-7}$$

$$a(bc)=(ab)c \quad \text{asszociatív törvény} \tag{1.2-8}$$

Érvényes még az

$$ab>a \tag{1.2-9}$$

egyenlőtlenség, ha

$$b>1.$$

Az összeadás és szorzás között az

$$a(b+c)=ab+ac \tag{1.2-10}$$

disztributív törvény teremt kapcsolatot.

1.3. Kivonás és negatív számok

A fizikai jelenségek leírásakor szükségünk van a kivonás műveletére is. Legyen $b\geq c$, akkor elvégezhetjük a

$$b-a=c \tag{1.3-1}$$

kivonást. 1.3-1-ben c az a szám, amely eleget tesz a

$$c+a=b \tag{1.3-2}$$

összefüggésnek.

Az összeadási és kivonási műveletet azonban célszerű egyetlen műveletté összefoglalni. Ekkor a szám kivonását egy negatív számmal való összeadással helyettesíthetjük. Legyen tehát

$$b+x=c, \tag{1.3-3}$$

ahol x azt a negatív számot jelenti, amelynek b-hez való hozzáadásával ugyanarra az eredményre jutunk, mint a kivonásával. Az 1.3-3-ban szereplő x-et

$$x=-a \tag{1.3-4}$$

-val is jelöljük.

Az összeadási műveletnek ez a formális kiterjesztése csak akkor célszerű, ha a negatív számokra vonatkozó műveleti szabályok azonosak a pozitív számokra vonatkozó műveleti szabályokkal.

A negatív számok definícióját célszerű a következő módon rögzíteni. Legyen $x=-a$, ha

$$x+a=0. \tag{1.3-5}$$

Tehát azt is írhatjuk, hogy

$$a+(-a)=a-a=0. \tag{1.3-6}$$

Feltételezve, hogy az összeadásra vonatkozó szabályok negatív számokra is érvényesek, látható, hogy az 1.3-5-tel definiált negatív számok bevezetésével 1.3-1-et valóban felírhatjuk az 1.3-2 formában. Hiszen

$$c=b-a=b+(-a), \tag{1.3-7}$$

tehát

$$c+a=\big(b+(-a)\big)+a. \tag{1.3-8}$$

Ebből 1.2-2 és 1.2-3 felhasználásával következik, hogy

$$c+a=b+\big((-a)+a\big)=b+0=b. \tag{1.3-9}$$

A kivonás tehát valóban helyettesíthető egy negatív számmal való összeadással.

1.4. Negatív számokat tartalmazó szorzatok

A negatív számok bevezetése az összeadási és kivonási műveletet egy általánosabb műveletté fogja össze. További hasznos műveletekhez jutunk, ha a *szorzást* kiterjesztjük negatív számokra. Azt találjuk, hogy pozitív és negatív számokat tartalmazó szorzatok úgy értelmezhetők, hogy egyrészt e szorzatok fizikai törvényeket fejezzenek ki, másrészt az általánosított szorzási eljárás eleget tegyen az 1.2-5-től 1.2-10 szabályoknak.

Az

$$a(-b)=c$$

szorzat, ahol a, $b>0$ és a egész szám, az alábbi n tag összegeként értelmezhető

$$\underbrace{(-b)+(-b)+\ldots(-b)}_{a \text{ tag}}=-ab. \tag{1.4-1}$$

A $(-a)b$ szorzat viszont csak a kommutatív törvény feltételezésével értelmezhető. Így feltesszük, hogy

$$(-a)b=b(-a)=-ba=-ab.$$

E definíció összhangban van a többi műveleti szabállyal.

Legyen

$$1+x=0, \tag{1.4-2}$$

tehát

$$x=-1.$$

Szorozzuk meg 1.4-2-t x-szel.

A szorzásra vonatkozó 1.2-3, 4, 5, 6 azonosságok alapján

$$x(1+x)=x \cdot 0=0.$$

Kihasználva a disztributivitást, kapjuk, hogy

$$x+x^2=0.$$

Az egyenlet mindkét oldalához 1-et adva:

$$1+(x+x^2)=1+0=1,$$

majd alkalmazva az asszociatív törvényt:

$$(1+x)+x^2=1.$$

Mivel kiinduló feltevésünk (1.4-2) szerint $1+x=0$, kapjuk, hogy

$$x^2=1, \tag{1.4-3}$$

vagyis

$$(-1)^2=1.$$

Hasonló módon tetszőleges $a>0$-ra

$$(1+x)a=0, \quad \text{ha} \quad x=-1,$$

tehát

$$a+xa=0,$$

és

$$xa=-a,$$

vagyis

$$-a=(-1)a. \tag{1.4-4}$$

Általában érvényes tehát, hogy tetszőleges szám (-1)-szerese egyenlő az illető szám negatívjával.

Az 1.4-4 szabály minden további nélkül alkalmazható negatív a-ra is. Ha a 1.4-4-et (-1)-gyel beszorozzuk, formálisan azt kapjuk, hogy

$$-(-a)=(-1)(-a)=(-1)^2a=a,$$

tehát

$$-(-a)=a,$$

vagyis egy negatív szám negatívja pozitív.

1.5. Több tagú összegek és az ezekből alkotott szorzat tulajdonságai

Az asszociatív és disztributív törvényt több tagú összegekből alkotott szorzatokra is kiterjeszthetjük. A következőkben a törvényekből következő néhány egyszerű összefüggést tárgyalunk.

Legyen

$$\begin{aligned} a_1+a_2+\ldots+a_k&=P_1,\\ b_1+b_2+\ldots+b_k&=P_2,\\ &\vdots\\ f_1+f_2+\ldots+f_k&=P_l, \end{aligned} \qquad 1.5\text{-}1$$

ahol $a_1, a_2, \ldots, b_1, b_2, \ldots, f_1, \ldots, f_k$ tetszőleges számok. A fenti összes szám S összegét az összeadás asszociativitása miatt rendkívül sokféleképpen számíthatjuk ki.

S meghatározása lehetséges például a következő csoportosításban:

$$S=P_1+P_2+\ldots+P_l, \qquad 1.5\text{-}2$$

vagyis a számokat soronként összeadjuk, majd összegezzük a kapott eredményeket. Megtehetjük azonban azt is, hogy először az egymás alatti számértékeket összegezzük:

$$\begin{aligned} a_1+b_1+\ldots+f_1&=Q_1,\\ a_2+b_2+\ldots+f_2&=Q_2,\\ &\vdots\\ a_k+b_k+\ldots+f_k&=Q_k. \end{aligned} \qquad 1.5\text{-}3$$

Így

$$S=Q_1+Q_2+\ldots+Q_k. \qquad 1.5\text{-}4$$

Áttekinthetőbb jelölést kapunk, ha az összegben szereplő mennyiségeket azonos betűvel jelöljük és egymástól kettős index használatával különböztetjük meg. Az első index jelenti, hogy az 1.5-1 téglalapséma hányadik sorában, a második index pedig, hogy hányadik oszlopában álló mennyiségről van szó.

Az 1.5-1 összegek helyett írhatjuk tehát, hogy

$$\begin{aligned} a_{11}+a_{12}+\ldots+a_{1k}&=P_1,\\ &\vdots \\ a_{l1}+a_{l2}+\ldots+a_{lk}&=P_l. \end{aligned} \qquad 1.5\text{-}5$$

Szumma jel felhasználásával az 1.5-1 és 1.5-3 formulákat igen tömör formában írhatjuk fel:

$$\sum_{j=1}^{k} a_{ij}=P_i, \qquad 1.5\text{-}6$$

és

$$\sum_{i=1}^{l} a_{ij}=Q_j. \qquad 1.5\text{-}7$$

1.5-6 és 1.5-7 jelölésével az S összeget az

$$S=\sum_{i=1}^{l} P_i=\sum_{i=1}^{l}\left(\sum_{j=1}^{k} a_{ij}\right) \qquad 1.5\text{-}8$$

és az

$$S=\sum_{j=1}^{k} Q_j=\sum_{j=1}^{k}\left(\sum_{i=1}^{l} a_{ij}\right) \qquad 1.5\text{-}9$$

alakban írhatjuk fel.

Tehát

$$S=\sum_{i=1}^{l}\left(\sum_{j=1}^{k} a_{ij}\right)=\sum_{j=1}^{k}\left(\sum_{i=1}^{l} a_{ij}\right), \qquad 1.5\text{-}10$$

vagyis a két index szerinti szummázás felcserélhető. Ennek kifejezésére a két index szerint történő szummázásra egyszerűbb jelölést vezethetünk be:

$$S=\sum_{i,j=1}^{l\ k} a_{ij}. \qquad 1.5\text{-}11$$

Az 1.5-11 ún. kettős szumma jól kifejezi, hogy az összegezés elvégzésekor a sorrend lényegtelen.

A disztributív törvény felhasználásával a több tagú összegek szorzatának tulajdonságait állapíthatjuk meg. Tegyük fel, hogy 1.5-5-ben az összegezendő elemek mindegyike egy A_i és B_j számsor elemeiből

$$a_{ij}=A_iB_j \qquad 1.5\text{-}12$$

szerint képzett szorzat.

Ez esetben

$$\sum_{i=1}^{l} a_{ij} = \sum_{i=1}^{l} A_i B_j.$$

A disztributivitás miatt

$$\sum_{i=1}^{l} A_i B_j = \left(\sum_{i=1}^{l} A_i \right) B_j. \qquad 1.5\text{-}13$$

1.5-13-at a második index szerint tovább összegezve,

$$\sum_{i,\, j=1}^{l,\, k} A_i B_j = \sum_{j=1}^{k} \left[\left(\sum_{i=1}^{l} A_i \right) B_j \right].$$

A disztributivitást még egyszer felhasználva,

$$\sum_{i,\, j=1}^{l\ k} A_i B_j = \left(\sum_{i=1}^{l} A_i \right) \left(\sum_{j=1}^{k} B_j \right). \qquad 1.5\text{-}14$$

Leolvashatjuk tehát, hogy több tagú összegek szorzatát egyetlen kettős szummával helyettesítjük.

Az 1.5-14 összefüggés érvényessége pusztán az összeadás asszociativitásának és a disztributív törvénynek a következménye.

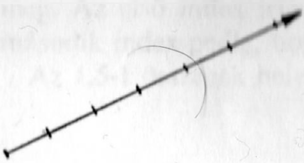
A fenti megállapítás fontos, mert a későbbiekben találkozunk olyan mennyiségekkel (pl. vektorokkal és tenzorokkal), amelyekre érvényes a fenti két törvény, de más alaptörvények (pl. a kommutatív törvény vagy a szorzás asszociativitása) nem érvényesek. Ilyen mennyiségekre az 1.5-14 szabályok érvényesek.

2. Vektorok és vektorműveletek

A skaláris mennyiségek a természetben előforduló jelenségek egy részének leírására alkalmasak, sok esetben azonban az egyetlen mérőszámmal történő jellemzés nem elegendő. Gyakran a mérőszámon kívül még egy irányt is meg kell adnunk.

Azokat a mennyiségeket, amelyek egy mérőszámmal és egy iránnyal adhatók meg, vektoroknak nevezzük. (Ilyen például a sebesség, gyorsulás, elektromos és mágneses térerősség stb.)

A vektorokat aláhúzott fett (kövér) betűkkel jelöljük. Az aláhúzás nélküli fett betűket a későbbiek során bevezetésre kerülő vektorreprezentációk jelölésére tartjuk fenn.



2.1. ábra

A vektorok szemléltetésére geometriai ábrázolásmódot szokás bevezetni. A vektorokat irányított szakaszokkal ábrázoljuk *(2.1. ábra)*. A szakasz hossza megadja az ábrázolt mennyiséghez rendelt mérőszámot (esetünkben 6,5), a nyíl pedig a mennyiség irányát jelzi.
A mérőszámot a vektor abszolút értékének nevezzük és

$$|\mathbf{a}| = a \qquad 2.1$$

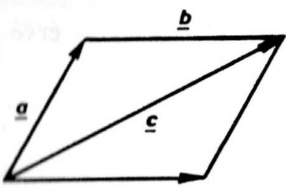
-val jelöljük. A vektorok abszolút értéke nem negatív szám. Bármely vektorra érvényes tehát az

$$|\mathbf{a}| \geqq 0$$

összefüggés. Azt a vektort, amelynek abszolút értéke zérus, nullvektornak nevezzük. Ha $|\mathbf{n}| = 0$, akkor $n = 0$-t is írhatunk. A nullvektornak nincs iránya. Az egyszerűség kedvéért azonban a nullvektort tetszőleges irányú vektornak is vehetjük.

Fizikai jelenségek közötti összefüggéseket sokszor vektorok közötti összefüggésekkel tükrözhetünk. A vektorműveletek bevezetésekor mindig az a célunk, hogy fizikai jelenségek leírására alkalmas formalizmust dolgozzunk ki.

2.1. Vektorok összegezése



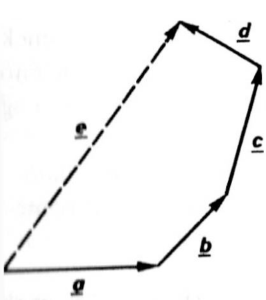
2.2. ábra

Az $\mathbf{a}$ és $\mathbf{b}$ vektorok

$$\mathbf{a} + \mathbf{b} = \mathbf{c} \qquad 2.1\text{-}1$$

összegét az ún. paralelogrammaszabály segítségével definiáljuk *(2.2. ábra)*. A 2.2. ábra alapján látható, hogy az $\mathbf{a}$ és $\mathbf{b}$ vektorok $\mathbf{c}$ összegét az $\mathbf{a}$ és $\mathbf{b}$ vektorokból alkotott paralelogrammának az $\mathbf{a}$ kezdőpontjából $\mathbf{b}$ végpontjába mutató átlója adja.

Több vektor összegét úgy képezzük, hogy az összeadás sorrendjének megfelelően a vektorokat egymás után mintegy felfűzzük. Az összegvektort az első tag kezdőpontjából az utolsó végpontjába mutató vektor adja. A *2.3. ábra* szerint



2.3. ábra

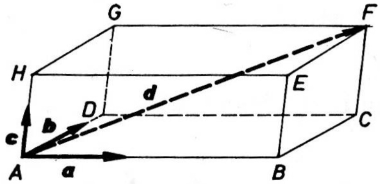
$$\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = \mathbf{e}.$$

A 2.2. ábráról leolvasható, hogy

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}. \qquad 2.1\text{-}2$$

tehát a vektorösszegzés *kommutatív*. A kommutativitás több összeadandó vektor esetén is fennáll.

Ugyancsak fennáll a vektorösszegezés *asszociativitása* is. Az asszociatív törvény érvényessége pl. három, nem egy síkba eső $\underline{a}$, $\underline{b}$, $\underline{c}$ vektor esetén egyszerűen leolvasható a *2.4. ábráról*.



2.4. ábra

A vektorösszeadást azért definiáltuk a fentiek szerint, mert ily módon sok jelenséget írhatunk le. Ezt a szabályt követik például az erők és sebességek is. Ha pl. egy $\underline{w}$ sebességgel mozgó vonathoz képest $\underline{v}$ sebességgel mozog egy test, akkor a test sebességét a földhöz képest a

$$\underline{V}=\underline{v}+\underline{w} \qquad 2.1\text{-}3$$

összeg határozza meg.

Világosan kell látnunk azonban azt, hogy az összeadási szabály csak geometriai szempontból puszta definíció. Amennyiben fizikai összefüggéseket állapítunk meg, akkor minden esetben kísérleti tapasztalat, hogy a mennyiségek hogyan összegezhetők, s hiába változtatnánk meg a vektorösszegezés szabályait, a sebességek továbbra is a régi módon viselkednének. A vektorösszeadás definícióját megváltoztatva, a 2.1-3 sebesség-összeadási törvény érvényességét veszítené.

2.2. Vektorok kivonása

Legyen

$$\underline{a}+\underline{x}=\underline{0}. \qquad 2.2\text{-}1$$

A 2.2-1 összefüggésnek eleget tevő $\underline{x}$ vektort az $\underline{a}$ vektor negatívjának nevezzük, és

$$\underline{x}=-\underline{a}$$

-val jelöljük.

A 2.2-1 definícióból következik, hogy $\underline{a}$ és $-\underline{a}$ azonos nagyságú, de ellentétes irányú vektorok.

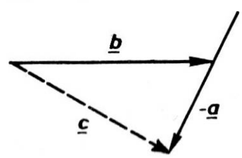
A számok kivonásának mintájára a

$$\underline{c}=\underline{b}+(-\underline{a})$$

összeget a $\underline{b}$ és $\underline{a}$ vektorok különbségeként is felfoghatjuk, és

$$\underline{c}=\underline{b}-\underline{a}$$

formában írhatjuk. A $\underline{c}$ különbségvektort úgy kapjuk, hogy a $\underline{b}$ vektor végpontjából az $\underline{a}$ vektort eredeti irányával ellentétes irányban mérjük fel *(2.5. ábra)*.



2.5. ábra

Mivel

$$\underline{a}+(\underline{b}-\underline{a})=\underline{b},$$

a $\underline{b}-\underline{a}$ olyan vektor, amit $\underline{a}$-hoz adva, $\underline{b}$-t kapunk.

2.3. Vektor szorzása számmal

Egyenlő vektorok összegét a természetes számok és vektorok szorzatának alábbi értelmezésével egyszerűbben írhatjuk fel

$$\underline{a}+\underline{a}=2\underline{a},\ \underline{a}+\underline{a}+\underline{a}=3\underline{a} \quad \ldots \text{ stb.} \qquad 2.3\text{-}1$$

A természetes számmal való szorzás tehát a vektor irányán nem változtat, csak hosszát többszörözi.

A 2.3-1 összefüggést tetszőleges szám és vektor szorzatára általánosítjuk. Legyen

$$\alpha\underline{a}=\underline{A}, \qquad 2.3\text{-}2$$

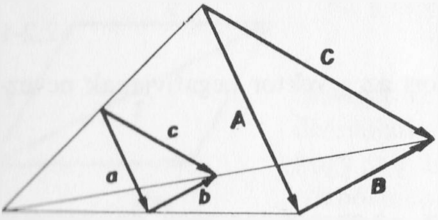
ahol

$$|\underline{A}|=|\alpha|\,|\underline{a}|, \qquad 2.3\text{-}3$$

valamint $\underline{A}$ és $\underline{a}$ iránya megegyezik, ha $\alpha>0$. Amennyiben $\alpha<0$, akkor $\underline{A}$ és $\underline{a}$ iránya ellentétes.

A 2.3-3 definícióból és az összeadás tulajdonságaiból következik, hogy a vektorok és számok szorzata disztributív az

$$\alpha(\underline{a}+\underline{b})=\alpha\underline{a}+\alpha\underline{b} \qquad 2.3\text{-}4$$



2.6. ábra

értelemben. A disztributív tulajdonság egyszerűen belátható a *2.6. ábra* alapján.

Az ábrán $\underline{a}$, $\underline{b}$, $\underline{c}$ egy háromszög oldalai, és irányításuknak megfelelően

$$\underline{a}+\underline{b}=\underline{c}. \qquad 2.3\text{-}5$$

Ha a háromszög oldalait $\alpha\neq 0$-val megszorozzuk, akkor egy hasonló háromszöghöz jutunk. (Az $\alpha<0$ esetben az új háromszöget tükrözéssel és méreteinek arányos változtatásával kapjuk meg.) Következésképpen

$$\alpha(\underline{a}+\underline{b})=\alpha\underline{a}+\alpha\underline{b}=\alpha\underline{c}. \qquad 2.3\text{-}6$$

A számmal (skalárral) való szorzás segítségével sok fizikai összefüggést kifejezhetünk.

Skalár- és vektormennyiségek szorzatával fejezhető ki például az impulzus, tömeg és sebesség közötti

$$\underline{\mathbf{p}} = m\underline{\mathbf{v}}$$

vagy az erő, gyorsulás és tömeg összefüggését meghatározó

$$\underline{\mathbf{F}} = m\underline{\mathbf{a}}$$

kapcsolat is.

2.3.1. A háromszög-egyenlőtlenség

Legyen

$$\underline{\mathbf{a}} + \underline{\mathbf{b}} = \underline{\mathbf{c}}.$$

Ekkor a három vektorból háromszög képezhető, így

$$a + b \geqq c. \qquad 2.3\text{-}7$$

A 2.3-7 formulában a dőlt betűk a megfelelő betűvel jelölt vektormennyiségek abszolút értékét jelentik. Az összefüggés következik abból, hogy egy háromszög két oldalának összege mindig nagyobb a harmadik oldalnál. Egyenlőség csak akkor jöhet létre, ha

$$\underline{\mathbf{a}} = \alpha\underline{\mathbf{b}}, \quad \text{azaz} \quad a = \alpha b.$$

Ez esetben

$$\underline{\mathbf{a}} + \underline{\mathbf{b}} = (1 + \alpha)\underline{\mathbf{b}}, \quad \text{azaz} \quad c = (1 + \alpha)b.$$

Tehát ekkor

$$a + b = c,$$

vagyis az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ vektorok ekkor egy irányúak.

2.3.2. Vektorok lineáris kombinációja

A vektor- és skalármennyiségek szorzatának értelmezésekor megállapítottuk, hogy a

$$\underline{\mathbf{b}} = \alpha\underline{\mathbf{a}} \qquad 2.3\text{-}8$$

vektorok $\underline{\mathbf{a}}$-val párhuzamosak. Ezeket a vektorokat a továbbiakban *azonos állású* vagy *egy egyenesbe eső* vektoroknak nevezzük.

Belátható, hogy minden az $\underline{\mathbf{a}} \neq \underline{\mathbf{0}}$-val egy egyenesbe eső $\underline{\mathbf{b}}$ előállítható a 2.3-8 alakban. Ugyanis az

$$\alpha = \frac{b}{a}$$

választás esetén

$$\alpha\underline{\mathbf{a}} = \frac{b}{a}\,\underline{\mathbf{a}} = \underline{\mathbf{b}}.$$

A fenti tétel megfordítása is érvényes. Az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok egy egyenesbe esnek, ha létezik olyan α, β számpár, amelyben legalább az egyik szám nem zérus, melyre

$$\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}=\underline{\mathbf{0}}, \qquad 2.3\text{-}9$$

ugyanis ha pl. $\alpha\neq 0$, akkor $\underline{\mathbf{a}}=\frac{\beta}{\alpha}\underline{\mathbf{b}}$.

Az a feltétel, hogy α, β valamelyike nem nulla, az

$$\alpha^2+\beta^2>0 \qquad 2.3\text{-}10$$

alakban is megfogalmazható.

2.3-9 és 2.3-10 szükséges és elégséges feltétele az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok párhuzamosságának. Vagyis $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ nem azonos állású vektorok, ha bármely

$$\alpha^2+\beta^2>0 \qquad 2.3\text{-}11$$

feltételt kielégítő α, β értékpárra

$$\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}\neq\underline{\mathbf{0}}. \qquad 2.3\text{-}12$$

Ez esetben az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok egy síkot feszítenek ki.

E sík vektorai előállíthatók

$$\underline{\mathbf{c}}=\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}} \qquad 2.3\text{-}13$$

alakban. A 2.3-13 alakú kifejezést az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok *lineáris kombinációjának* nevezzük.

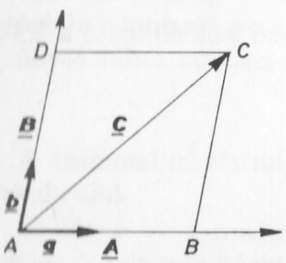
A következőkben bebizonyítjuk, hogy az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ által kifeszített síkba eső összes vektor előállítható az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok lineáris kombinációjaként.

Legyen ugyanis $\underline{\mathbf{c}}$ tetszőleges, az $\underline{\mathbf{a}}\underline{\mathbf{b}}$ síkba eső vektor. Vegyünk fel egy paralelogrammát, amelynek $\underline{\mathbf{c}}$ az egyik átlója, két oldala pedig beleesik $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ egyenesébe.

A *2.7. ábra* jelöléseivel $\overrightarrow{AB}=\underline{\mathbf{A}}$; $\overrightarrow{AD}=\underline{\mathbf{B}}$ és $\overrightarrow{AC}=\underline{\mathbf{c}}$. Tehát

$$\underline{\mathbf{c}}=\underline{\mathbf{A}}+\underline{\mathbf{B}}. \qquad 2.3\text{-}14$$

Mivel azonban $\underline{\mathbf{A}}$ és $\underline{\mathbf{a}}$, illetve $\underline{\mathbf{B}}$ és $\underline{\mathbf{b}}$ azonos állású vektorok,



2.7. ábra

$$\underline{\mathbf{A}}=\alpha\underline{\mathbf{a}}, \quad \text{és} \quad \underline{\mathbf{B}}=\beta\underline{\mathbf{b}}, \qquad 2.3\text{-}15$$

ahol

$$\alpha=\frac{A}{a}, \quad \text{és} \quad \beta=\frac{B}{b},$$

vagyis

$$\underline{\mathbf{c}}=\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}} \qquad 2.3\text{-}16$$

alakban írható — állításunknak megfelelően.

E tétel megfordításaként 2.3-16 alapján megállapítjuk, hogy három vektor, $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$, akkor és csak akkor esik egy síkba, ha létezik olyan α, β, γ számhármas, ahol

$$\alpha^2+\beta^2+\gamma^2\neq 0 \qquad 2.3\text{-}17$$

és amelyekkel

$$\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}+\gamma\underline{\mathbf{c}}=\underline{\mathbf{0}}. \qquad 2.3\text{-}18$$

A fenti állítás egyszerű átfogalmazásából következik, hogy amennyiben 2.3-17-et kielégítő tetszőleges α, β, γ számhármasra

$$\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}+\gamma\underline{\mathbf{c}}\neq\underline{\mathbf{0}}, \qquad 2.3\text{-}19$$

akkor $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ nem fekszik egy síkban.

* * *

A következőkben belátjuk, hogy a 2.3-17 és 2.3-19 feltételeknek eleget tevő, *tehát három, nem egy síkba eső* $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ *és* $\underline{\mathbf{c}}$ *vektor lineáris kombinációja segítségével tetszőleges* $\underline{\mathbf{d}}$ *vektort előállíthatunk.*

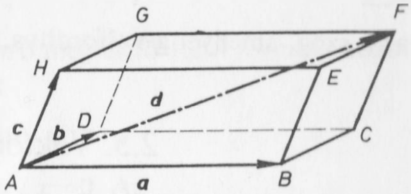
Legyen $\underline{\mathbf{d}}$ a tér egy tetszőleges vektora. Képzeljünk el egy paralelepipedont, amelynek egy csúcsba összefutó élei rendre párhuzamosak a 2.3-17 és 2.3-19-nek eleget tevő $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorokkal, testátlója pedig $\underline{\mathbf{d}}$.[1]

A *2.8. ábra* jelöléseivel legyen

$$\underline{\mathbf{A}}=\overrightarrow{AB},\quad \underline{\mathbf{B}}=\overrightarrow{AD},\quad \underline{\mathbf{C}}=\overrightarrow{AH},$$

valamint

$$\underline{\mathbf{d}}=\overrightarrow{AF}.$$



2.8. ábra

Az ábra alapján látható, hogy

$$\underline{\mathbf{A}}+\underline{\mathbf{B}}+\underline{\mathbf{C}}=\underline{\mathbf{d}}. \qquad 2.3\text{-}20$$

Mivel $\underline{\mathbf{A}}$ és $\underline{\mathbf{a}}$; $\underline{\mathbf{B}}$ és $\underline{\mathbf{b}}$, valamint $\underline{\mathbf{C}}$ és $\underline{\mathbf{c}}$ rendre azonos állású vektorok, így

$$\underline{\mathbf{A}}=\alpha\underline{\mathbf{a}};\quad \underline{\mathbf{B}}=\beta\underline{\mathbf{b}};\quad \underline{\mathbf{C}}=\gamma\underline{\mathbf{c}},$$

ahol

$$\alpha=\frac{A}{a};\quad \beta=\frac{B}{b};\quad \gamma=\frac{C}{c}.$$

[1] A fenti megfogalmazásokat egyszerűbb formára hozhatnánk, ha az eleve nulla vektorokat kihagynánk. Így pl. 2.3-18 helyett az egyszerűbb 2.3-16-ot használhatjuk, ha feltesszük, hogy $\underline{\mathbf{a}}\neq\underline{\mathbf{0}}$ és $\underline{\mathbf{b}}\neq\underline{\mathbf{0}}$. Azzal, hogy kettő, (α, β) helyett három, (α, β, γ) együtthatót használunk, olyan kifejezésekhez jutunk, amelyek akkor is érvényben maradnak, ha az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok között nulla vektorok is szerepelhetnek.

Ennek alapján 2.3-20 az

$$\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}+\gamma\underline{\mathbf{c}}=\underline{\mathbf{d}} \qquad 2.3\text{-}21$$

alakra hozható. Mivel $\underline{\mathbf{d}}$-re semmilyen megkötést sem tettünk, 2.3-21 éppen az eredeti állítást jelenti.

*

Tapasztalati tény, hogy a térben eljárásunk nem folytatható, azaz minden térbeli vektor előállítható három adott, nem komplanáris vektor lineáris kombinációjaként. Ez annak kifejeződése, hogy terünk háromdimenziós.

2.4. Vektorok által alkotott szög

Egy pontból induló két vektor által alkotott szögön a metsző irányvonalak által alkotott egyik szöget értjük. Az egyértelműséget úgy biztosítjuk, hogy az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ által bezárt szögnek azon legkisebb elfordulás szögét nevezzük, amellyel $\underline{\mathbf{a}}$-t $\underline{\mathbf{b}}$ irányába fordíthatjuk el. E definíció akkor is fenntartható, ha a vektorok irányvonala nem metszi egymást. Ebben az esetben is

$$\alpha=(\underline{\mathbf{a}}, \underline{\mathbf{b}})\sphericalangle$$

az a szög, amellyel $\underline{\mathbf{a}}$-t elfordítva, $\underline{\mathbf{b}}$-vel párhuzamossá válik.

2.5. Vektorok skaláris szorzása

A gyakorlatban találkozunk olyan skaláris mennyiségekkel, amelyek vektorokból sajátos módon származtathatók. Ilyen például a munka számértéke. Ha az F erő s elmozdulás során fejti ki hatását egy testre, akkor a munkavégzés

$$W=Fs\cos\vartheta, \qquad 2.5\text{-}1$$

ahol ϑ az erő és az elmozdulásvektor által alkotott szög.

Bevezetjük az

$$Fs\cos\vartheta=\underline{\mathbf{F}}\underline{\mathbf{s}} \qquad 2.5\text{-}2$$

jelölésmódot. Az $\underline{\mathbf{F}}\underline{\mathbf{s}}$ szorzatot az erő és elmozdulásvektor *skaláris* szorzatának nevezzük.

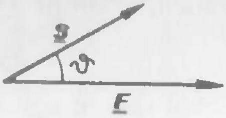
Általában az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ vektorok

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=A$$

skaláris szorzatán az

$$A=ab\cos\vartheta \qquad 2.5\text{-}3$$

mennyiséget értjük, ahol ϑ az $\mathbf{a}$, $\mathbf{b}$ vektorok által bezárt szög (*2.9. ábra*).



2.9. ábra

Az $\mathbf{a}$ vektornak önmagával vett skaláris szorzatát $\mathbf{aa}=\mathbf{a}^2$-tel jelöljük. A definíció alapján tehát

$$\mathbf{aa}=a^2.$$

A skaláris szorzás definícióját az indokolja, hogy alkalmas a valóságban előforduló összefüggések leírására.

2.5.1. A skaláris szorzat tulajdonságai

Először megvizsgáljuk, hogy a valós számok szorzására vonatkozó 1.2-(5, 6, 7, 8) műveleti szabályok közül melyek érvényesek a vektorok skaláris szorzatára.

A 2.5-3 definícióból következik, hogy

$$\mathbf{ab}=\mathbf{ba}, \qquad 2.5\text{-}4$$

tehát a skaláris szorzás *kommutatív*. Tetszőleges vektornak az $\mathbf{n}$ nullvektorral képezett skaláris szorzata zérus:

$$\mathbf{an}=0. \qquad 2.5\text{-}5$$

A 2.5-5 egyenlőség és a számokra vonatkozó megfelelő 1.2-8 formula között egy fontos különbség van: számok esetén az

$$ax=0 \quad (a\neq 0)$$

összefüggésből egyértelműen következik, hogy $x=0$. Az

$$\mathbf{ax}=0 \quad (\mathbf{a}\neq\mathbf{0}) \qquad 2.5\text{-}6$$

vektoregyenletből azonban nem következik az $\mathbf{x}=\mathbf{0}$ eredmény.

Definíció szerint ugyanis

$$\mathbf{ax}=ax\cos\vartheta. \qquad 2.5\text{-}7$$

2.5-6 teljesülhet $\mathbf{a}\neq\mathbf{0}$ és $\mathbf{x}\neq\mathbf{0}$ esetén is, ha

$$\cos\vartheta=0, \quad \text{azaz} \quad \vartheta=90°,$$

vagyis ha az $\mathbf{a}$ és $\mathbf{x}$ vektorok merőlegesek egymásra.

Az $\mathbf{x}$ vektor zérus voltára csak akkor következtethetünk, ha az

$$\mathbf{ax}=0 \qquad 2.5\text{-}8$$

egyenlőség tetszőleges $\mathbf{a}$-ra fennáll. Valóban, ha 2.5-8 tetszőleges $\mathbf{a}$ vektorra fennáll, akkor az $\underline{\mathbf{a}}=\underline{\mathbf{x}}$-re is teljesül. Tehát

$$\underline{\mathbf{x}}^2=0.$$

Ebből viszont következik, hogy

$$\underline{\mathbf{x}}=\underline{\mathbf{0}}.$$

A fenti tétel egy fontos alkalmazása a következő: Legyen

$$\underline{\mathbf{a}}\underline{\mathbf{x}}=\underline{\mathbf{a}}\underline{\mathbf{y}} \qquad 2.5\text{-}9$$

tetszőleges $\underline{\mathbf{a}}$-ra, akkor

$$\underline{\mathbf{a}}(\underline{\mathbf{x}}-\underline{\mathbf{y}})=0$$

minden $\underline{\mathbf{a}}$-ra, s így

$$\underline{\mathbf{x}}-\underline{\mathbf{y}}=\underline{\mathbf{0}},$$

vagyis

$$\underline{\mathbf{x}}=\underline{\mathbf{y}}. \qquad 2.5\text{-}10$$

Utóbbi eredményünk fontos szerepet játszik a vektoregyenletek megoldásakor.

A skalárszorzás eredménye nem vektormennyiség — a művelet kivezet a vektorok köréből —, nincs tehát olyan vektor, amellyel szorozva, tetszőleges vektor változatlan marad.

A vektorok skaláris szorzatával kapcsolatban az eredeti értelemben vett egységről nem beszélhetünk. Célszerű azonban tetszőleges irányú vektorokhoz egységvektorokat bevezetni a következő módon.

Az $\underline{\mathbf{a}}\neq\underline{\mathbf{0}}$ vektorhoz tartozó egységvektor

$$\underline{\mathbf{k}}=\frac{\underline{\mathbf{a}}}{a}. \qquad 2.5\text{-}11$$

2.5-11 alapján látható, hogy $\underline{\mathbf{k}}$ és $\underline{\mathbf{a}}$ azonos irányú vektorok, valamint hogy $k=1$. (Innen az egységvektor elnevezés.)

A $\underline{\mathbf{k}}$ vektor segítségével $\underline{\mathbf{a}}$-t az

$$\underline{\mathbf{a}}=a\underline{\mathbf{k}}$$

formában fejezhetjük ki.

Ennek segítségével az $\underline{\mathbf{a}}\underline{\mathbf{b}}=ab\cos\vartheta$ skalárszorzat $b\cos\vartheta$ tényezőjéről felismerhetjük, hogy

$$\underline{\mathbf{k}}\underline{\mathbf{b}}=b\cos\vartheta. \qquad 2.5\text{-}12$$

A 2.5-12 szorzatot a $\underline{\mathbf{b}}$ *vektor* $\underline{\mathbf{a}}$*-ra vett vetületének nevezzük, és*

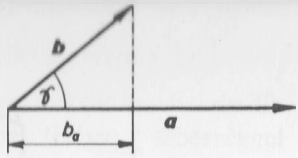
$$\underline{\mathbf{k}}\underline{\mathbf{b}}=b\cos\vartheta=b_a \qquad 2.5\text{-}13$$

-val jelöljük.

A vetületet vektorként is kezelhetjük, definíció szerint

$$\underline{\mathbf{b}}_a=\underline{\mathbf{k}}b_a, \qquad 2.5\text{-}14$$

vagyis $\underline{\mathbf{b}}_a$ az $\underline{\mathbf{a}}$ irányába eső b_a abszolút értékű vektor. A vetületvektor geometriai jelentését a *2.10. ábra* mutatja.



2.10. ábra

Megállapíthatjuk azt is, hogy

$$-b \leq b_a \leq b. \qquad 2.5\text{-}15$$

A definícióból a $ba_b = ab_a$ összefüggés is következik, ahol a_b az $\underline{\mathbf{a}}$ vektor $\underline{\mathbf{b}}$-re vett vetülete.

Az asszociatív törvény nem érvényes a skalárszorzatra. Három vektor szorzatát is csak úgy értelmezhetjük, ha megadjuk a szorzások sorrendjét. Az $(\underline{\mathbf{a}}\underline{\mathbf{b}})\underline{\mathbf{c}}$ szorzat például $\underline{\mathbf{c}}$ irányú vektor, az $\underline{\mathbf{a}}(\underline{\mathbf{b}}\underline{\mathbf{c}})$ viszont $\underline{\mathbf{a}}$ irányú vektor. Így a kétféle szorzat csak egészen különleges esetekben egyenlő.

A vektorösszeadásra és skaláris szorzásra érvényes az

$$(\underline{\mathbf{a}}+\underline{\mathbf{b}})\underline{\mathbf{c}} = \underline{\mathbf{a}}\underline{\mathbf{c}} + \underline{\mathbf{b}}\underline{\mathbf{c}} \qquad 2.5\text{-}16$$

összefüggéssel kifejezhető disztributív törvény.

A bizonyítás elvégzéséhez vezessük be a következő jelöléseket:

$$\underline{\mathbf{a}}+\underline{\mathbf{b}} = \underline{\mathbf{A}}, \quad \text{és} \quad (\underline{\mathbf{a}}+\underline{\mathbf{b}})\underline{\mathbf{c}} = Q. \qquad 2.5\text{-}17$$

Egyszerűség kedvéért foglalkozzunk először azzal az esettel, amikor $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ komplanáris vektorok. A *2.11. ábra* alapján egyszerűen belátható, hogy

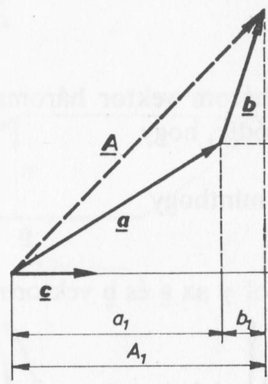
$$Q = A_1 c, \qquad 2.5\text{-}18$$

ahol A_1 az $\underline{\mathbf{A}}$ vektor $\underline{\mathbf{c}}$-re vett vetülete, és

$$A_1 = a_1 + b_1.$$

Tehát

$$Q = (a_1 + b_1)c = a_1 c + b_1 c. \qquad 2.5\text{-}19$$



2.11. ábra

Az a_1 és b_1 vetületek segítségével

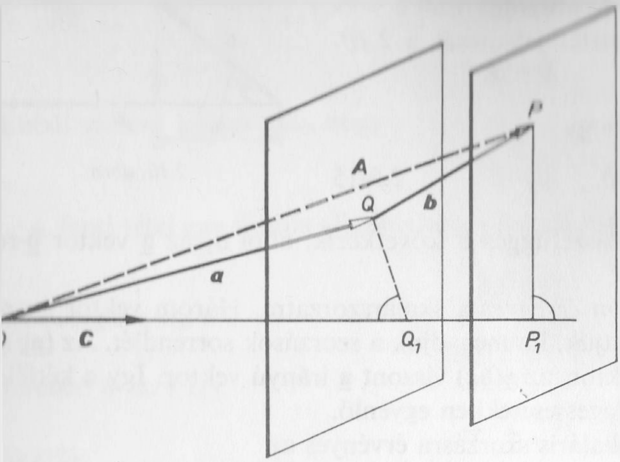
$$\underline{\mathbf{a}}\underline{\mathbf{c}} = a_1 c, \quad \underline{\mathbf{b}}\underline{\mathbf{c}} = b_1 c. \qquad 2.5\text{-}20$$

Következésképpen

$$Q = (\underline{\mathbf{a}}+\underline{\mathbf{b}})\underline{\mathbf{c}} = \underline{\mathbf{a}}\underline{\mathbf{c}} + \underline{\mathbf{b}}\underline{\mathbf{c}}, \qquad 2.5\text{-}21$$

ami éppen a szorzat disztributivitását jelenti.

2.5-16 azonban akkor is helytálló, ha $\underline{\mathbf{c}}$ nincs az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok síkjában. Szerkesszünk $\underline{\mathbf{c}}$-re két merőleges síkot, amelyek egyike az $\underline{\mathbf{A}} = \underline{\mathbf{a}} + \underline{\mathbf{b}}$ vektorok



2.12. ábra

P végpontján, a másik az $\underline{\mathbf{a}}$ vektor Q végpontján megy át (*2.12. ábra*). Legyen a síkok és a $\underline{\mathbf{c}}$ vektor egyenesének döféspontja P_1 és Q. Leolvasható, hogy, ha A_1 az $\underline{\mathbf{A}}$ vektor $\underline{\mathbf{c}}$-re vett vetülete és

$$a_1 = OQ_1 \quad \text{és} \quad b_1 = Q_1P_1,$$

akkor

$$A_1 = a_1 + b_1.$$

Ezután a síkbeli bizonyítás változtatás nélkül megismételhető a térbeli esetre.

2.5.2. Alkalmazás. (A cosinustétel)

A skaláris szorzás felhasználásával egyszerűen adódik a háromszögekre vonatkozó cosinustétel. Legyen $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ három vektor úgy, hogy

$$\underline{\mathbf{b}} - \underline{\mathbf{a}} = \underline{\mathbf{c}}; \tag{2.5-22}$$

a három vektor háromszöget képez. A fenti összefüggést négyzetre emelve, adódik, hogy

$$\underline{\mathbf{c}}^2 = \underline{\mathbf{a}}^2 + \underline{\mathbf{b}}^2 - 2\underline{\mathbf{a}}\underline{\mathbf{b}},$$

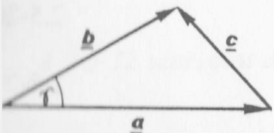
és minthogy

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = ab \cos \gamma,$$

ahol γ az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok által bezárt szög, azt kapjuk, hogy

$$c^2 = a^2 + b^2 - 2ab \cos \gamma. \tag{2.5-23}$$

Ez a háromszögre vonatkozó cosinustétel. Továbbá 2.5-23-ból következik, hogy[1]



2.13. ábra

$$\cos \gamma = \frac{a^2 + b^2 - c^2}{2ab}. \tag{2.5-24}$$

[1] A gondolatmenetet megfordítva, a fenti egyenletet a szög különleges definíciójának is felfoghatjuk.

2.6. A vektoriális szorzat

Találkozunk olyan fizikai törvényekkel, amelyekben két vektor egy harmadik vektort sajátságos módon határoz meg. Így a $\underline{\mathbf{B}}$ mágneses térben $\underline{\mathbf{v}}$ sebességgel mozgó e töltésre ható erőt a következő formula alapján határozhatjuk meg:

$$F=\frac{e}{c}\,vB\sin\vartheta. \qquad 2.6\text{-}1$$

A formulában ϑ a $\underline{\mathbf{v}}$ és $\underline{\mathbf{B}}$ vektorok által alkotott szög. A képlet alapján csak az erő nagysága adódik. Az $\underline{\mathbf{F}}$ erő iránya tapasztalat szerint merőleges a $\underline{\mathbf{v}}$ és $\underline{\mathbf{B}}$ által kifeszített síkra és a $\underline{\mathbf{v}}$, $\underline{\mathbf{B}}$ és $\underline{\mathbf{F}}$ vektorok jobbcsavart alkotnak. A 2.6-1 formulát szokás az

$$\underline{\mathbf{F}}=\frac{e}{c}(\underline{\mathbf{v}}\times\underline{\mathbf{B}}) \qquad 2.6\text{-}2$$

alakban is felírni, ahol

$$\underline{\mathbf{v}}\times\underline{\mathbf{B}}$$

a fentieknek megfelelő irányú és

$$|\underline{\mathbf{v}}\times\underline{\mathbf{B}}|=vB\sin\vartheta$$

abszolút értékű vektort jelent.

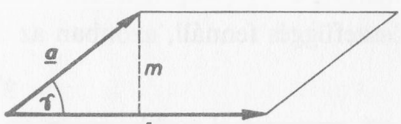
$\underline{\mathbf{v}}\times\underline{\mathbf{B}}$-t a $\underline{\mathbf{v}}$ és $\underline{\mathbf{B}}$ vektorok *vektoriális szorzatának* nevezzük. Használatos még a külső vagy keresztszorzat elnevezés és a $[\underline{\mathbf{v}}\underline{\mathbf{B}}]$ jelölés is.

Általában az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok vektoriális szorzatán a

$$\underline{\mathbf{c}}=\underline{\mathbf{a}}\times\underline{\mathbf{b}} \qquad 2.6\text{-}3$$

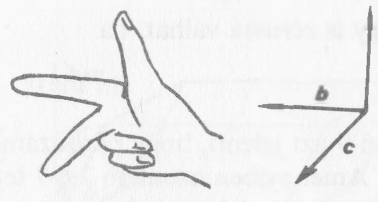
vektort értjük, ahol

$$c=ab\sin\gamma \qquad 2.6\text{-}4$$



2.14. ábra

(γ az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok által bezárt szög). A *2.14. ábra* alapján megállapíthatjuk, hogy a $\underline{\mathbf{c}}$ vektor nagysága éppen az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok által alkotott paralelogramma területének számértéke.



2.15. ábra

A $\underline{\mathbf{c}}$ vektor irányát az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok síkjára merőlegesnek vesszük úgy, hogy $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ az adott sorrendben jobbrendszert alkot[1] *(2.15. ábra)*.

[1] Jobbrendszeren a jobb kezünk hüvelyk-, mutató és középső ujjához hasonló sorrendű rendszereket értjük (ábra). Azt, hogy az adott sorrendű $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok jobbrendszert alkotnak-e, könnyen ellenőrizhetjük úgy is, hogy $\underline{\mathbf{c}}$ helyébe egy „jobbcsavart" képzelünk. Képzeletben forgassuk az $\underline{\mathbf{a}}$ vektort $\underline{\mathbf{b}}$ felé. Amennyiben az $\underline{\mathbf{a}}$ helyére képzelt csavar erre a forgatásra $\underline{\mathbf{c}}$ irányába mozdulna, akkor $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ jobbrendszert alkot.

Megjegyezzük, hogy a vektoriális szorzatot úgy is definiálhatnánk, hogy a 2.6-3 képletben az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok balcsavart alkothassanak. E konvenció éppen úgy alkalmas lenne fizikai összefüggések leírására, mint az általunk megadott konvenciók. Ha a konvenciót változtatjuk, a törvények leírásában pusztán előjelcserék jönnek létre. Fontos azonban, hogy egy konvencióhoz ragaszkodjunk, mert egyébként előjelzavarok lépnek fel.

A vektoriális szorzat definíciója ismét nem matematikai szükségszerűség következménye. A 2.6-1 Lorentz-törvény és más hasonló szerkezetű fizikai törvények leírása teszi hasznossá a vektorok közötti, vektorértékű eredményt adó újabb szorzási művelet bevezetését.

2.6.1. A vektoriális szorzat tulajdonságai

Röviden áttekintjük a vektoriális szorzás tulajdonságait. Vezérfonalként célszerű ismét a számok szorzására vonatkozó 1.2-(5, 6, 7, 8) szabályokkal történő összehasonlítás.

A vektoriális szorzás *nem kommutatív*. Az irány meghatározásából következik, hogy

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=-\underline{\mathbf{b}}\times\underline{\mathbf{a}}, \qquad 2.6\text{-}5$$

$\underline{\mathbf{b}}$-t ugyanis éppen ellenkező irányú forgatással vihetjük át $\underline{\mathbf{a}}$-ba, mint $\underline{\mathbf{a}}$-t $\underline{\mathbf{b}}$-be. Az ilyen típusú szorzási műveletet *antikommutatívnak* szokás nevezni.

Az

$$\underline{\mathbf{a}}\times\underline{\mathbf{0}}=\underline{\mathbf{0}} \qquad 2.6\text{-}6$$

összefüggés fennáll, azonban az

$$\underline{\mathbf{a}}\times\underline{\mathbf{y}}=\underline{\mathbf{0}} \quad (a\neq 0) \qquad 2.6\text{-}7$$

egyenlőségből nem következik, hogy $\underline{\mathbf{y}}=\underline{\mathbf{0}}$, hiszen

$$|\underline{\mathbf{a}}\times\underline{\mathbf{y}}|=ay\sin\vartheta=0 \qquad 2.6\text{-}8$$

úgy is zérussá válhat, ha

$$\sin\vartheta=0, \quad \text{azaz} \quad \vartheta=0 \quad \text{vagy} \quad \pi. \qquad 2.6\text{-}9$$

2.6-9 azt jelenti, hogy párhuzamos vektorok keresztszorzata zérus.

Amennyiben azonban 2.6-7 tetszőleges $\underline{\mathbf{a}}\neq\underline{\mathbf{0}}$ esetén fennáll, akkor ebből már következik, hogy $\underline{\mathbf{y}}=\underline{\mathbf{0}}$.

Megállapíthatjuk azt is, hogy a vektoriális szorzás műveletére nem létezik olyan vektor, amellyel tetszőleges vektort szorozva, a szorzandó nem változik. (Nem létezik tehát az egységszámhoz hasonlóan viselkedő vektor.) Ez következménye annak, hogy a vektoriális szorzás eredménye mindkét tényezőre merőleges.

Három vektor

$$\underline{\mathbf{A}}=\underline{\mathbf{a}}\times(\underline{\mathbf{b}}\times\underline{\mathbf{c}}) \qquad 2.6\text{-}10$$

hármas vektoriális szorzata ismét vektort ad eredményül. A hármas vektorszorzatra az *asszociatív törvény nem* érvényes. A művelet tulajdonságainak vizsgálatára még visszatérünk (4.7. fejezet).

Az összeadásra és vektoriális szorzásra vonatkozó *disztributivitás* az

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}+\underline{\mathbf{c}})=\underline{\mathbf{a}}\times\underline{\mathbf{b}}+\underline{\mathbf{a}}\times\underline{\mathbf{c}} \qquad 2.6\text{-}11$$

értelemben fennáll, bizonyítását azonban csak később adjuk meg.

2.7. A hármas vegyes szorzat

Több vektor összeszorzásakor előfordulhat az alábbi típusú, ún. hármas vegyes szorzat:

$$\underline{\mathbf{a}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=v. \qquad 2.7\text{-}1$$

2.7-1 szerint a $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok vektoriális szorzatát kell az $\underline{\mathbf{a}}$ vektorral skalárisan szorozni.

Az eddigi definíciók alapján ezek a műveletek rendre elvégezhetők. Az eredménynek azonban egyszerű geometriai jelentése is van.

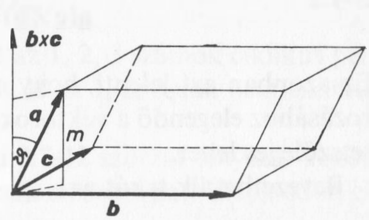
Ha az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok egy síkban feküsznek, akkor bármely kettő vektoriális szorzata a síkra és így a harmadik vektorra is merőleges. Így a harmadik vektorral való skaláris szorzat nullát ad. Tehát

$$\underline{\mathbf{a}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=0,$$

ha $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ komplanáris vektorok. (A fenti eredmény akkor is érvényes, ha a vektorok között nulla vektor is előfordul.)

Legyenek $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ nem komplanáris vektorok. Ezek a térben egy paralelepipedont feszítenek ki *(2.16. ábra)*.

A $\underline{\mathbf{b}}\times\underline{\mathbf{c}}$ vektoriális szorzat abszolút értéke a $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok által kifeszített paralelogramma területének számértékével egyenlő. Az ábra szerint pedig az $\underline{\mathbf{a}}$ vektor $\underline{\mathbf{b}}\times\underline{\mathbf{c}}$ irányú vetülete éppen a paralelepipedon magasságával egyenlő, tehát



2.16. ábra

$$m=a\cos\vartheta. \qquad 2.7\text{-}2$$

A paralelepipedon alapterülete

$$T=|\underline{\mathbf{b}}\times\underline{\mathbf{c}}|. \qquad 2.7\text{-}3$$

A térfogat

$$V = |\underline{\mathbf{b}} \times \underline{\mathbf{c}}|\, a \cos \theta, \qquad 2.7\text{-}4$$

2.7-4 éppen az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok vegyes szorzatának abszolút értéke.

Általában azt írhatjuk tehát, hogy

$$|\underline{\mathbf{a}}(\underline{\mathbf{b}} \times \underline{\mathbf{c}})| = |v| = V. \qquad 2.7\text{-}5$$

Számítsuk ki a paralelepipedon térfogatát most úgy, hogy az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok által meghatározott paralelogrammát tekintjük alapnak. Természetesen a térfogat ugyanannyi marad, tehát

$$|\underline{\mathbf{a}} \times \underline{\mathbf{b}}|\, c \cos \gamma = V, \qquad 2.7\text{-}6$$

γ most az $\underline{\mathbf{a}} \times \underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok által bezárt szög. 2.7-6 azonban az

$$|(\underline{\mathbf{a}} \times \underline{\mathbf{b}})\underline{\mathbf{c}}| = V \qquad 2.7\text{-}7$$

szorzat abszolút értéke.

Hasonló eredményt kapunk, ha az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorokból tetszőleges sorrendű hármasszorzatokat képezünk. Megállapítható tehát, hogy három vektor vegyes szorzatának abszolút értéke a vektorok tetszőleges sorrendje esetén azonos.

Ha az előjeleket is figyelembe vesszük, a hármas vegyes szorzat értéke attól függően pozitív vagy negatív, hogy az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok az adott sorrendben jobbrendszert alkotnak, vagy sem.

Ha $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ jobbrendszert alkot, akkor $\underline{\mathbf{a}}$ 180°-nál kisebb szöget zár be $\underline{\mathbf{b}} \times \underline{\mathbf{c}}$ irányával, tehát v pozitív, vagyis a vegyes szorzat a test térfogatát adja.

Tehát ha $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ jobbrendszert alkot, akkor

$$\underline{\mathbf{a}}(\underline{\mathbf{b}} \times \underline{\mathbf{c}}) = \underline{\mathbf{c}}(\underline{\mathbf{a}} \times \underline{\mathbf{b}}) = \underline{\mathbf{b}}(\underline{\mathbf{c}} \times \underline{\mathbf{a}}) = V,$$

de

$$\underline{\mathbf{a}}(\underline{\mathbf{c}} \times \underline{\mathbf{b}}) = \underline{\mathbf{c}}(\underline{\mathbf{b}} \times \underline{\mathbf{a}}) = \underline{\mathbf{b}}(\underline{\mathbf{a}} \times \underline{\mathbf{c}}) = -V. \qquad 2.7\text{-}8$$

Ez azonban azt jelenti, hogy a vegyes szorzat értékének egyértelmű meghatározásához elegendő a vektorok sorrendjét megszabni, a műveleti jelek sorrendje tetszőleges lehet.

Bevezethetjük tehát az

$$\underline{\mathbf{a}}(\underline{\mathbf{b}} \times \underline{\mathbf{c}}) = (\mathbf{a}, \underline{\mathbf{b}}, \underline{\mathbf{c}}) \qquad 2.7\text{-}9$$

rövidebb jelölést.

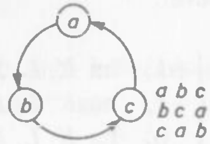
A fenti eredmények komplanáris vektorokra is érvényesek, ahol a vegyes szorzat eltűnik.

2.7.1. Ciklikus permutáció

2.7-8 alapján látható, hogy az $\underline{a}$, $\underline{b}$, $\underline{c}$ vektorok hármas vegyes szorzatának előjele a három vektor sorrendjétől függ. A sorrend könnyebb áttekintésére bevezetjük a ciklikus permutáció fogalmát.

Írjuk fel az a, b, c mennyiségeket egy kör mentén *(2.17. ábra)*, valamely — például az óramutató járásával ellentétes — irányban haladva. Az a, b, c sorrendhez tartozó ciklikus permutációnak nevezzük a három elem olyan permutációit, amelyekben tetszőleges helyről indulva, a megadott irányban haladva soroljuk fel az elemeket. a, b, c ciklikus permutációi tehát:

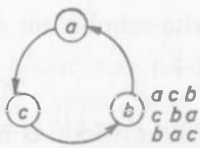
a-ból indulva a, b, c,
b-ből indulva b, c, a,
c-ből indulva c, a, b.



2.17. ábra

Fordított sorrendben haladva, az a, b, c sorrendhez tartozó nem ciklikus permutációkhoz jutunk. Ezek:

a-val kezdve a, c, b,
b-vel kezdve b, a, c,
c-vel kezdve c, b, a.



2.17. a ábra

Az így kapott permutációk pl. az a, c, b permutáció ciklikus permutációiként kaphatók meg.

A ciklikus permutációk felhasználásával egyszerűbben felírhatjuk 2.1-7 eredményeket.

Legyen

$$\underline{a}=\underline{a}^{(1)},\quad \underline{b}=\underline{a}^{(2)},\quad \underline{c}=\underline{a}^{(3)} \qquad 2.7\text{-}11$$

három jobbrendszert alkotó vektor. Így

$$\underline{a}^{(K)}(\underline{a}^{(L)}\times\underline{a}^{(M)})=\pm V, \qquad 2.7\text{-}12$$

ahol a pozitív előjel akkor érvényes, ha K, L, M az 1, 2, 3 számok ciklikus permutációja, a negatív előjel pedig akkor, ha K, L, M az 1, 2, 3-nak nem ciklikus permutációja.

Amennyiben a 2.7-11 vektorok valamelyike a 2.7-12 szorzat tényezői között kétszer szerepel, akkor a három vektor által meghatározott paralelepipedon elfajuló, térfogata s így a 2.7-12 szorzat értéke zérus.

Például ez az eset fordul elő, ha $K=L=1$ és $M=2$. Ebben az esetben azonban K, L, M az 1, 2, 3 számoknak nem permutációja!

Általában, ha 2.7-12-ben K, L, M az 1, 2, 3 számoknak nem permutációja, akkor a vegyes szorzat értéke zérus.

2.7.2. A Levi—Civita-szimbólum

Összefoglalva a fentieket:

$$\underline{\mathbf{a}}^{(K)}(\underline{\mathbf{a}}^{(L)}\times\underline{\mathbf{a}}^{(M)})=\begin{cases}+|V|, & \text{ha } K, L, M \text{ az } 1, 2, 3 \text{ ciklikus permutációja,}\\ -|V|, & \text{ha } K, L, M \text{ az } 1, 2, 3 \text{ nem ciklikus permutációja,}\\ 0, & \text{ha } K, L, M \text{ az } 1, 2, 3 \text{ értékeknek nem permutációja.}\end{cases} \qquad 2.7\text{-}13$$

A 2.7-13 összefüggés egyszerűbben felírható az ε_{KLM} ún. Levi—Civita-szimbólum bevezetésével.

$$\varepsilon_{KLM}=\begin{cases}+1, & \text{ha } K, L, M \text{ az } 1, 2, 3 \text{ ciklikus permutációja,}\\ -1, & \text{ha } K, L, M \text{ az } 1, 2, 3 \text{ nem ciklikus permutációja,}\\ 0, & \text{ha } K, L, M \text{ az } 1, 2, 3\text{-nak nem permutációja.}\end{cases} \qquad 2.7\text{-}14$$

A Levi—Civita-szimbólum segítségével a hármas vegyes szorzat értékét az

$$\varepsilon_{KLM}\underline{\mathbf{a}}^{(K)}\underline{\mathbf{a}}^{(L)}\underline{\mathbf{a}}^{(M)}=V \qquad 2.7\text{-}15$$

formula adja, ahol V az $\underline{\mathbf{a}}^{(j)}$ ($j=1, 2, 3$) vektorok által meghatározott paralelepipedon köbtartalmát jelenti.

2.7.3. A vektoriális szorzat disztributivitása

A vektoriális szorzatra vonatkozó

$$(\underline{\mathbf{a}}+\underline{\mathbf{b}})\times\underline{\mathbf{c}}=\underline{\mathbf{a}}\times\underline{\mathbf{c}}+\underline{\mathbf{b}}\times\underline{\mathbf{c}} \qquad 2.7\text{-}16$$

disztributív törvényt a skaláris szorzatra már kimutatott disztributivitás segítségével bizonyítjuk.

Első lépésként belátjuk, hogy a hármas vegyes szorzat mindhárom tényezőjében disztributív.[1]

Minthogy

$$(\underline{\mathbf{a}}+\underline{\mathbf{b}})\underline{\mathbf{x}}=\underline{\mathbf{a}}\underline{\mathbf{x}}+\underline{\mathbf{b}}\underline{\mathbf{x}}, \qquad 2.7\text{-}17$$

az $\underline{\mathbf{a}}+\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$, $\underline{\mathbf{x}}$ vektorok hármas vegyes szorzatára fennáll az

$$(\underline{\mathbf{a}}+\underline{\mathbf{b}}, \underline{\mathbf{c}}, \underline{\mathbf{x}})=(\underline{\mathbf{a}}+\underline{\mathbf{b}})(\underline{\mathbf{c}}\times\underline{\mathbf{x}})=(\underline{\mathbf{a}}, \underline{\mathbf{c}}, \underline{\mathbf{x}})+(\underline{\mathbf{b}}, \underline{\mathbf{c}}, \underline{\mathbf{x}}) \qquad 2.7\text{-}18$$

[1] Ezt az eljárást Katus Gábor első éves fizikushallgató javasolta.

összefüggés. A hármas vegyes szorzat értéke nem változik, ha tényezőit ciklikusan permutáljuk, így 2.7-18 az

$$(\underline{x}, (\underline{a}+\underline{b}), \underline{c}) = (\underline{x}, \underline{a}, \underline{c}) + (\underline{x}, \underline{b}, \underline{c}) \qquad 2.7\text{-}19$$

alakban is felírható.

Ez viszont azt jelenti, hogy

$$\underline{x}((\underline{a}+\underline{b})\times\underline{c}) = \underline{x}(\underline{a}\times\underline{c}+\underline{b}\times\underline{c}). \qquad 2.7\text{-}20$$

2.7-20 tetszőleges $\underline{x}$-re érvényes, tehát

$$(\underline{a}+\underline{b})\times\underline{c} = \underline{a}\times\underline{c}+\underline{b}\times\underline{c}, \qquad 2.7\text{-}21$$

vagyis a vektoriális szorzat disztributív.

Az asszociatív és disztributív törvények több tagú kifejezésekből alkotott szorzatokra vonatkozó alkalmazásából következik, hogy

$$\left(\sum_k \underline{a}^{(k)}\right)\times\left(\sum_l \underline{b}^{(l)}\right) = \sum_{k,l}\left(\underline{a}^{(k)}\times\underline{b}^{(l)}\right). \qquad 2.7\text{-}22$$

A 2.7-22 tökéletes megfelelője a számok esetén bizonyított 1.4-14 szabálynak, vagyis hogy két több tagú kifejezés szorzatát úgy kapjuk, hogy egyik többtag minden tagját a másik többtag minden tagjával megszorozzuk.

3. A derékszögű koordináta-rendszerek

A vektorokat nagyságukkal és irányukkal jellemezzük. A vektorok irányát általában valamilyen rögzített irányokhoz viszonyítva lehet megadni. Egyértelműen jellemezhető egy vektor iránya a Földön például a vízszintes síkkal alkotott szöggel, valamint a vektor vízszintes síkra vett vetülete és pl. az északi irány által alkotott szöggel.

A következőkben a vektorok jellemzésére más módszert választunk. A 2.3-2 fejezetben láttuk, hogy a tér tetszőleges $\underline{x}$ vektora előállítható

$$\underline{x} = \alpha\underline{a} + \beta\underline{b} + \gamma\underline{c} \qquad 3.1$$

alakban, tehát az $\underline{a}$, $\underline{b}$, $\underline{c}$ nem komplanáris vektorok lineáris kombinációjaként. Az α, β, γ számok egyértelműen jellemzik az $\underline{x}$ vektort. Az $\underline{a}$, $\underline{b}$, $\underline{c}$ nem komplanáris alapvektorok lerögzítése után tehát minden vektor egyértelműen jellemezhető három szám segítségével.

Az $\underline{a}$, $\underline{b}$, $\underline{c}$ alapvektorok egy $\mathcal{K}$ koordináta-rendszert feszítenek ki. Adjuk meg ugyanis a rendszer origóját, és vegyük fel a tengelyeket az $\underline{a}$, $\underline{b}$, $\underline{c}$ vektorok irányában. A két-két alapvektor által kifeszített síkokat a $\mathcal{K}$ rendszer koordinátasíkjainak, az α, β, γ számokat pedig az $\underline{\mathbf{x}}$ vektor $\mathcal{K}$-beli koordinátáinak nevezzük.

Különösen egyszerű helyzetet állítunk elő, ha alapvektorként három, páronként egymásra merőleges egységvektort választunk.

Megállapodás szerint tehát

$$e^{(1)}=e^{(2)}=e^{(3)}=1, \tag{3.2}$$

azaz

$$\underline{\mathbf{e}}^{(1)^2}=\underline{\mathbf{e}}^{(2)^2}=\underline{\mathbf{e}}^{(3)^2}=1. \tag{3.3}$$

Az egységvektorok merőlegessége miatt

$$\underline{\mathbf{e}}^{(1)}\underline{\mathbf{e}}^{(2)}=\underline{\mathbf{e}}^{(2)}\underline{\mathbf{e}}^{(3)}=\underline{\mathbf{e}}^{(3)}\underline{\mathbf{e}}^{(1)}=0. \tag{3.4}$$

3.1. A Kronecker-szimbólum

A 3.3, 3.4 formulákat a következő módon foglalhatjuk össze:

$$\underline{\mathbf{e}}^{(i)}\underline{\mathbf{e}}^{(k)}=\begin{cases}1, & \text{ha} \quad i=k,\\ 0, & \text{ha} \quad i\neq k\end{cases} \quad (i,\ k=1,\ 2,\ 3). \tag{3.1-1}$$

Használatos az

$$\underline{\mathbf{e}}^{(i)}\underline{\mathbf{e}}^{(k)}=\delta_{ik} \quad (i,\ k=1,\ 2,\ 3) \tag{3.1-2}$$

jelölés is, ahol δ_{ik} a Kronecker-szimbólum, melynek definíciója:

$$\delta_{ik}=\begin{cases}1, & \text{ha} \quad i=k,\\ 0, & \text{ha} \quad i\neq k.\end{cases} \tag{3.1-3}$$

A 3.1-3 definíció általános, nemcsak az i, $k=1$, 2, 3, hanem tetszőleges i, k értékekre értelmezi a Kronecker-deltát.

3.2. Ortogonális koordináták

A szemléletből következik, hogy az $\underline{\mathbf{e}}^{(1)}$, $\underline{\mathbf{e}}^{(2)}$, $\underline{\mathbf{e}}^{(3)}$ vektorok nincsenek egy síkban. Egyszerűen bizonyíthatjuk, hogy a 2.3-19 kritériumnak is eleget tesznek.

Legyen

$$\underline{\mathbf{x}}=\alpha\underline{\mathbf{e}}^{(1)}+\beta\underline{\mathbf{e}}^{(2)}+\gamma\underline{\mathbf{e}}^{(3)}. \tag{3.2-1}$$

Emeljük négyzetre az egyenletet. A skaláris szorzásra érvényes a disztributív törvény, így a szorzást tagonként végezhetjük. 3.1-1 felhasználásával ez az

$$\underline{x}^2 = \alpha^2(\underline{e}^{(1)})^2 + \beta^2(\underline{e}^{(2)})^2 + \gamma(\underline{e}^{(3)})^2 \qquad 3.2\text{-}2$$

eredményre vezet, vagyis tetszőleges α, β, γ esetén, melyre

$$\alpha^2 + \beta^2 + \gamma^2 > 0,$$

$$\underline{x} \neq \underline{0},$$

tehát $\underline{e}^{(1)}$, $\underline{e}^{(2)}$, $\underline{e}^{(3)}$ valóban nem egysíkú vektorok.

Tetszőleges $\underline{a}$ vektort előállíthatunk tehát az $\underline{e}^{(1)}$, $\underline{e}^{(2)}$, $\underline{e}^{(3)}$ vektorok lineáris kombinációjaként.

Legyen

$$\underline{a} = a_1\underline{e}^{(1)} + a_2\underline{e}^{(2)} + a_3\underline{e}^{(3)}. \qquad 3.2\text{-}3$$

Az a_1, a_2, a_3 értékeket az $\underline{a}$ vektor $\mathcal{K}$ koordináta-rendszerben vett koordinátáinak nevezzük.

Az a_k $(k=1, 2, 3)$ értékeket és szemléletes jelentésüket megkapjuk, ha a 3.2-3 egyenlőséget rendre beszorozzuk $\underline{e}^{(k)}$-val $(k=1, 2, 3)$. 3.1-2 felhasználásával azt kapjuk, hogy

$$\underline{a}\underline{e}^{(k)} = a_k \quad (k=1, 2, 3). \qquad 3.2\text{-}4$$

Az a_k értékek tehát az $\underline{a}$ vektornak az egyes koordinátatengelyekre vett vetületével egyenlők. Az $\underline{a}$ vektort a

$$\mathcal{K}(\underline{a}) = a_1, a_2, a_3 \qquad 3.2\text{-}5$$

számhármas jellemzi. Az egyszerűbb írásmód kedvéért azt is írhatjuk, hogy

$$a_1, a_2, a_3 = \mathbf{a}.$$

Ne tévesszük össze **a**-*t* $\underline{a}$-val! Az **a** egy számhármas, amely az iránnyal és abszolút értékkel rendelkező $\underline{a}$ vektort egy $\mathcal{K}$ koordináta-rendszerben reprezentálja, $\underline{a}$ pedig egy fizikai mennyiség. 3.2-5 helyett azt is írhatjuk, hogy

$$\mathcal{K}(\underline{a}) = \mathbf{a}. \qquad 3.2\text{-}6$$

a-t az $\underline{a}$ vektor $\mathcal{K}$-beli reprezentációjának nevezzük.

3.2.1. Az alapvektorok reprezentációja

Határozzuk meg az alapvektorok koordinátáit:

$$\mathcal{K}(\underline{e}^{(k)}) = \mathbf{e}^{(k)} \quad (k=1, 2, 3),$$

$$\mathbf{e}^{(k)} = e_1^{(k)}, e_2^{(k)}, e_3^{(k)} \quad (k=1, 2, 3).$$

Mivel

$$\underline{e}^{(k)} = \sum_{l=1}^{3} e_l^{(k)}\underline{e}^{(l)}, \qquad 3.2\text{-}6$$

a koordináták jelentésének 3.2-4 értelmezése alapján

$$e_l^{(k)} = \delta_{kl} \quad (k, l = 1, 2, 3). \qquad 3.2\text{-}7$$

Kevésbé tömör formában felírva 3.2-7 azt jelenti, hogy

$$\mathbf{e}^{(1)} = 1, 0, 0; \quad \mathbf{e}^{(2)} = 0, 1, 0; \quad \mathbf{e}^{(3)} = 0, 0, 1. \qquad 3.2\text{-}8$$

4. Vektorműveletek derékszögű koordináták segítségével

A vektorműveleteket koordináta-rendszertől függetlenül vezettük be. Definíciók konstruálásakor ez mindig lényeges követelmény, hiszen a reprezentációkat csak technikai okok miatt vezetjük be. A reprezentáció jellegzetességeinek azonban el kell tűnniük a formulákból, ha objektív összefüggéseket írunk le. Ezért csak reprezentációtól független műveleti szabályok esetén nyerhetünk olyan matematikai apparátust, amely a természet összefüggéseinek tükrözésére alkalmas.

Hasznos azonban azon rövid műveleti szabályok megkeresése, amelyek segítségével a reprezentációtól független, adott műveletek a reprezentációk segítségével hajthatók végre.

A következőkben a különböző vektorműveletek reprezentációival foglalkozunk.

4.1. Összeadás

Legyen

$$\underline{\mathbf{a}} + \underline{\mathbf{b}} = \underline{\mathbf{c}} \qquad 4.1\text{-}1$$

és

$$\mathcal{K}(\underline{\mathbf{a}}) = \mathbf{a} = a_1, a_2, a_3, \quad \mathcal{K}(\underline{\mathbf{b}}) = \mathbf{b} = b_1, b_2, b_3, \quad \mathcal{K}(\underline{\mathbf{c}}) = \mathbf{c} = c_1, c_2, c_3.$$

Tehát

$$\underline{\mathbf{a}} = \sum_{k=1}^{3} a_k \underline{\mathbf{e}}^{(k)}, \quad \underline{\mathbf{b}} = \sum_{k=1}^{3} b_k \underline{\mathbf{e}}^{(k)} \quad \text{és} \quad \underline{\mathbf{c}} = \sum_{k=1}^{3} c_k \underline{\mathbf{e}}^{(k)}. \qquad 4.1\text{-}2$$

4.1-2-t beírva 4.1-1-be kapjuk, hogy

$$\sum_{k=1}^{3} (a_k + b_k) \underline{\mathbf{e}}^{(k)} = \sum_{k=1}^{3} c_k \underline{\mathbf{e}}^{(k)}. \qquad 4.1\text{-}3$$

4.1-3 két oldalának összehasonlításából következik, hogy

$$a_k + b_k = c_k \quad (k = 1, 2, 3), \tag{4.1-4}$$

vagyis két vektor összegének (az eredő vektornak) komponenseit a megfelelő komponensek összege adja. A 4.1-4 formulákat az

$$\mathbf{a} + \mathbf{b} = \mathbf{c} \tag{4.1-5}$$

jelöléssel foglalhatjuk össze.

4.1-1 és 4.1-5 tulajdonképpen ugyanazt jelenti. Mindkét formula a két vektor összegezését fejezi ki. 4.1-1 arra utal, hogy az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok összege hogyan határozható meg a paralelogrammaszabály segítségével, 4.1-5 pedig arra, hogy a $\mathcal{K}(\underline{\mathbf{a}}) = \mathbf{a}$ és $\mathcal{K}(\underline{\mathbf{b}}) = \mathbf{b}$ reprezentációk segítségével hogyan határozható meg az összegvektor reprezentációja.

4.2. Szorzás skalárral

Legyen $\underline{\mathbf{c}} = \alpha \underline{\mathbf{a}}$ és

$$\underline{\mathbf{a}} = \sum_{k=1}^{3} a_k \underline{\mathbf{e}}^{(k)}.$$

A disztributivitás értelmében

$$\alpha \underline{\mathbf{a}} = \sum_{k=1}^{3} \alpha a_k \underline{\mathbf{e}}^{(k)},$$

vagyis a $\underline{\mathbf{c}}$ vektor komponensei az $\underline{\mathbf{a}}$ vektor komponenseinek α-szorosai. Skalárral tehát komponensenként szorzunk.

Tehát

$$\mathbf{c} = \alpha \mathbf{a}. \tag{4.2-1}$$

Az összeadási és szorzási művelet elvégzéséből következik, hogy a

$$\underline{\mathbf{b}} = \sum_{k=1}^{3} \alpha_\nu \underline{\mathbf{a}}_\nu \tag{4.2-2}$$

lineáris kombináció a $\mathcal{K}$ rendszerben a

$$\mathbf{b} = \sum_{k=1}^{3} \alpha_\nu \mathbf{a}_\nu \tag{4.2-3}$$

formula alapján határozható meg.

3*

4.3. A skaláris szorzat reprezentációja

Legyen

$$\underline{\mathbf{a}}=\sum_{k=1}^{3} a_k \underline{\mathbf{e}}^{(k)}; \quad \underline{\mathbf{b}}=\sum_{k=1}^{3} b_k \underline{\mathbf{e}}^{(k)}. \tag{4.3-1}$$

Ekkor az 1.4. fejezet értelmében a skaláris szorzat disztributivitását figyelembe véve:

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\sum_{k,\,l}^{3} a_k b_l \underline{\mathbf{e}}^{(k)}\underline{\mathbf{e}}^{(l)}. \tag{4.3-2}$$

Az $\underline{\mathbf{e}}^{(k)}$ ($k=1, 2, 3$) vektorok merőlegessége miatt a különböző alapvektorok szorzatát tartalmazó tagok eltűnnek, és így 4.3-2 az

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\sum_{k=1}^{3} a_k b_k \tag{4.3-3}$$

alakra egyszerűsödik. 4.3-3 tulajdonképpen a skalárszorzás $\mathcal{K}$ koordináta-rendszerbeli elvégzésére ad műveleti utasítást. Bevezethetjük az

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\mathbf{ab} \tag{4.3-4}$$

jelölést is. Itt a bal oldal $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ skalárszorzatát, a jobb oldal pedig a skalárszorzás $\mathcal{K}$ rendszerbeli kiszámítási módját jelenti, vagyis

$$ab\cos\vartheta=\sum_{k=1}^{3} a_k b_k. \tag{4.3-5}$$

Mivel az $ab\cos\vartheta$ érték csak az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektoroktól függ, a

$$\sum_{k=1}^{3} a_k b_k$$

értéknek — tetszőleges derékszögű koordináta-rendszert használva — azonosnak kell lennie.

Legyenek a $\mathcal{K}$ koordináta-rendszer alapvektorai $\underline{\mathbf{e}}^{(1)}$, $\underline{\mathbf{e}}^{(2)}$, $\underline{\mathbf{e}}^{(3)}$, a $\mathcal{K}'$ koordináta-rendszeré pedig $\underline{\mathbf{e}}^{(1)'}$, $\underline{\mathbf{e}}^{(2)'}$, $\underline{\mathbf{e}}^{(3)'}$. Az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ vektorok mind a $\mathcal{K}$, mind a $\mathcal{K}'$ koordináta-rendszerben előállíthatók.

Legyen

$$\underline{\mathbf{a}}=\sum_{k=1}^{3} a_k \underline{\mathbf{e}}^{(k)}=\sum_{k=1}^{3} a'_k \underline{\mathbf{e}}^{(k)'},$$

$$\underline{\mathbf{b}}=\sum_{k=1}^{3} b_k \underline{\mathbf{e}}^{(k)}=\sum_{k=1}^{3} b'_k \underline{\mathbf{e}}^{(k)'}.$$

Tehát

$$\mathcal{K}(\mathbf{a})=\mathbf{a}=a_1, a_2, a_3, \quad \mathcal{K}(\mathbf{b})=\mathbf{b}=b_1, b_2, b_3,$$

$$\mathcal{K}'(\mathbf{a})=\mathbf{a}'=a_1', a_2', a_3', \quad \mathcal{K}'(\mathbf{b})=\mathbf{b}'=b_1', b_2', b_3'.$$

Az $\underline{\mathbf{a}}\underline{\mathbf{b}}$ skalárszorzatot mindkét koordináta-rendszerbeli reprezentációval kifejezhetjük, és

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\sum_{k=1}^{3} a_k b_k=\sum_{k=1}^{3} a_k' b_k' \tag{4.3-6}$$

kell legyen. 4.3-6 direkt bizonyítására később térünk ki.

4.4. A vektoriális szorzás elvégzése derékszögű koordinátákkal

a) Határozzuk meg először egy jobbsodrású $\mathcal{K}$ koordináta-rendszer

$$\underline{\mathbf{e}}^{(1)}; \quad \underline{\mathbf{e}}^{(2)}; \quad \underline{\mathbf{e}}^{(3)}$$

alapvektorainak vektoriális szorzatát.

A vektoriális szorzat és az egységvektorok definíciója alapján

$$\begin{aligned}
&\underline{\mathbf{e}}^{(1)}\times\underline{\mathbf{e}}^{(2)}=\underline{\mathbf{e}}^{(3)}, \quad \underline{\mathbf{e}}^{(2)}\times\underline{\mathbf{e}}^{(1)}=-\underline{\mathbf{e}}^{(3)},\\
&\underline{\mathbf{e}}^{(2)}\times\underline{\mathbf{e}}^{(3)}=\underline{\mathbf{e}}^{(1)}, \quad \underline{\mathbf{e}}^{(3)}\times\underline{\mathbf{e}}^{(2)}=-\underline{\mathbf{e}}^{(1)},\\
&\underline{\mathbf{e}}^{(3)}\times\underline{\mathbf{e}}^{(1)}=\underline{\mathbf{e}}^{(2)}, \quad \underline{\mathbf{e}}^{(1)}\times\underline{\mathbf{e}}^{(3)}=-\underline{\mathbf{e}}^{(2)}.
\end{aligned} \tag{4.4-1}$$

Továbbá, minthogy tetszőleges egységvektor önmagával való vektoriális szorzata zérus,

$$\underline{\mathbf{e}}^{(k)}\times\underline{\mathbf{e}}^{(k)}=\underline{0} \quad (k=1, 2, 3). \tag{4.4-2}$$

4.4-1 és 4.4-2 a Levi—Civita-szimbólum segítségével az

$$\underline{\mathbf{e}}^{(K)}\times\underline{\mathbf{e}}^{(L)}=\sum_{j=1}^{3} \varepsilon_{KLj}\underline{\mathbf{e}}^{(j)} \tag{4.4-3}$$

alakban foglalható össze, ti. a jobb oldalon legfeljebb egy tag nem zérus, mert ha $K=L$, akkor $\varepsilon_{KKj}=0$ miatt minden tag eltűnik; amennyiben $K\neq L$, akkor egyetlen tag marad meg, amelynek indexére $j\neq K, L$. Szorozzuk 4.4-3-at $\underline{\mathbf{e}}^{(M)}$-mel; 3.1-1 segítségével az

$$\underline{\mathbf{e}}^{(M)}(\underline{\mathbf{e}}^{(K)}\times\underline{\mathbf{e}}^{(L)})=\varepsilon_{KLM} \tag{4.4-4}$$

eredményre jutunk. A 4.4-4 összefüggést direkt módon is megkaphatjuk, ha figyelembe vesszük, hogy az egységvektorok egységnyi térfogatú kockát határoznak meg.

A vektorok koordinátás előállítása alapján 4.4-4 az $\underline{\mathbf{e}}^{(K)}\times\underline{\mathbf{e}}^{(L)}$ vektoriális szorzat $\mathcal{K}$ koordináta-rendszerbeli reprezentációjának M-edik komponensét jelenti. Tehát

$$(\underline{\mathbf{e}}^{(K)}\times\underline{\mathbf{e}}^{(L)})_M=\varepsilon_{KLM}. \tag{4.4-5}$$

b) Tetszőleges vektorok vektoriális szorzata. Legyen

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=\underline{\mathbf{c}} \tag{4.4-6}$$

és

$$\underline{\mathbf{a}}=\sum_{i=1}^{3} a_i\underline{\mathbf{e}}^{(i)},\quad \underline{\mathbf{b}}=\sum_{j=1}^{3} b_j\underline{\mathbf{e}}^{(j)},\quad \underline{\mathbf{c}}=\sum_{k=1}^{3} c_j\underline{\mathbf{e}}^{(k)}. \tag{4.4-7}$$

4.4-7-et beírva 4.4-6-ba:

$$\sum_{i=1}^{3} a_i\underline{\mathbf{e}}^{(i)}\times\sum_{j=1}^{3} b_j\underline{\mathbf{e}}^{(j)}=\sum_{k=1}^{3} c_k\underline{\mathbf{e}}^{(k)}. \tag{4.4-8}$$

A disztributív törvény 2.7-24 szerinti felhasználásával

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=\sum_{i,j=1}^{3} a_ib_j(\underline{\mathbf{e}}^{(i)}\times\underline{\mathbf{e}}^{(j)})=\sum_{k=1}^{3} c_k\underline{\mathbf{e}}^{(k)}. \tag{4.4-9}$$

Tehát 4.4-3 felhasználásával

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=\sum_{i,j,k=1}^{3} \varepsilon_{ijk}a_ib_j\underline{\mathbf{e}}^{(k)}=\sum_{k=1}^{3} c_k\underline{\mathbf{e}}^{(k)}. \tag{4.4-10}$$

Megszorozva az egyenleteket $\underline{\mathbf{e}}^{(M)}$-mel és figyelembe véve, hogy $\underline{\mathbf{e}}^{(M)}\underline{\mathbf{e}}^{(k)}=\delta_{Mk}$,

$$c_M=(\underline{\mathbf{a}}\times\underline{\mathbf{b}})\underline{\mathbf{e}}^{(M)}=\sum_{i,j=1}^{3} a_ib_j\varepsilon_{Mij}. \tag{4.4-11}$$

Részletezve a 4.4-11 formulát, megkapjuk az $\underline{\mathbf{a}}\times\underline{\mathbf{b}}$ vektoriális szorzat $\mathcal{K}$ koordináta-rendszerbeli előállítását a tényezők komponenseivel:

$$\begin{aligned} c_1&=a_2b_3-a_3b_2,\\ c_2&=a_3b_1-a_1b_3,\\ c_3&=a_1b_2-a_2b_1. \end{aligned} \tag{4.4-12}$$

A 4.4-12 formulákat összefoglalóan

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \qquad 4.4\text{-}13$$

-vel is jelölhetjük. (Az aláhúzás nélküli fett betű a $\mathcal{K}$ rendszerbeli reprezentációt jelenti.)

Megjegyzés: A 4.4-12 formulák megjegyzését megkönnyíti, hogy minden egyes képletben az első három index az 1, 2, 3 számok ciklikus permutációja.

4.5. A hármas vegyes szorzat kifejezése koordináták segítségével. A determináns fogalma

Az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorokból képzett

$$(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}}) = v \qquad 4.5\text{-}1$$

hármas vegyes szorzat értékét a vektorok $\mathcal{K}$ rendszerben vett $\mathcal{K}(\underline{\mathbf{a}}) = \mathbf{a}$, $\mathcal{K}(\underline{\mathbf{b}}) = \mathbf{b}$, $\mathcal{K}(\underline{\mathbf{c}}) = \mathbf{c}$ reprezentációi segítségével a skaláris szorzatra vonatkozó 4.3-5 és a vektoriális szorzatra vonatkozó 4.4-10 összefüggés alapján határozhatjuk meg. Ezek felhasználásával

$$v = (\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{k,\, l,\, m=1}^{3} \varepsilon_{klm} a_k b_l c_m. \qquad 4.5\text{-}2$$

4.5-2-t részletesen kiírva,

$$v = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1 - a_1 b_3 c_2. \qquad 4.5\text{-}3$$

A 4.5-3 formula megjegyzését megkönnyíti, ha a vektorok komponenseit a

$$D = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \det(\mathbf{a}, \mathbf{b}, \mathbf{c}) \qquad 4.5\text{-}4$$

sémába rendezzük. A 4.5-4 formulával definiált mennyiséget *determinánsnak* nevezzük. A fogalmat később általánosabban is definiáljuk. A determinánsok jelentős szerepet játszanak az algebrában. A 4.5-4-ben szereplő det jelölés a determináns szó rövidítése. A 4.5-3 kifejtésből látható, hogy a hármasszorzat a vektorkomponensek pozitív és negatív előjellel vett bizonyos hármasszorzatainak összegét jelenti.

A 4.5-4 elrendezést egészítsük ki az első két oszlop ismételt leírásával

$$\begin{array}{ccccc} a_1 & a_2 & a_3 & a_1 & a_2 \\ b_1 & b_2 & b_3 & b_1 & b_2 \\ c_1 & c_2 & c_3 & c_1 & c_2 \end{array} \qquad 4.5\text{-}5$$

Képezzünk a folytonos vonalak (főátlók) mentén fekvő elemekből hármasszorzatokat, és lássuk el őket pozitív előjellel. A szaggatott vonalak (mellékátlók) mentén fekvő elemekből képzett hármasszorzatokat pedig lássuk el negatív előjellel.

Az így nyert hattagú összeg éppen az ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$) hármas vegyes szorzat értékét adja.

4.6. Vektor előállítása három, nem komplanáris vektorból

Legyen $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ három, nem egy síkba eső vektor, tehát

$$\underline{\mathbf{a}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=\det(\mathbf{a}, \mathbf{b}, \mathbf{c})\neq 0. \qquad 4.6\text{-}1$$

Már megmutattuk, hogy tetszőleges $\underline{\mathbf{A}}$ vektor előállítható az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok lineáris kombinációjaként.

Legyen

$$\underline{\mathbf{A}}=x\underline{\mathbf{a}}+y\underline{\mathbf{b}}+z\underline{\mathbf{c}}. \qquad 4.6\text{-}2$$

Az x, y, z szorzótényezőket a következőképpen határozhatjuk meg. Szorozzuk 4.6-2-t $\underline{\mathbf{b}}\times\underline{\mathbf{c}}$-vel. Minthogy

$$\underline{\mathbf{b}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=\underline{\mathbf{c}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=0,$$

azt kapjuk, hogy

$$\underline{\mathbf{A}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=x\underline{\mathbf{a}}(\underline{\mathbf{b}}\times\underline{\mathbf{c}}). \qquad 4.6\text{-}3$$

Tehát

$$x=\frac{(\underline{\mathbf{A}}, \underline{\mathbf{b}}, \underline{\mathbf{c}})}{(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}})}. \qquad 4.6\text{-}4$$

Hasonló módon, ha $\underline{\mathbf{c}}\times\underline{\mathbf{a}}$-val, ill. $\underline{\mathbf{a}}\times\underline{\mathbf{b}}$-vel szorzunk, az

$$y=\frac{(\underline{\mathbf{a}}, \underline{\mathbf{A}}, \underline{\mathbf{c}})}{(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}})}, \qquad 4.6\text{-}5$$

ill.

$$z=\frac{(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{A}})}{(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}})} \qquad 4.6\text{-}6$$

eredményt kapjuk. 4.6-5 és 4.6-6-ban a számlálóban szereplő hármas vegyes szorzatokat a szimmetria kedvéért az

$$\underline{\mathbf{A}}(\underline{\mathbf{c}}\times\underline{\mathbf{a}})=\underline{\mathbf{a}}(\underline{\mathbf{A}}\times\underline{\mathbf{c}})$$

és

$$\underline{\mathbf{A}}(\underline{\mathbf{a}}\times\underline{\mathbf{b}})=\underline{\mathbf{a}}(\underline{\mathbf{b}}\times\underline{\mathbf{A}})$$

azonosságoknak megfelelően átalakítottuk.

Az x, y, z faktorok meghatározásakor alkalmazott módszert háromismeretlenes egyenletrendszerek megoldására is felhasználhatjuk. A 4.6-2 egyenletet tetszőleges reprezentációban felírva, azt kapjuk, hogy

$$\begin{aligned} a_1x+b_1y+c_1z&=A_1,\\ a_2x+b_2y+c_2z&=A_2,\\ a_3x+b_3y+c_3z&=A_3. \end{aligned} \tag{4.6-7}$$

Az egyenletrendszer megoldását a 4.6-4, 4.6-5, 4.6-6 formulák szolgáltatják, csak az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$, $\underline{\mathbf{A}}$ vektorokat a megfelelő **a**, **b**, **c**, **A** reprezentációkkal kell helyettesítenünk. Tehát

$$x=\frac{\det(\mathbf{A},\mathbf{b},\mathbf{c})}{\det(\mathbf{a},\mathbf{b},\mathbf{c})},\quad y=\frac{\det(\mathbf{a},\mathbf{A},\mathbf{c})}{\det(\mathbf{a},\mathbf{b},\mathbf{c})},\quad z=\frac{\det(\mathbf{a},\mathbf{b},\mathbf{A})}{\det(\mathbf{a},\mathbf{b},\mathbf{c})}. \tag{4.6-8}$$

A megoldások 4.6-8 alakú determinánsokkal történő előállítását *Cramer-szabálynak* nevezzük. A későbbiekben látni fogjuk, hogy a módszer általánosított formában tetszőleges számú ismeretlent tartalmazó egyenletre is alkalmazható.

4.7. A vektoriális hármasszorzat

Vizsgáljuk meg az

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=\underline{\mathbf{d}} \tag{4.7-1}$$

alakú hármasszorzatot. Azonnal látható, hogy 4.7-1-ben a zárójel nem felesleges, mert a vektoriális szorzás nem asszociatív művelet.

Ezt abból láthatjuk, hogy 4.7-1-et $\underline{\mathbf{a}}$-val szorozva, az

$$\underline{\mathbf{a}}\underline{\mathbf{d}}=0$$

eredményt kapjuk, vagyis $\underline{\mathbf{d}}$ a $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok síkjában fekszik. Az $(\underline{\mathbf{a}}\times\underline{\mathbf{b}})\times\underline{\mathbf{c}}=\underline{\mathbf{d}}'$ vektor viszont az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok síkjában van. $\underline{\mathbf{d}}$ és $\underline{\mathbf{d}}'$ általában tehát nem is azonos síkú vektorok.

Mivel a 4.7-1 szorzás egy a $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ síkban fekvő vektorra vezet, írhatjuk, hogy

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=\beta\underline{\mathbf{b}}+\gamma\underline{\mathbf{c}}. \qquad 4.7\text{-}2$$

A β és γ szorzótényezők a következőképpen határozhatók meg. Állítsuk elő az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorokat az $\underline{\mathbf{e}}^{(K)}$ ($K=1, 2, 3$) ortogonális egységvektorok segítségével. Tehát legyen

$$\underline{\mathbf{a}}=\sum_{k=1}^{3} a_k\underline{\mathbf{e}}^{(k)},\quad \underline{\mathbf{b}}=\sum_{l=1}^{3} b_l\underline{\mathbf{e}}^{(l)},\quad \underline{\mathbf{c}}=\sum_{m=1}^{3} c_m\underline{\mathbf{e}}^{(m)}. \qquad 4.7\text{-}3$$

4.7-3-at 4.7-1-be behelyettesítve és a disztributivitást felhasználva:

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=\sum_{k,\,l,\,m} a_k b_l c_m[\underline{\mathbf{e}}^{(k)}\times(\underline{\mathbf{e}}^{(l)}\times\underline{\mathbf{e}}^{(m)})], \qquad 4.7\text{-}4$$

vagyis tetszőleges vektorok hármasszorzata kifejezhető az alapvektorok hármasszorzatával.

Vezessük be az

$$\underline{\mathbf{E}}_{klm}=\underline{\mathbf{e}}^{(k)}\times(\underline{\mathbf{e}}^{(l)}\times\underline{\mathbf{e}}^{(m)}) \qquad 4.7\text{-}5$$

jelölést. A különböző indexű $\underline{\mathbf{E}}_{klm}$-ek vagy zérus értékűek, vagy egységvektorok:

$$\underline{\mathbf{E}}_{klm}=0,\quad \text{ha}\quad \begin{cases} l=m, \\ k\neq l\neq m, \end{cases} \qquad 4.7\text{-}6$$

hiszen $l=m$ esetén az $\underline{\mathbf{e}}^{(l)}\times\underline{\mathbf{e}}^{(m)}$ szorzat értéke zérus, $k\neq l\neq m$ esetén pedig $\underline{\mathbf{e}}^{(l)}\times\underline{\mathbf{e}}^{(m)}$ párhuzamos $\underline{\mathbf{e}}^{(k)}$-val. $\underline{\mathbf{E}}_{klm}$ nem zérus értékei a következők:

$$\underline{\mathbf{E}}_{llm}=\underline{\mathbf{e}}^{(l)}\times(\underline{\mathbf{e}}^{(l)}\times\underline{\mathbf{e}}^{(m)})=-\underline{\mathbf{e}}^{(m)}, \qquad 4.7\text{-}7$$

$$\underline{\mathbf{E}}_{mlm}=\underline{\mathbf{e}}^{(m)}\times(\underline{\mathbf{e}}^{(l)}\times\underline{\mathbf{e}}^{(m)})=\underline{\mathbf{e}}^{(l)}. \qquad 4.7\text{-}8$$

Az összes lehetséges eseteket az

$$\underline{\mathbf{E}}_{klm}=\underline{\mathbf{e}}^{(l)}\delta_{km}-\underline{\mathbf{e}}^{(m)}\delta_{kl} \qquad 4.7\text{-}9$$

formulával foglalhatjuk össze. Ennek felhasználásával 4.7-4 az

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}\times\underline{\mathbf{c}})=\sum_{k,\,l,\,m} a_k b_l c_m(\underline{\mathbf{e}}^{(l)}\delta_{km}-\underline{\mathbf{e}}^{(m)}\delta_{kl}) \qquad 4.7\text{-}10$$

alakban írható fel. Minthogy

$$\sum_{k,\,l} a_k c_m\delta_{km}=\sum a_k c_k=\underline{\mathbf{a}}\underline{\mathbf{c}},$$

$$\sum_{k,\,l} a_k b_l\delta_{kl}=\sum a_k b_k=\underline{\mathbf{a}}\underline{\mathbf{b}},$$

azt kapjuk, hogy

$$\underline{a}\times(\underline{b}\times\underline{c})=\underline{b}(\underline{a}\underline{c})-\underline{c}(\underline{a}\underline{b}). \qquad 4.7\text{-}11$$

Hasonló módon

$$(\underline{a}\times\underline{b})\times\underline{c}=\underline{b}(\underline{a}\underline{c})-\underline{a}(\underline{b}\underline{c}). \qquad 4.7\text{-}12$$

A 4.7-11 és 4.7-12 formulákat a vektoriális hármasszorzat kifejtési szabályának nevezzük.

A 4.7-11 és 4.7-12 formulák megjegyzését megkönnyíti, ha a következő szabály szerint járunk el. Kezdjük a felírást a középső tényezővel. Ezt szorozzuk a másik két tényező skalárszorzatával. A negatív tagot úgy kapjuk, hogy a zárójelben levő másik tényezőt szorozzuk a megmaradó két tényező skalárszorzatával. A fenti szabályok a gyakorlati felhasználás szempontjából nagyon fontosak, ezért célszerű megjegyezni őket.

4.8. Vektorok négyesszorzatai

Az alkalmazások során gyakran találkozunk a következő típusú négyesszorzatokkal

a)
$$(\underline{a}\times\underline{b})\times(\underline{c}\times\underline{d})=\underline{E}. \qquad 4.8\text{-}1$$

A 4.8-1 típusú négyesszorzatot a következőképpen fejthetjük ki. Vezessük be az

$$\underline{A}=\underline{a}\times\underline{b} \qquad 4.8\text{-}2$$

jelölést. Ezzel

$$\underline{E}=(\underline{a}\times\underline{b})\times(\underline{c}\times\underline{d})=\underline{A}\times(\underline{c}\times\underline{d})=\underline{c}(\underline{A}\underline{d})-\underline{d}(\underline{A}\underline{c}). \qquad 4.8\text{-}3$$

4.8-3-ba visszaírva 4.8-2-t:

$$\underline{E}=[(\underline{a}\times\underline{b})\underline{d}]\underline{c}-[(\underline{a}\times\underline{b})\underline{c}]\underline{d}=(\underline{a},\underline{b},\underline{d})\underline{c}-(\underline{a},\underline{b},\underline{c})\underline{d}. \qquad 4.8\text{-}4$$

Hasonló módon a

$$\underline{B}=\underline{c}\times\underline{d}$$

jelölés bevezetésével az

$$\underline{E}=(\underline{a}\times\underline{b})\times\underline{B}=\underline{b}(\underline{a}\underline{B})-\underline{a}(\underline{b}\underline{B})=\underline{b}(\underline{a},\underline{c},\underline{d})-\underline{a}(\underline{b},\underline{c},\underline{d}) \qquad 4.8\text{-}5$$

eredményt kapjuk.

Tehát

$$(\underline{a}\times\underline{b})\times(\underline{c}\times\underline{d})=-\underline{a}(\underline{b},\underline{c},\underline{d})+\underline{b}(\underline{a},\underline{c},\underline{d})=-\underline{d}(\underline{a},\underline{b},\underline{c})+\underline{c}(\underline{a},\underline{b},\underline{d}). \qquad 4.8\text{-}6$$

b) A négyesszorzatok másik fontos típusa az

$$A=(\underline{a}\times\underline{b})(\underline{c}\times\underline{d}) \qquad 4.8\text{-}7$$

típusú szorzat. 4.8-7 a vegyes szorzat

$$A=(\underline{d}, (\underline{a}\times\underline{b}), \underline{c})$$

átalakításával és a 4.7-12 kifejtési tétel segítségével az

$$A=(\underline{a}\underline{c})(\underline{b}\underline{d})-(\underline{a}\underline{d})(\underline{b}\underline{c}) \qquad 4.8\text{-}8$$

alakban is felírható.

4.8.1 Reciprok vektorok

A négyesszorzat egy fontos alkalmazása a következő. Legyen $\underline{a}$, $\underline{b}$, $\underline{c}$ három adott, nem komplanáris vektor, tehát

$$(\underline{a}, \underline{b}, \underline{c})=v\neq 0. \qquad 4.8\text{-}10$$

Alkossunk belőlük páronként vektoriális szorzatokat, és osszuk ezeket v-vel.

Legyen tehát

$$\underline{A}=\frac{1}{v}(\underline{b}\times\underline{c}),$$

$$\underline{B}=\frac{1}{v}(\underline{c}\times\underline{a}), \qquad 4.8\text{-}11$$

$$\underline{C}=\frac{1}{v}(\underline{a}\times\underline{b}).$$

Az így definiált $\underline{A}$, $\underline{B}$, $\underline{C}$ vektorok rendre merőlegesek a $\underline{b}$, $\underline{c}$; $\underline{c}$, $\underline{a}$ és $\underline{a}$, $\underline{b}$ vektorok síkjára.

A 4.8-6 szabály szerint

$$\underline{A}\times\underline{B}=\frac{1}{v}[(\underline{b}\times\underline{c})\times(\underline{c}\times\underline{a})]=\frac{1}{v^2}[-\underline{b}(\underline{c}, \underline{c}, \underline{a})+\underline{c}(\underline{b}, \underline{c}, \underline{a})].$$

Minthogy

$$(\underline{c}, \underline{c}, \underline{a})=0 \text{ és } (\underline{b}, \underline{c}, \underline{a})=v,$$

azt kapjuk, hogy

$$\underline{A}\times\underline{B}=\frac{1}{v}\underline{c}.$$

Megszorozva ezt az összefüggést

$$(\underline{a}\times\underline{b})=\underline{C}v\text{-vel},$$

$$(\underline{A}, \underline{B}, \underline{C})v=\frac{1}{v}(\underline{a}\times\underline{b})\underline{c}=1.$$

tehát

$$(\underline{\mathbf{A}}, \underline{\mathbf{B}}, \underline{\mathbf{C}}) = \frac{1}{v} = V. \qquad 4.8\text{-}12$$

Hasonló módon kapjuk, hogy

$$\underline{\mathbf{B}} \times \underline{\mathbf{C}} = \frac{1}{v} \underline{\mathbf{a}} \quad \text{és} \quad \underline{\mathbf{C}} \times \underline{\mathbf{A}} = \frac{1}{v} \underline{\mathbf{b}}.$$

Így 4.8-10, 11, 12-ből azt kapjuk, hogy

$$\underline{\mathbf{a}} = \frac{1}{V} (\underline{\mathbf{B}} \times \underline{\mathbf{C}}),$$

$$\underline{\mathbf{b}} = \frac{1}{V} (\underline{\mathbf{C}} \times \underline{\mathbf{A}}), \qquad 4.8\text{-}13$$

$$\underline{\mathbf{c}} = \frac{1}{V} (\underline{\mathbf{A}} \times \underline{\mathbf{B}}),$$

ahol

$$(\underline{\mathbf{A}}, \underline{\mathbf{B}}, \underline{\mathbf{C}}) = V$$

4.8-13 a 4.8-10 és 4.8-11 megfordítása. A két összefüggés szimmetriája miatt az $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$, $\underline{\mathbf{C}}$ vektorokat az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok *reciprok rendszerének* nevezzük. A szimmetria folytán látható, hogy az

$$\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}} \quad \text{pedig az} \quad \underline{\mathbf{A}}, \underline{\mathbf{B}}, \underline{\mathbf{C}}$$

vektorok reciprok rendszerét alkotja. E megfordíthatóság indokolja a *reciprok* elnevezést. Az elnevezés eredete még jobban kitűnik ortogonális $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorhármas esetén. Ekkor ugyanis 4.8-11-ből $|\underline{\mathbf{A}}| = \frac{1}{a}$, $|\underline{\mathbf{B}}| = \frac{1}{b}$, $|\underline{\mathbf{C}}| = \frac{1}{c}$.

5. Analitikus geometria

5.1. A helyzetvektor és a görbe egyenletének fogalma

Az eddig kifejtett vektorformalizmus alkalmas különböző geometriai problémák egyszerű tárgyalására. A tér pontjainak helyzetét jellemezhetjük a tér egy kiválasztott pontjából a kérdéses P pontokhoz vezető $\underline{\mathbf{r}}$ vektorok segítségével. Ezeket a vektorokat *helyzetvektoroknak* vagy *helyvektoroknak* nevezzük.

Az eddigiekben a vektorokkal végzett műveletek esetén csak a vektorok nagysága és iránya volt fontos számunkra. A helyvektorok esetén a tér pontjai csak akkor jellemezhetők egyértelműen, ha az egyes helyvektorokat mindig ugyanabból a pontból mérjük fel. *A helyvektorokat ilyen értelemben „kötött vektoroknak" nevezzük.* Vektorok segítségével görbéket, illetve felületeket jellemezhetünk, ha megfelelő módon megadjuk azokat a helyzetvektorokat, amelyek az alakzat pontjait meghatározzák. Egy görbét tehát megadhatunk mint azon pontok halmazát, amelyek helyzetvektorai

$$\underline{\mathbf{r}}=\underline{\mathbf{r}}(p) \qquad 5.1\text{-}1$$

alakban állíthatók elő, ahol p egy független paraméter. A paraméter számértékének megfelelően kapjuk a görbe pontjait. Így tehát az

$$\underline{\mathbf{r}}_1=\underline{\mathbf{r}}(p_1), \quad \underline{\mathbf{r}}_2=\underline{\mathbf{r}}(p_2) \quad \text{stb.}$$

egy görbe pontjainak helyzetvektorai a

$$p=p_1, p_2, \ldots$$

paraméterértékeknek megfelelően.

Kényelmi okokból $\underline{\mathbf{r}}(p)$-t futóvektornak nevezzük. A futóvektor tehát nem egyetlen vektort jelent, hanem p változásakor „végigfut" a görbe összes pontjain.

Felületet két paraméter segítségével határozhatunk meg. Egy felület pontjai eleget tesznek egy

$$\underline{\mathbf{r}}=\underline{\mathbf{r}}(p, q) \qquad 5.1\text{-}2$$

összefüggésnek, ahol

$\underline{\mathbf{r}}$ a felület futópontjainak helyzetvektora. Így a felület pontjai a fenti képletből a p és q paraméterértékek segítségével számíthatók, azaz

$$\underline{\mathbf{r}}_{11}=\underline{\mathbf{r}}(p_1, q_1), \underline{\mathbf{r}}_{12}=\underline{\mathbf{r}}(p_1, q_2), \ldots, \underline{\mathbf{r}}_{kl}=\underline{\mathbf{r}}(p_k, q_l)$$

a felület pontjai, ahol

$$p=p_1, p_2, \ldots, p_k, \ldots$$
$$q=q_1, q_2, \ldots, q_l, \ldots$$

adott paraméterértékek.

A fenti módszert a görbék és felületek *parametrikus* előállításának nevezzük. Az 5.1-1, illetve 5.1-2 egyenletek koordinátareprezentációiból egyes esetekben a paraméterek eliminálhatók. Az

$$F(\mathbf{r})=0 \qquad 5.1\text{-}3$$

alakú összefüggést a felület implicit egyenletének nevezzük. Itt F valamilyen vektorfüggvény. A felületet azok a helyzetvektorok alkotják, amelyek 5.1-3-at kielégítik.

Vegyük észre, hogy a felületek és a görbék paraméteres reprezentációja között összefüggés áll fenn.

5.1-2 felületet határoz meg. Azonban, ha az egyik paramétert rögzítjük, akkor

$$\mathbf{r}_1 = \mathbf{r}(p_1, q) \qquad 5.1\text{-}4$$

összefüggés p_1 adott értéke mellett 5.1-1 szerint egy görbét határoz meg. Ez a görbe az 5.1-2 felületen fekszik. Így különböző rögzített $p = p_1, p_2, \ldots$ értékeket helyettesitve 5.1-2-be, a felületen egy görbesereget kapunk. Hasonló módon egy másik görbesereghez jutunk, ha a második paraméter értékét rögzítjük, az elsőt pedig független paraméternek tekintjük. Így az

$$\mathbf{r} = \mathbf{r}(q_k, p) \quad (k = 1, 2, 3, \ldots) \qquad 5.1\text{-}5$$

görbéket kapjuk. Ezek a görbék metszik az 5.1-4 görbéket. Az 5.1-4 és 5.1-5 görbeseregek hálózatot képeznek az 5.1-2 felületen — és ezek segítségével egy felületi koordináta-rendszerhez juthatunk. Erre a kérdésre a későbbiekben még visszatérünk. A következőkben egyes görbéket és felületeket vektorkifejezésekkel jellemzünk.

5.1.1. Az egyenes egyenlete

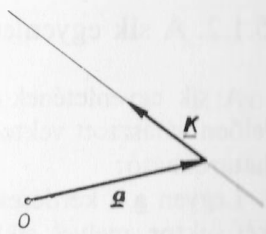
Egy egyenest egyértelműen meghatározhatunk például egy pontjával és az irányával. Legyen $\underline{\mathbf{a}}$ egy egyenes valamely pontjának helyvektora, a $\underline{\mathbf{K}}$ vektor pedig legyen párhuzamos az egyenessel.

Az *5.1. ábra* alapján látható, hogy az egyenes tetszőleges pontjához eljuthatunk, ha az $\underline{\mathbf{a}}$ helyvektorú pontból a $\underline{\mathbf{K}}$ vektor irányában haladunk.

Képletben kifejezve:

$$\underline{\mathbf{r}} = \underline{\mathbf{a}} + p\underline{\mathbf{K}}. \qquad 5.1\text{-}6$$

5.1-6-ot az egyenes paraméteres egyenletének nevezzük. A formulában p egy paraméter, amely $-\infty$-től $+\infty$-ig minden értéket felvehet, $\underline{\mathbf{r}}$ pedig az egyenes futóvektora. Amennyiben p összes lehetséges értékein végigfutunk, $\underline{\mathbf{r}}$ az egyenes összes pontjait megadja.



5.1. ábra

Az egyenes egyenletét paraméter nélkül is kifejezhetjük. Állítsuk elő ugyanis a $\underline{\mathbf{K}}$ vektort

$$\underline{\mathbf{K}} = \underline{\mathbf{L}} \times \underline{\mathbf{M}} \qquad 5.1$$

alakban. Itt $\underline{\mathbf{L}}$ és $\underline{\mathbf{M}}$ a $\underline{\mathbf{K}}$ vektor egyenesére merőleges síkba eső vektorok. 5.1-7-ből következik, hogy

$$\underline{\mathbf{K}}\underline{\mathbf{L}} = \underline{\mathbf{K}}\underline{\mathbf{M}} = 0,$$

tehát 5.1-6-ot $\underline{\mathbf{L}}$-lel, illetve $\underline{\mathbf{M}}$-mel szorozva, azt kapjuk, hogy

$$\underline{\mathbf{r}}\underline{\mathbf{L}}=\alpha, \quad \text{és} \quad \underline{\mathbf{r}}\underline{\mathbf{M}}=\beta, \tag{5.1-8}$$

ahol

$$\alpha=\underline{\mathbf{a}}\underline{\mathbf{L}}, \quad \text{és} \quad \beta=\underline{\mathbf{a}}\underline{\mathbf{M}}. \tag{5.1-9}$$

Mivel az 5.1-8 egyenletpárt az egyenes pontjai és csak az egyenes pontjai elégítik ki, 5.1-8-at is az egyenes egyenletének tekinthetjük. Ezt a formát az egyenes *implicit egyenletének* nevezzük.

Legyen

$$\begin{gathered}\mathcal{K}(\underline{\mathbf{r}})=\mathbf{r}=x_1, x_2, x_3, \quad \mathcal{K}(\underline{\mathbf{L}})=\mathbf{L}=L_1, L_2, L_3,\\ \mathcal{K}(\underline{\mathbf{M}})=\mathbf{M}=M_1, M_2, M_3\end{gathered} \tag{5.1-10}$$

az $\underline{\mathbf{r}}$, $\underline{\mathbf{L}}$ és $\underline{\mathbf{M}}$ vektorok $\mathcal{K}$ koordináta-rendszerben vett reprezentációja.

Az 5.1-10 reprezentáció segítségével 5.1-11 az

$$\left.\begin{aligned}L_1x_1+L_2x_2+L_3x_3=\alpha\\ M_1x_1+M_2x_2+M_3x_3=\beta\end{aligned}\right\} \tag{5.1-12}$$

alakban írható fel. Az L_i és M_i $(i=1, 2, 3)$ értékek között fennáll az

$$|\underline{\mathbf{K}}|\neq 0\text{-t} \tag{5.1-13}$$

kifejező

$$(L_2M_3-L_3M_2)^2+(L_3M_1-L_1M_3)^2+(L_1M_2-L_2M_1)^2>0$$

összefüggés.

Az 5.1-12 egyenletek az egyenes 5.1-8 implicit egyenletének $\mathcal{K}$ rendszerbeli reprezentációját jelentik.

5.1-12 egyenletrendszerből kifejezhetjük x_1-et és x_2-t x_3 segítségével. Ekkor az egyenes egyenletének explicit alakjához jutunk.

5.1.2. A sík egyenlete

A sík egyenletének meghatározásához vegyük figyelembe, hogy két megfelelően választott vektor lineáris kombinációjaként egy sík minden pontja meghatározható:

Legyen $\underline{\mathbf{a}}$ a kérdéses $\mathcal{S}$ sík tetszőleges pontjának helyvektora, $\underline{\mathbf{L}}$ és $\underline{\mathbf{M}}$ pedig két vektor, melyek síkja párhuzamos az $\mathcal{S}$ síkkal, és amelyekre

$$\underline{\mathbf{K}}=\underline{\mathbf{L}}\times\underline{\mathbf{M}}\neq\underline{\mathbf{0}}. \tag{5.1-14}$$

Ekkor a sík tetszőleges pontjának helyvektora előállítható az

$$\underline{\mathbf{r}}=\underline{\mathbf{a}}+p\underline{\mathbf{L}}+q\underline{\mathbf{M}} \tag{5.1-15}$$

alakban, ahol a p és q paraméterek értéke $-\infty$ és $+\infty$ között minden értéket felvehet *(5.2. ábra)*.

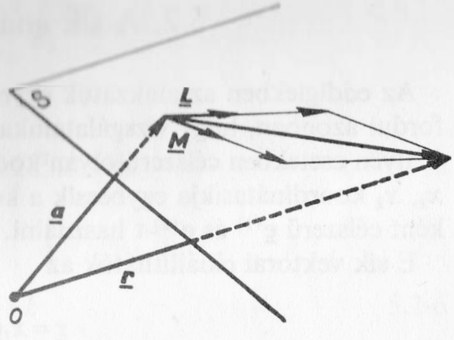
Az 5.1-15 egyenletet a sík paraméteres egyenletének nevezzük.

5.1-15 egyenletet $\underline{\mathbf{K}}$-val skalárisan szorozva, a

$$\underline{\mathbf{K}}\underline{\mathbf{r}} = \alpha, \qquad 5.1\text{-}16$$

ahol

$$\alpha = \underline{\mathbf{K}}\underline{\mathbf{a}}. \qquad 5.1\text{-}17$$



5.2. ábra

A $\underline{\mathbf{K}}$ vektort — mivel $\underline{\mathbf{K}}$ merőleges az $\mathcal{S}$ síkra — a sík normálvektorának nevezzük. 5.1-16 a sík implicit egyenlete.

Az

$$\mathbf{r} = \mathcal{K}(\underline{\mathbf{r}}), \quad \mathbf{K} = \mathcal{K}(\underline{\mathbf{K}})$$

koordinátareprezentáció segítségével a sík implicit egyenlete a

$$K_1x_1 + K_2x_2 + K_3x_3 = \alpha, \qquad 5.1\text{-}18$$

ahol

$$K_1^2 + K_2^2 + K_3^2 > 0$$

alakban is felírható.

Amennyiben $K_3 \neq 0$, az 5.1-18 egyenletből kifejezhető x_3, és a sík egyenletének

$$x_3 = \frac{\alpha - K_1x_1 - K_2x_2}{K_3} \qquad 5.1\text{-}19$$

explicit alakjához jutunk.

Bevezetve a

$$C = \frac{\alpha}{K_3}\,; \quad B = \frac{-K_1}{K_3}, \quad A = \frac{-K_2}{K_3} \qquad 5.1\text{-}20$$

jelöléseket, a sík explicit egyenlete az

$$x_3 = Ax_1 + Bx_2 + C \qquad 5.1\text{-}21$$

formára hozható.

A $K_3 \neq 0$ feltétel csak annyit jelent, hogy a vizsgált sík nem tartalmazza az x_3 tengelyt. Az utóbbi esetben 5.1-21 helyett azt kapjuk, hogy

$$K_1x_1 + K_2x_2 = \alpha, \quad x_3 \text{ tetszőleges.} \qquad 5.1\text{-}22$$

Az eddigiekben az alakzatok egyenletét a térben írtuk le. Igen gyakran előfordul azonban, hogy vizsgálatainkat elegendő egyetlen síkra korlátoznunk. Ilyen esetekben célszerű olyan koordináta-rendszert választanunk, amelynek x_1, x_2 koordinátasíkja egybeesik a kérdéses síkkal, valamint $\underline{\mathbf{L}}$ és $\underline{\mathbf{M}}$ vektorokként célszerű $\underline{\mathbf{e}}^{(1)}$ és $\underline{\mathbf{e}}^{(2)}$-t használni. Ekkor $\underline{\mathbf{K}}=\underline{\mathbf{e}}^{(3)}$.

E sík vektorai előállíthatók az

$$\underline{\mathbf{r}}=x_1\underline{\mathbf{e}}^{(1)}+x_2\underline{\mathbf{e}}^{(2)} \qquad 5.2\text{-}1$$

formában.

5.2.1. Az egyenes egyenlete

A síkban fekvő, a sík $\underline{\mathbf{a}}$ helyvektorú pontján átmenő és a $\underline{\mathbf{K}}$ vektorral párhuzamos egyenesének egyenletét 5.1-1 szerint az

$$\underline{\mathbf{r}}=\underline{\mathbf{a}}+p\underline{\mathbf{K}} \qquad 5.2\text{-}2$$

alakban írhatjuk fel. $\underline{\mathbf{L}}$ vektorként $\underline{\mathbf{e}}^{(3)}$-at választva, az $\underline{\mathbf{M}}=\underline{\mathbf{n}}$ vektor az x_1, x_2 síkba eső és az 5.2-2 egyenesre merőleges vektor.

Az egyenes 5.1-3 implicit egyenlete ekkor

$$\underline{\underline{\mathbf{n}}}\underline{\underline{\mathbf{r}}}=\underline{\underline{\mathbf{n}}}\underline{\underline{\mathbf{a}}} \qquad 5.2\text{-}3$$

alakot ölt. (Csak egy egyenletet kapunk, mert $\underline{\mathbf{L}}=\underline{\mathbf{e}}^{(3)}$ merőleges $\underline{\mathbf{a}}$-ra is.) Az $\underline{\mathbf{n}}$ vektort az egyenes normálvektorának nevezzük.

A

$$\mathcal{K}(\underline{\mathbf{n}})=\mathbf{n}=n_1,\, n_2;\quad \mathcal{K}(\underline{\mathbf{a}})=\mathbf{a}=a_1,\, a_2 \quad \text{és} \quad \mathcal{K}(\underline{\mathbf{r}})=\mathbf{r}=r_1,\, r_2 \qquad 5.2\text{-}4$$

reprezentációban az 5.2-3 egyenlet az

$$\mathbf{nr}=\mathbf{na}$$

formát ölti. Ezt az egyenletet koordinátás alakban kifejtve:

$$n_1x_1+n_2x_2=n_1a_1+n_2a_2. \qquad 5.2\text{-}5$$

A harmadik koordináta mindenütt zérus, ezért mindenütt elhagytuk.

5.2-5-ből egyszerűen meghatározhatjuk az egyenes egyenletének explicit alakját, ha $n_2\neq 0$. 5.2-5-ből x_2-t kifejezve, az

$$x_2 = -\frac{n_1}{n_2}x_1 + \frac{n_1a_1 + n_2a_2}{n_2}$$

formához jutunk. A

$$-\frac{n_1}{n_2} = m \quad \text{és} \quad b = \frac{n_1a_1 + n_2a_2}{n_2}$$

jelölésekkel éppen az egyenes ismert

$$x_2 = mx_1 + b \qquad 5.2\text{-}6$$

iránytényezős alakját kapjuk.

Abban az esetben, ha m, $b \neq 0$, akkor b-vel osztva, 5.2-6-ot az

$$\frac{x_1}{p_1} + \frac{x_2}{p_2} = 1 \qquad 5.2\text{-}7$$

alakban is felírhatjuk, ahol

$$p_1 = -\frac{b}{m}, \quad p_2 = b.$$

5.2-7-et az egyenes tengelymetszetes egyenletének nevezzük, mert p_1 és p_2 az egyenes által a tengelyekből kimetszett szakaszok hossza.

5.2.2. A kör egyenlete

Egyszerűen felírhatjuk az x_1, x_2 síkban fekvő $\underline{a}$ középpontú kör egyenletét is. Az *5.3. ábra* alapján ugyanis

$$(\underline{r} - \underline{a})^2 = R^2, \qquad 5.2\text{-}8$$

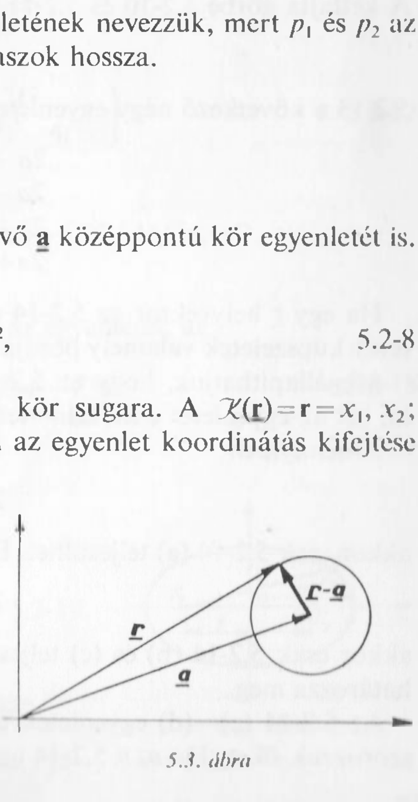
ahol $\underline{r}$ a futópont helyvektora, R pedig a kör sugara. A $\mathcal{K}(\underline{r}) = \mathbf{r} = x_1, x_2$; $\mathcal{K}(\underline{a}) = \mathbf{a} = a_1, a_2$ koordinátareprezentációban az egyenlet koordinátás kifejtése az

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 = R^2 \qquad 5.2\text{-}9$$

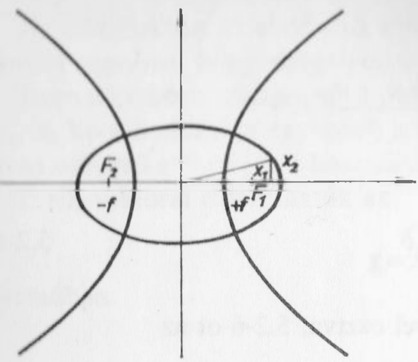
alakot ölti. 5.2-9-ről leolvasható, hogy az origó középpontú ($\underline{a} = \underline{0}$) kör egyenlete

$$x_1^2 + x_2^2 = R^2$$

alakú.



5.3. ábra



5.4. *ábra*

Az ellipszis azon pontok mértani helye, amelyek két ponttól — a fókuszpontoktól — vett u, v távolságainak összege állandó. Jelöljük ezt az állandót $2a$-val. Ekkor az ellipszis pontjaira *(5.4. ábra)*

$$u+v=2a. \qquad 5.2\text{-}10$$

A hiperbola azon pontok mértani helye, amelyek két ponttól vett u, v távolságai különbségének abszolút értéke állandó. Tehát a hiperbola pontjaira

$$|u-v|=2a. \qquad 5.2\text{-}11$$

Az 5.2-10 és 5.2-11 összefüggések az ellipszis és hiperbola futópontjával ($\underline{\mathbf{r}}$) és a fókuszok $\underline{\mathbf{f}}_1$ és $\underline{\mathbf{f}}_2$ helyvektoraival is kifejezhetők, hiszen

$$u=|\underline{\mathbf{r}}-\underline{\mathbf{f}}_1|, \quad v=|\underline{\mathbf{r}}-\underline{\mathbf{f}}_2|. \qquad 5.2\text{-}12$$

A kétfajta görbe 5.2-10 és 5.2-11 egyenleteit egyetlen egyenletté egyesíthetjük:

$$\pm u \pm v=2a. \qquad 5.2\text{-}13$$

5.2-13 a következő négy egyenletet adja egyszerre:

$$\begin{aligned} 2a-u-v&=0 \quad \text{(a)},\\ 2a-u+v&=0 \quad \text{(b)},\\ 2a+u-v&=0 \quad \text{(c)},\\ 2a+u+v&=0 \quad \text{(d)}. \end{aligned} \qquad 5.2\text{-}14$$

Ha egy $\underline{\mathbf{r}}$ helyvektor az 5.2-14 egyenletek bármelyikét kielégíti, akkor $\underline{\mathbf{r}}$ a fenti kúpszeletek valamely pontja lehet.

Megállapíthatjuk, hogy az 5.2-14 (d) egyenlet semmiképpen sem elégíthető ki, ezt az egyenletet csak szimmetriaokokból írtuk fel.

Amennyiben

$$2a>|\underline{\mathbf{f}}_1-\underline{\mathbf{f}}_2|,$$

akkor csak 5.2-14 (a) teljesülhet. Ez pedig egy ellipszis pontjaihoz vezet. Ha

$$2a<|\underline{\mathbf{f}}_1-\underline{\mathbf{f}}_2|,$$

akkor csak 5.2-14 (b) és (c) teljesülhet. E két egyenlet egy hiperbola két ágát határozza meg.

Az 5.2-14 (a)—(d) egyenletek egyetlen egyenletté foghatók össze, ha összeszorozzuk őket. Ha u, v 5.2-14 egyik egyenletének megoldása, akkor

$$(2a-u-v)(2a-u+v)(2a+u-v)(2a+u+v)=0. \qquad 5.2\text{-}15$$

Az 5.2-15 szorzat szimmetrikus eredményre vezet, ha az első és negyedik, valamint a második és harmadik tényezőt összeszorozzuk. A szorzás eredményeként a

$$(4a^2-(u+v)^2)(4a^2-(u-v)^2)=0,$$

illetve a

$$(4a^2-u^2-v^2-2uv)(4a^2-u^2-v^2+2uv)=0$$

eredményhez jutunk. Ebből

$$(4a^2-u^2-v^2)^2-4u^2v^2=0. \qquad 5.2\text{-}16$$

5.2-16 az u, v közötti összefüggést kényelmes formában adja, minthogy racionális kifejezésre vezet. Használjunk olyan koordináta-rendszert, amelyben

$$f_1=0,-f, \quad f_2=0,+f,$$

tehát a gyújtópontok az x_1 tengelyen az origótól szimmetrikusan balra és jobbra helyezkednek el. Ebben a reprezentációban 5.2-12 így írható *(5.5. ábra):*

$$u^2=(x_1+f)^2+x_2^2,$$

$$v^2=(x_1-f)^2+x_2^2,$$

tehát

$$\left.\begin{aligned} u^2+v^2&=2x_1^2+2x_2^2+2f^2\\ u^2v^2&=(x_1^2+x_2^2+f^2)^2-4f^2x_1^2\end{aligned}\right\}. \qquad 5.2\text{-}17$$

Ezerint 5.2-16 a

$$(4a^2-2x_1^2-2x_2^2-2f^2)^2-4(x_1^2+x_2^2+f^2)^2+16f^2x_1^2=0$$

alakot ölti.

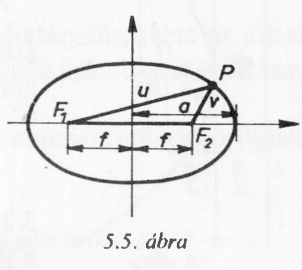
Elvégezve a négyzetre emelést és rendezve az egyenletet, az

$$(a^2-f^2)x_1^2+a^2x_2^2=a^2(a^2-f^2) \qquad 5.2\text{-}18$$

formulához jutunk.

Amennyiben $a^2-f^2\neq 0$,

$$\frac{x_1^2}{a^2}+\frac{x_2^2}{a^2-f^2}=1. \qquad 5.2\text{-}19$$



5.5. ábra

5.2-19 az ellipszis és hiperbola közös egyenlete.

Más formát kapunk, ha az

$$|a^2 - f^2| = b^2$$

jelölést vezetjük be.

Ekkor az ellipszis esetében, ahol $f < a$, azt kapjuk, hogy

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1. \qquad 5.2\text{-}20$$

Hiperbola esetén pedig

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1. \qquad 5.2\text{-}21$$

Az $a = f$ esetben 5.2-18 szerint

$$x_2 = 0.$$

Tehát az ellipszis, illetve hiperbola egyenessé fajul el. Az egyenes a gyújtópontokon halad át. Az F_1F_2 szakaszt elfajult ellipszisnek, az egyenes F_1F_2 szakaszon kívüli részeit az elfajult hiperbola két ágának tekinthetjük.

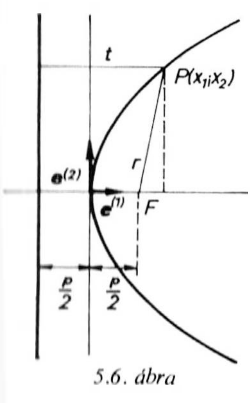
5.2-20-at és 5.2-21-et az ellipszis, illetve hiperbola kanonikus egyenletének nevezzük.

5.2.4. A parabola egyenlete

Befejezésül a parabola egyenletét határozzuk meg. Parabolának azon pontok mértani helyét nevezzük, amelyek egy rögzített ponttól és egy rá nem illeszkedő egyenestől azonos távolságban vannak.

Az egyenest a parabola vezéregyenesének, a pontot a parabola fókuszának nevezzük.

A parabola egyenletét koordinátareprezentációban írjuk fel. Vegyünk fel egy derékszögű koordináta-rendszert úgy, hogy az x_1 tengely legyen merőleges a parabola vezéregyenesére és menjen át a fókuszponton, az origó pedig felezze a vezéregyenes és a fókuszpont távolságát *(5.6. ábra)*.



5.6. ábra

A fókuszpont koordinátáit jelöljük $\frac{p}{2}$, 0-val. Ebben az esetben az origó triviálisan a parabola pontja. Az ábra alapján $|\mathbf{r}| = t$, tehát $p^2 = t^2$, vagyis

$$\left(x_1 + \frac{p}{2}\right)^2 = \left(x_1 - \frac{p}{2}\right)^2 + x_2^2. \qquad 5.2\text{-}19$$

5.2-19-ben a kijelölt műveleteket elvégezve és rendezve, a parabola

$$x_2^2 = 2px_1 \qquad 5.2\text{-}20$$

kanonikus egyenletéhez jutunk.

5.3. Síkbeli polárkoordináták

Síkgörbék ábrázolásakor egyes esetekben hasznos, ha a helyzetvektort nem az x_1, x_2 koordinátákkal, hanem az origótól mért

$$r = \sqrt[+]{x_1^2 + x_2^2}$$

távolsággal, valamint az $\underline{r}$ vektor és az x_1 tengely közötti φ szöggel jellemezzük.[1]

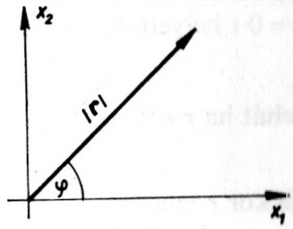
Így tehát

$$x_1 = r\cos\varphi,$$
$$x_2 = r\sin\varphi; \qquad 5.3\text{-}1$$

r-et és φ-t az $\underline{r}$ vektor polárkoordinátáinak nevezzük *(5.7. ábra)*.

Azt is írhatjuk, hogy

$$\mathcal{K}_p(\underline{r}) = r, \varphi.$$



5.7. ábra

A φ szöget az egyértelműség biztosítása céljából a $0-2\pi$ tartományra korlátozhatnánk. Célszerűbb azonban tetszőleges φ értéket megengedni, s ez esetben $\varphi+2\pi, \varphi+4\pi, \ldots, \varphi+2k\pi$ értékek ugyanazt a vektort jellemzik.

Az 5.3-1 formulák megfordításával az

$$r = \sqrt{x_1^2 + x_2^2}, \qquad 5.3\text{-}2$$

$$\varphi = \arccos\frac{x_1}{\sqrt{x_1^2+x_2^2}} = \arcsin\frac{x_2}{\sqrt{x_1^2+x_2^2}} \qquad 5.3\text{-}3$$

képletek adódnak.

5.3-3 esetén azért szükséges a látszólag dupla meghatározás, mert az arc sin és arc cos függvények a $0-2\pi$ tartományban több értékűek, egyértelmű meghatározást csak együttes felhasználásukkal érhetünk el.

A továbbiakban meghatározzuk néhány síkbeli alakzat polárkoordinátás egyenletét.

[1] A gyökjel alatt elhelyezett + jel arra utal, hogy a négyzetgyök pozitív értékét kell venni.

5.3.1. Az egyenes polárkoordinátás egyenlete

Az egyenes 5.2-7 iránytényezős alakját az $m = \operatorname{tg} \alpha$ összefüggés figyelembevételével átrendezhetjük az

$$x_2 \cos \alpha - x_1 \sin \alpha = c \tag{5.3-4}$$

alakra, ahol $c = b \cos \alpha$ állandó mennyiség.

x_1 és x_2 helyére beírva 5.3-1 kifejezéseket, az egyenes egyenlete az

$$r \sin (\varphi - \alpha) = c \tag{5.3-5}$$

alakot ölti. Ha

$$\sin (\varphi - \alpha) \neq 0,$$

tehát

$$\varphi \neq \alpha, \quad \pi + \alpha,$$

akkor oszthatunk $\sin (\varphi - \alpha)$val is, így az

$$r = \frac{c}{\sin (\varphi - \alpha)} \tag{5.3-6}$$

képletet kapjuk. Az 5.3-6 egyenlet olyan görbét ad, amely nem az origón megy keresztül. Az origón áthaladó görbék esetében 5.3-5-re kell visszatérni. Ha $c = 0$-t helyettesítünk be, azt kapjuk, hogy

$$r \sin (\varphi - \alpha) = 0, \tag{5.3-7}$$

tehát ha $r \neq 0$, akkor

$$\sin (\varphi - \alpha) = 0.$$

Ekkor r tetszőleges, $\varphi = \alpha, \pi + \alpha$ így 5.3-7 az origón áthaladó egyenes két félegyenesét adja.

5.3.2. Az ellipszis, hiperbola és parabola polárkoordinátás egyenlete

Az ellipszis és hiperbola közös egyenletét az 5.2-13 egyenletek segítségével

$$\pm u \pm v = 2a \tag{5.3-8}$$

alakban írhatjuk fel. Ha valamelyik gyújtópontot vesszük origónak, és az x_1 tengely a fókuszokat összekötő egyenes, akkor *(5.8., 5.9. ábra)*

$$u = r, \quad v = \sqrt{r^2 + 4f^2 - 4fr \cos \varphi}. \tag{5.3-9}$$

5.3-8-ból következik, hogy

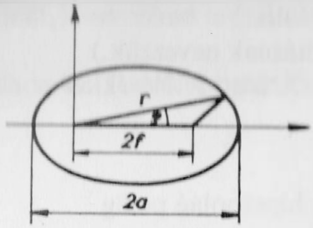
$$v^2 = (2a \mp u)^2,$$

tehát 5.3-9 segítségével

$$r^2+4f^2-4fr\cos\varphi = r^2+4a^2 \mp 4ar,$$

vagyis

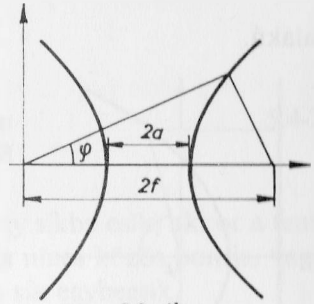
$$r=\frac{a^2-f^2}{\pm a - f\cos\varphi}. \qquad 5.3\text{-}10$$



5.8. ábra

Megjegyezzük, hogy 5.3-10 csak olyan φ értékekre határoz meg pontokat, ahol $r \geqq 0$-nak adódik. Tehát, ha $a > f$, csak a + jel ad pontokat, éspedig

$$r=\frac{a^2-f^2}{a-f\cos\varphi} \geqq 0 \quad \text{minden } \varphi\text{-re}. \qquad 5.3\text{-}11$$



5.9. ábra

Ez az ellipszis egyenlete. 5.3-11-ből következik, hogy *(5.8. ábra)*

$$r=\frac{a^2-f^2}{a-f}=a+f, \quad \text{ha} \quad \varphi=0,$$

ill.

$$r=\frac{a^2-f^2}{a+f}=a-f, \quad \text{ha} \quad \varphi=\pi.$$

Tehát r az $a-f$ és $a+f$ értékek között ingadozik.

Az $a<f$ esetekre *(5.9. ábra)*

$$r=\frac{a^2-f^2}{\pm a-f\cos\varphi}>0, \quad \text{ha} \quad \cos\varphi > +\frac{a}{f}, \qquad 5.3\text{-}12$$

— ebben a tartományban a rádiuszvektor a hiperbola mindkét ágát metszi — és

$$r=\frac{-(a^2-f^2)}{-a-f\cos\varphi}>0, \quad \text{ha} \quad -\frac{a}{f}<\cos\varphi<\frac{a}{f}, \qquad 5.3\text{-}13$$

ebben a tartományban a rádiuszvektor a hiperbolának csak egyik ágát metszi.

Ezeket az egyenleteket gyakran felhasználjuk például a csillagászatban. Az 5.3-11 kifejezés adja ugyanis egy bolygó pályáját a Naptól való távolság és az azimut függvényeként.

Szokásos még az 5.3-11, 5.3-12 és 5.3-13 egyenletek számlálóját és nevezőjét is a-val végigosztani, valamint a

$$b^2=|a^2-f^2|, \quad p=\frac{a^2-f^2}{a} \quad \text{és} \quad e=\frac{f}{a}$$

jelöléseket bevezetni. (p-t a görbe paraméterének, e-t pedig numerikus excentricitásnak nevezzük.)

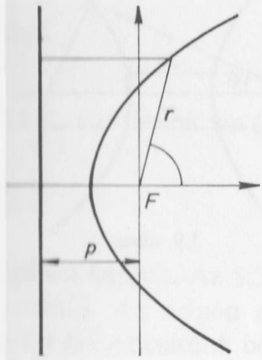
A fenti jelölésekkel az ellipszis egyenlete

$$r=\frac{p}{1-e\cos\varphi}, \qquad 5.3\text{-}14$$

a hiperboláé pedig

$$r=\frac{p}{\pm 1-e\cos\varphi} \qquad 5.3\text{-}15$$

alakú.



5.10. ábra

Az *5.10. ábra* és a parabola definíciójának felhasználásával, koordináta-rendszerünk origóját ismét a fókuszpontba helyezve, a parabola egyenlete az

$$r=p+r\cos\varphi \qquad 5.3\text{-}16$$

alakban írható fel. Kifejezve r-et:

$$r=\frac{p}{1-\cos\varphi}. \qquad 5.3\text{-}17$$

Látható tehát, hogy a három görbe egyenlete egyöntetűen az

$$r=\frac{p}{\pm 1-e\cos\varphi} \qquad 5.3\text{-}18$$

formában írható — ahol a nevezőben szereplő kettős előjel csak a hiperbola esetén érvényes —, amikor is

	p	e
ellipszis, ha	>0,	<1,
hiperbola, ha	<0,	>1,
parabola, ha	>0,	$=1$.

Megjegyezzük még, hogy 5.3-14 $f=0$ esetén triviálisan egy origó középpontú kör egyenlete.

5.4. Három sík közös pontjának meghatározása

Legyen a három sík egyenlete

$$\underline{\mathbf{K}}_1\underline{\mathbf{r}}=\alpha_1,$$
$$\underline{\mathbf{K}}_2\underline{\mathbf{r}}=\alpha_2, \qquad 5.4\text{-}1$$
$$\underline{\mathbf{K}}_3\underline{\mathbf{r}}=\alpha_3.$$

Keressük azt az $\underline{\mathbf{r}}_0$ helyvektorú pontot, amely mindhárom síkon rajta van, tehát amely mindhárom egyenletet kielégíti.

Amennyiben a $\underline{\mathbf{K}}_1$, $\underline{\mathbf{K}}_2$, $\underline{\mathbf{K}}_3$ nem komplanárisak, akkor 4.8-11-nek megfelelően definiálhatjuk a

$$\underline{\mathbf{k}}_l = \varepsilon_{lmn} = (\underline{\mathbf{K}}_m \times \underline{\mathbf{K}}_n)/V$$

reciprok vektorokat, ahol

$$V = \underline{\mathbf{K}}_1(\underline{\mathbf{K}}_2 \times \underline{\mathbf{K}}_3).$$

Minthogy

$$\underline{\mathbf{K}}_j \underline{\mathbf{k}}_l = \delta_{jl},$$

látjuk, hogy az

$$\underline{\mathbf{r}}_0 = \alpha_1 \underline{\mathbf{k}}_1 + \alpha_2 \underline{\mathbf{k}}_2 + \alpha_3 \underline{\mathbf{k}}_3 \qquad 5.4\text{-}2$$

vektor 5.4-1 megoldása.

Abban az esetben, ha a három $\underline{\mathbf{K}}_i$ normálvektor egy síkba esik, akkor a fenti módszer nem alkalmazható, és a három síknak vagy nincs közös pontja, vagy egy egyenesben metszik egymást, illetve mindhárom sík egybeesik.

5.5. Sík és egyenes metszéspontja

Legyen

$$\underline{\mathbf{r}} = \underline{\mathbf{a}} + p\underline{\mathbf{K}} + q\underline{\mathbf{L}} \qquad 5.5\text{-}1$$

egy sík,

$$\underline{\mathbf{r}} = \underline{\mathbf{b}} - s\underline{\mathbf{M}} \qquad 5.5\text{-}2$$

pedig egy egyenes paraméteres egyenlete. (A negatív előjelet a későbbi egyszerűbb jelölések kedvéért vezettük be.) A két alakzat közös pontja eleget tesz az

$$\underline{\mathbf{a}} + p\underline{\mathbf{K}} + q\underline{\mathbf{L}} = \underline{\mathbf{b}} - s\underline{\mathbf{M}}, \qquad 5.5\text{-}3$$

illetve

$$p\underline{\mathbf{K}} + q\underline{\mathbf{L}} + s\underline{\mathbf{M}} = \underline{\mathbf{b}} - \underline{\mathbf{a}} \qquad 5.5\text{-}4$$

egyenletnek.

Amennyiben $\underline{\mathbf{K}}$, $\underline{\mathbf{L}}$, $\underline{\mathbf{M}}$ nem egy síkba esnek, ismét bevezethetjük a $\underline{\mathbf{k}}$, $\underline{\mathbf{l}}$, $\underline{\mathbf{m}}$ reciprok vektorokat, és így

$$p = (\underline{\mathbf{b}} - \underline{\mathbf{a}})\underline{\mathbf{k}}; \quad q = (\underline{\mathbf{b}} - \underline{\mathbf{a}})\underline{\mathbf{l}}; \quad s = (\underline{\mathbf{b}} - \underline{\mathbf{a}})\underline{\mathbf{m}}. \qquad 5.5\text{-}5$$

Amennyiben a

$$\det(\underline{\mathbf{K}}, \underline{\mathbf{L}}, \underline{\mathbf{M}}) \neq 0$$

feltétel nem teljesül, az egyenesnek és a síknak vagy nincs egyetlen közös pontja sem, vagy a teljes egyenes a síkban fekszik.

A következőkben néhány eddig tárgyalt alakzat távolságának analitikus meghatározásával foglalkozunk.

5.6.1. Két pont távolsága

Két pont távolsága a

$$d = |\underline{\mathbf{r}}_2 - \underline{\mathbf{r}}_1| \tag{5.6-1}$$

formulával adható meg. Koordinátás felbontásban

$$\mathcal{K}(\underline{\mathbf{r}}_2) = \mathbf{r}_2 = x_1, x_2, x_3; \quad \mathcal{K}(\underline{\mathbf{r}}_1) = \mathbf{r}_1 = y_1, y_2, y_3;$$

$$d = \underset{+}{\sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2}}. \tag{5.6-2}$$

5.6.2. Két párhuzamos sík távolsága

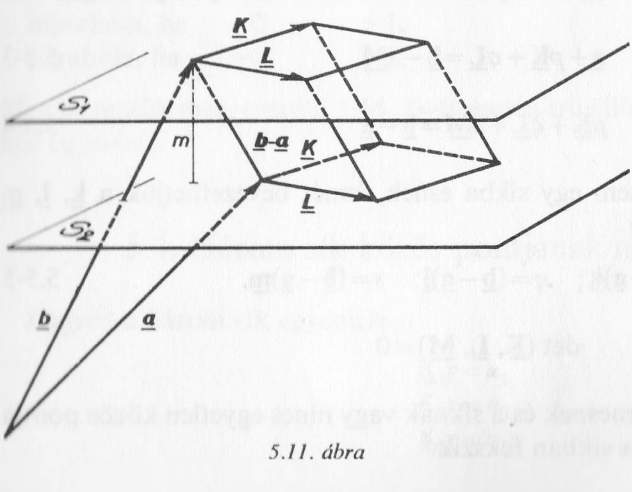
Legyen a két, $\mathcal{S}_1$ és $\mathcal{S}_2$ sík egyenlete

$$\mathcal{S}_1: \quad \underline{\mathbf{r}} = \underline{\mathbf{a}} + p\underline{\mathbf{K}} + q\underline{\mathbf{L}} \tag{5.6-3}$$

és

$$\mathcal{S}_2: \quad \underline{\mathbf{r}} = \underline{\mathbf{b}} + p\underline{\mathbf{K}} + q\underline{\mathbf{L}}. \tag{5.6-4}$$

Mivel az 5.6-3-ban és 5.6-4-ben szereplő $\underline{\mathbf{K}}$, $\underline{\mathbf{L}}$ azonos, a két sík párhuzamos egymással. A két sík távolságát a két síkra merőleges egyenesnek a síkok közé eső szakasza adja meg.



5.11. ábra

Konstruáljunk egy paralelepipedont, amelynek oldalai $\underline{\mathbf{b}} - \underline{\mathbf{a}}$, $\underline{\mathbf{K}}$, $\underline{\mathbf{L}}$ *(5.11. ábra)*.

Az 5.11. ábráról leolvasható, hogy ezen paralelepipedon a két sík közé simul.

A paralelepipedon térfogatát a 2.7. fejezet alapján a

$$V = |(\underline{\mathbf{b}} - \underline{\mathbf{a}}, \underline{\mathbf{K}}, \underline{\mathbf{L}})| \tag{5.6-5}$$

vegyes szorzat abszolút értéke adja. Az S_1, ill. S_2 síkban fekvő oldallapok területe

$$T=|\underline{\mathbf{K}}\times\underline{\mathbf{L}}|,$$

és mivel

$$V=T\cdot m,$$

a két sík távolsága az

$$m=\frac{|(\underline{\mathbf{b}}-\underline{\mathbf{a}},\ \underline{\mathbf{K}},\ \underline{\mathbf{L}})|}{|\underline{\mathbf{K}}\times\underline{\mathbf{L}}|} \qquad 5.6\text{-}6$$

formulával határozható meg.

5.6.3. Kitérő egyenesek távolsága

Legyen a két s_1 és s_2 kitérő egyenes egyenlete

$$s_1\colon\quad \underline{\mathbf{r}}=\underline{\mathbf{a}}+p\underline{\mathbf{K}}, \qquad 5.6\text{-}7$$

$$s_2\colon\quad \underline{\mathbf{r}}=\underline{\mathbf{b}}+q\underline{\mathbf{L}}, \qquad 5.6\text{-}8$$

és tegyük fel, hogy $\underline{\mathbf{K}}\times\underline{\mathbf{L}}=\underline{\mathbf{M}}\neq\underline{\mathbf{0}}$, tehát az egyenesek nem párhuzamosak. A két egyenes tetszés szerinti P és Q pontját összekötő vektor

$$\overrightarrow{PQ}=\underline{\mathbf{r}}(P)-\underline{\mathbf{r}}(Q)=\underline{\mathbf{a}}-\underline{\mathbf{b}}+p_0\underline{\mathbf{K}}-q_0\underline{\mathbf{L}} \qquad 5.6\text{-}9$$

(p_0 és q_0 jelöli a két egyenes P és Q pontjához tartozó paraméterértéket).

A kitérő egyenesek távolsága definíció szerint a két egyenes által a mindkettőre merőleges egyenesből kivágott szakasz hosszával egyenlő. (Kimutatható, hogy a két kitérő egyenes pontjait összekötő szakaszok között ez a legrövidebb.)

Szükségünk van tehát arra a $\overrightarrow{PQ}$ vektorra, amely mindkét egyenesre merőleges. Ez a szakasz párhuzamos a $\underline{\mathbf{K}}\times\underline{\mathbf{L}}$ vektorral, vagyis

$$\overrightarrow{PQ}=\underline{\mathbf{r}}(P)-\underline{\mathbf{r}}(Q)=S(\underline{\mathbf{K}}\times\underline{\mathbf{L}}), \qquad 5.6\text{-}10$$

ahol S egyelőre ismeretlen arányossági tényező. Beírva ezt az 5.6-9 egyenletbe, az

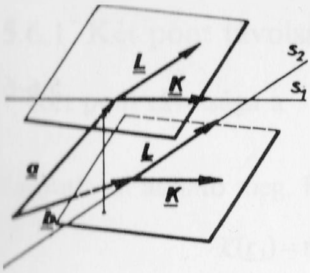
$$S(\underline{\mathbf{K}}\times\underline{\mathbf{L}})=\underline{\mathbf{a}}-\underline{\mathbf{b}}+p_0\underline{\mathbf{K}}+q_0\underline{\mathbf{L}} \qquad 5.6\text{-}11$$

egyeneshez jutunk, amely az S, p_0, q_0 ismeretlenekre megoldható. Szorozzuk be 5.6-11-et $\underline{\mathbf{K}}\times\underline{\mathbf{L}}$-lel skalárisan, és fejezzük ki S-et

$$S=\frac{|((\underline{\mathbf{a}}-\underline{\mathbf{b}}),\ \underline{\mathbf{K}},\ \underline{\mathbf{L}})|}{|(\underline{\mathbf{K}}\times\underline{\mathbf{L}})|^2}\ .$$

A kitérő egyenesek d távolságát az $S(\underline{\mathbf{K}}\times\underline{\mathbf{L}})$ vektor abszolút értéke adja, tehát

$$d=\frac{|((\underline{\mathbf{a}}-\underline{\mathbf{b}}),\underline{\mathbf{K}},\underline{\mathbf{L}})|}{|\underline{\mathbf{K}}\times\underline{\mathbf{L}}|}, \qquad 5.6\text{-}12$$



5.12. ábra

Vizsgáljuk meg most a fentiek geometriai jelentését *(5.12. ábra)*. Vegyük fel az első egyenes $\underline{\mathbf{a}}$ és a második egyenes $\underline{\mathbf{b}}$ pontjában az $\underline{\mathbf{L}}$, ill. $\underline{\mathbf{K}}$ vektorokat.

Ezek a vektorok két egymással párhuzamos síkot feszítenek ki, amelyek tartalmazzák a kitérő egyeneseket. Mivel mindkét sík merőleges a $\underline{\mathbf{K}}\times\underline{\mathbf{L}}$ vektorra, nyilvánvaló, hogy a kitérő egyenesek távolsága megegyezik a két sík távolságával. A két sík távolsága pedig 5.6-6 alapján

$$d=\frac{|((\underline{\mathbf{a}}-\underline{\mathbf{b}}),\underline{\mathbf{K}},\underline{\mathbf{L}})|}{|\underline{\mathbf{K}}\times\underline{\mathbf{L}}|},$$

ami megegyezik előző eredményünkkel.

5.6.4. Pont és sík távolsága

Pont és sík távolsága definíció szerint a pontból a síkra bocsátott merőleges szakasz hossza. Legyen a P pont helyvektora $\underline{\mathbf{a}}$, az $\mathcal{S}$ sík egyenlete pedig

$$\underline{\mathbf{r}}=\underline{\mathbf{b}}+p\underline{\mathbf{K}}+q\underline{\mathbf{L}}.$$

Fektessünk az $\underline{\mathbf{a}}$ helyvektorú P ponton keresztül az $\mathcal{S}$ síkkal párhuzamos síkot. Az így kapott $\mathcal{S}'$ sík egyenlete

$$\underline{\mathbf{r}}=\underline{\mathbf{a}}+p\underline{\mathbf{K}}+q\underline{\mathbf{L}}.$$

A két sík távolsága megegyezik a pont és a sík távolságával, tehát a távolság a

$$d=\frac{|((\underline{\mathbf{a}}-\underline{\mathbf{b}}),\underline{\mathbf{K}},\underline{\mathbf{L}})|}{|\underline{\mathbf{K}}\times\underline{\mathbf{L}}|} \qquad 5.6\text{-}13$$

formulával fejezhető ki.

5.6.5. Pont és egyenes távolsága

Legyen a P pont helyvektora $\mathbf{a}$, és az egyenes egyenlete

$$\mathbf{r}=\mathbf{b}+p\mathbf{K}.$$

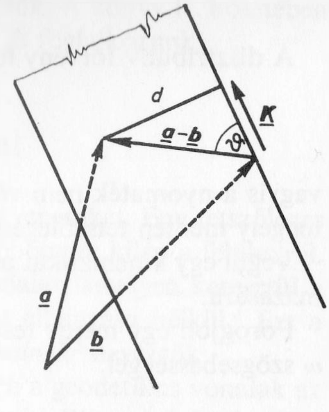
Amennyiben a P pont nincs az egyenesen, akkor az egyenes és a pont egy síkot határoz meg. Mind a $\mathbf{K}$, mind pedig az $\mathbf{a}-\mathbf{b}$ vektor benne van ebben a síkban, tehát a sík egyenlete felírható az

$$\mathbf{r}=\mathbf{b}+p\mathbf{K}+q(\mathbf{a}-\mathbf{b}) \qquad 5.6\text{-}14$$

alakban. A pont és egyenes távolságát a pontból az egyenesre bocsátott merőleges szakasz hossza adja meg. Az *5.13. ábra* alapján

$$d=|\mathbf{a}-\mathbf{b}|\sin\vartheta,$$

ahol ϑ a $\mathbf{K}$ és $\mathbf{a}-\mathbf{b}$ vektorok által bezárt szög.



5.13. ábra

5.6.6. Alkalmazások

Vizsgáljunk egy, az origó körül szabadon elforgatható merev testet, amelyre az $\mathbf{r}_0$ helyvektorú P pontban egy $\mathbf{F}$ erő hat.

Az $\mathbf{F}$ erő forgatónyomatéka az origóra vonatkozóan definíció szerint

$$\mathbf{M}=\mathbf{r}_0\times\mathbf{F}. \qquad 5.6\text{-}15$$

Ha a test nem forgatható szabadon, hanem csak egy az origón átmenő $\mathbf{k}$ egységvektorral meghatározott irányú tengely körül forgatható el, akkor a tengelyre ható nyomaték

$$M_k=\mathbf{k}\mathbf{M}=\mathbf{k}(\mathbf{r}_0\times\mathbf{F})=(\mathbf{k},\,\mathbf{r}_0,\,\mathbf{F}). \qquad 5.6\text{-}16$$

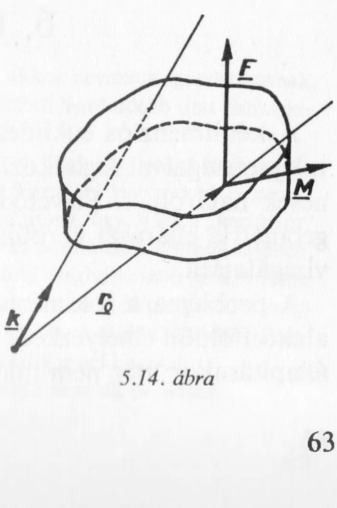
Ha az erőt hatásvonala mentén eltoljuk, akkor $\mathbf{r}_0$ helyett az

$$\mathbf{r}=\mathbf{r}_0+q\mathbf{F} \qquad 5.6\text{-}17$$

helyvektorral kell számolnunk *(5.14. ábra)*.

Amennyiben a nyomatékot nem az origóra, hanem a tengely más pontjára számítjuk, akkor az erő támadáspontja és a forgási centrum közötti vektor

$$\mathbf{r}'=\mathbf{r}_0+q\mathbf{F}+p\mathbf{k}. \qquad 5.6\text{-}18$$



5.14. ábra

Itt p és q kétféle eltolást jellemző paraméter. Az eltolt erőnek az eltolt forgási centrumra ható nyomatéka

$$(\underline{\mathbf{k}}, \underline{\mathbf{r}}', \underline{\mathbf{F}}) = (\underline{\mathbf{k}}, (\underline{\mathbf{r}}_0 + p\underline{\mathbf{k}} + q\underline{\mathbf{F}}), \underline{\mathbf{F}}). \qquad 5.6\text{-}19$$

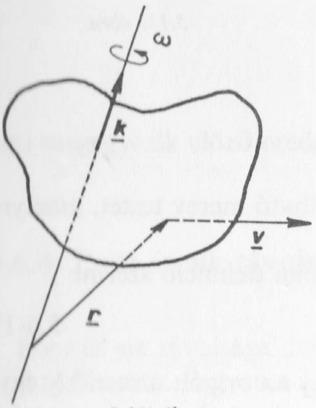
A disztributív törvény felhasználásával adódik, hogy

$$(\underline{\mathbf{r}}_0, \underline{\mathbf{k}}, \underline{\mathbf{F}}) = (\underline{\mathbf{r}}', \underline{\mathbf{k}}, \underline{\mathbf{F}}), \qquad 5.6\text{-}20$$

vagyis a nyomaték nem változik, ha az erőt hatásvonala, ill. a forgáspontot a tengely mentén tetszőlegesen eltoljuk.

Végül egy kinematikai példát említünk a vektoriális szorzat gyakorlati alkalmazására.

Forogjon egy merev test egy $\underline{\mathbf{k}}$ egységvektor irányába mutató tengely körül ω szögsebességgel.



5.15. ábra

A forgást jellemezhetjük az

$$\underline{\boldsymbol{\omega}} = \omega\underline{\mathbf{k}} \qquad 5.6\text{-}21$$

szögsebességvektorral is. (A vektor iránya a forgási tengely irányát, nagysága pedig a szögsebességet adja meg.)

Ha a tengely az origón megy keresztül, akkor az $\underline{\mathbf{r}}$ koordinátavektorral jellemzett pont sebességét a

$$\underline{\mathbf{v}} = \underline{\boldsymbol{\omega}} \times \underline{\mathbf{r}} \qquad 5.6\text{-}22$$

formula adja meg *(5.15. ábra)*.

6. Gömbi geometria

A kétdimenziós euklideszi geometria a síkfelületeken elhelyezkedő idomok tulajdonságaival foglalkozik. Az egyszerűbb alakzatokat ekkor pontok és egyenesek határolják. Felvetődik a kérdés, hogyan általánosíthatók az euklideszi geometria eljárásai görbült felületeken elhelyezkedő idomok tulajdonságainak vizsgálatára.

A probléma a földméréssel kapcsolatban merült fel, ugyanis a közel gömb alakú Földön elhelyezkedő idomok méretei közötti pontos összefüggések megállapításakor már nem mindig elegendő a Föld felszínének síkkal történő közelítése. Jelentkezik ez a probléma a csillagászatban is. Az égitestek látszólagos távolságai közötti összefüggések megállapítása az éggömbön fekvő pontok geometriájának kidolgozását tette szükségessé.

A következőkben a gömbi geometriával foglalkozunk. A könyv II. kötetében tetszőleges görbült felület geometriájának leírásával is foglalkozunk.

6.1. A geodetikus vonal

A görbült felületek általában nem tartalmaznak egyeneseket. Egy tetszőleges $\mathcal{S}$ görbe felület A és B pontjait összekötő egyenes többnyire kilép a felületből. A két pont között azonban meghúzhatjuk azt a vonalat, amelyen keresztül a legrövidebb úton juthatunk A-ból B-be az $\mathcal{S}$ felület elhagyása nélkül.[1] Ezt a görbét az A és B pontokat összekötő *geodetikus vonalnak*[2] nevezzük.

A későbbiek során látni fogjuk, hogy az $\mathcal{S}$ felületen a geodetikus vonalak az egyenesekhez hasonló tulajdonságokkal rendelkeznek. Bizonyos értelemben tehát a geodetikus vonalak a görbült felületek „egyeneseinek" tekinthetők.

Gömbfelületen a geodetikus vonalak a gömb főkörei. Az O középpontú gömb A és B pontját összekötő geodetikus vonalat tehát a gömb és az AOB sík metszetgörbéje szolgáltatja.

6.2. A gömbháromszög

Metssze egymást a g_1 és g_2 geodetikus vonal a P pontban. Jelöljük a g_1 és g_2 görbék P pontban vett érintőit t_1-gyel és t_2-vel. Az érintők hajlásszögét definíciószerűen a két geodetikus vonal hajlásszögének tekintjük, tehát

$$(g_1, g_2)\sphericalangle = (t_1, t_2)\sphericalangle .$$

Legyen g_1, g_2, g_3 három geodetikus vonal, amelyek legalább három metszésponttal rendelkeznek, és jelöljük

[1] Pontosabban fogalmazva az AB pontokat összekötő vonalat akkor nevezzük geodetikusnak, ha a vonal tetszőleges $P_1 P_2$ szakaszára érvényes, hogy a P_1 és P_2 közötti legrövidebb utat határozza meg.

A definíció általánosítása egyes esetekben arra az eredményre vezet, hogy két pontot több geodetikus vonallal is összeköthetünk. Pl. egy gömbfelület A és B pontját összekötő legrövidebb ív geodetikus vonal, azonban az A pontból B-be a gömböt megkerülve is eljuthatunk, úgy, hogy a „kerülő út" nem túlságosan távol fekvő pontjait összekötő ívdarabok az adott két pont között a legrövidebbek.

A gömbfelület két pontját összekötő legrövidebb ív a két pont által meghatározott gömbi főkör egy íve. A fenti definíció értelmében azonban geodetikus vonalnak tekintjük a főkör gömböt megkerülő ívét is. Ez utóbbi ív már nem a legrövidebb az adott két pont között. A gömbi főkörön választott ív csak akkor a legrövidebb két pont között, ha a hozzá tartozó szög π-nél kisebb.

[2] Az elnevezés utal arra, hogy a fogalmat a földmérésben (geodézia) vezették be először.

g_2g_3 egyik metszéspontját A-val,
g_3g_1 egyik metszéspontját B-vel,
g_1g_2 egyik metszéspontját C-vel.

Az ABC pontok a gömbön egy háromszöget alkotnak, amelynek oldalai

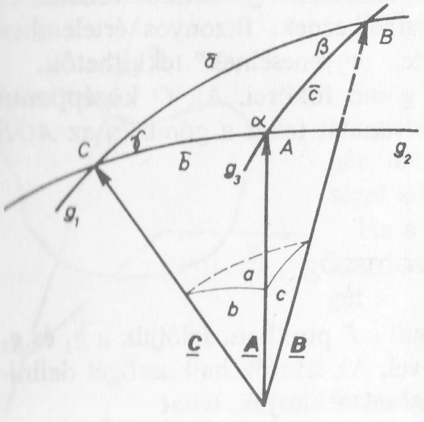
$$\hat{a}=\widehat{BC} \quad \text{a } g_1 \text{ ívdarabja,}$$

$$\hat{b}=\widehat{CA} \quad \text{a } g_2 \text{ ívdarabja,}$$

$$\hat{c}=\widehat{AB} \quad \text{a } g_3 \text{ ívdarabja.}$$

A háromszög szögei pedig

$$\alpha=(g_2, g_3)\sphericalangle, \quad \beta=(g_3, g_1)\sphericalangle, \quad \gamma=(g_2, g_1)^1\sphericalangle.$$



6.1. ábra

A gömbháromszög oldalait alkotó ívek egyszerűen kifejezhetők az alábbi formában *(6.1. ábra)*.

$$\hat{a}=Ra, \quad \hat{b}=Rb, \quad \hat{c}=Rc, \qquad 6.2\text{-}1$$

ahol R a gömb sugara és

$$a=(COB)\sphericalangle, \quad b=(COA)\sphericalangle,$$
$$c=(AOB)\sphericalangle. \qquad 6.2\text{-}2$$

A gömbháromszögek tárgyalásakor tehát az oldalakat célszerű a megfelelő főkörívekhez tartozó középponti szögekkel kifejezni.

6.3. A gömbháromszög trigonometriája

A következőkben a gömbháromszög adatai közötti összefüggéseket határozunk meg.

Az ABC gömbháromszöget egyértelműen jellemezhetjük a gömb középpontjából az A, B, C csúcsok irányába mutató tetszőleges hosszúságú $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$, $\underline{\mathbf{C}}$ vektorokkal. Az előzőekben megadott definíciók szerint

[1] Megjegyezzük, hogy a fenti jelöléseket ciklikusan permutálva, a harmadik szögre $\gamma'=(g_1, g_2)\sphericalangle$-t kapnánk. Azonban $\gamma'=\pi-\gamma$ a háromszög külső szöge. Amennyiben a belső szögekre akarunk szorítkozni, akkor a fenti aszimmetrikus írásmódot kell használnunk.

$$\left.\begin{aligned}
a &= (\underline{\mathbf{B}}, \underline{\mathbf{C}})\sphericalangle, &&\text{tehát} \quad \cos a = \frac{(\underline{\mathbf{B}}\underline{\mathbf{C}})}{BC}, \\
b &= (\underline{\mathbf{C}}, \underline{\mathbf{A}})\sphericalangle, &&\text{tehát} \quad \cos b = \frac{(\underline{\mathbf{C}}\underline{\mathbf{A}})}{CA}, \\
c &= (\underline{\mathbf{A}}, \underline{\mathbf{B}})\sphericalangle, &&\text{tehát} \quad \cos c = \frac{(\underline{\mathbf{A}}\underline{\mathbf{B}})}{AB},
\end{aligned}\right\} \qquad 6.3\text{-}1$$

ahol

$$A = |\underline{\mathbf{A}}|, \quad B = |\underline{\mathbf{B}}|, \quad C = |\underline{\mathbf{C}}|.$$

Hasonló módon

$$\left.\begin{aligned}
\sin a &= \frac{|\underline{\mathbf{B}} \times \underline{\mathbf{C}}|}{BC}, \\
\sin b &= \frac{|\underline{\mathbf{C}} \times \underline{\mathbf{A}}|}{AB}, \\
\sin c &= \frac{|\underline{\mathbf{B}} \times \underline{\mathbf{A}}|}{AB}.
\end{aligned}\right\} \qquad 6.3\text{-}2$$

A gömbháromszög szögei azonosak az oldalakat kimetsző főkörök síkjainak szögével. Így például

$$\cos \gamma = \frac{(\underline{\mathbf{C}} \times \underline{\mathbf{A}})(\underline{\mathbf{C}} \times \underline{\mathbf{B}})}{|\underline{\mathbf{C}} \times \underline{\mathbf{A}}||\underline{\mathbf{C}} \times \underline{\mathbf{B}}|},$$

$$\sin \gamma = \frac{|(\underline{\mathbf{C}} \times \underline{\mathbf{A}}) \times (\underline{\mathbf{C}} \times \underline{\mathbf{B}})|}{|\underline{\mathbf{C}} \times \underline{\mathbf{A}}||\underline{\mathbf{C}} \times \underline{\mathbf{B}}|}. \qquad 6.3\text{-}3$$

Az α és β szög megfelelő szögfüggvényei a 6.3-3 formulákból a betűk ciklikus permutációjával állíthatók elő.

A vektoriális négyesszorzatra vonatkozó 4.8-6 azonosság alapján

$$(\underline{\mathbf{C}} \times \underline{\mathbf{A}}) \times (\underline{\mathbf{C}} \times \underline{\mathbf{B}}) = \underline{\mathbf{C}}(\underline{\mathbf{A}}, \underline{\mathbf{B}}, \underline{\mathbf{C}}), \qquad 6.3\text{-}4$$

tehát

$$|(\underline{\mathbf{C}} \times \underline{\mathbf{A}}) \times (\underline{\mathbf{C}} \times \underline{\mathbf{B}})| = C(\underline{\mathbf{A}}, \underline{\mathbf{B}}, \underline{\mathbf{C}}). \qquad 6.3\text{-}5$$

Így 6.3-2 és 6.3-3 alapján

$$\frac{\sin c}{\sin \gamma}=\frac{|\underline{\mathbf{B}}\times\underline{\mathbf{A}}|\;|\underline{\mathbf{C}}\times\underline{\mathbf{A}}|\;|\underline{\mathbf{C}}\times\underline{\mathbf{B}}|}{ABC(\underline{\mathbf{A}},\underline{\mathbf{B}},\underline{\mathbf{C}})}. \qquad 6.3\text{-}6$$

A 6.3-6 kifejezés nem változik, ha a jobb oldalán az $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$, $\underline{\mathbf{C}}$ vektorokat ciklikusan permutáljuk. Következésképpen

$$\frac{\sin\alpha}{\sin a}=\frac{\sin\beta}{\sin b}=\frac{\sin\gamma}{\sin c}. \qquad 6.3\text{-}7$$

A 6.3-7 összefüggést a gömbháromszögekre vonatkozó sinustételnek nevezzük.

A *cosinustétel* megfelelőjét az alábbi módon kaphatjuk meg. 4.8-8 alapján

$$(\underline{\mathbf{C}}\times\underline{\mathbf{A}})(\underline{\mathbf{C}}\times\underline{\mathbf{B}})=\underline{\mathbf{A}}\underline{\mathbf{B}}C^2-(\underline{\mathbf{C}}\underline{\mathbf{B}})(\underline{\mathbf{C}}\underline{\mathbf{A}}). \qquad 6.3\text{-}8$$

6.3-8-at $|\underline{\mathbf{C}}\times\underline{\mathbf{A}}|\;|\underline{\mathbf{C}}\times\underline{\mathbf{B}}|$-vel elosztva 6.3-1 és 6.3-2 felhasználásával a

$$\cos\gamma=\frac{\cos c}{\sin b\sin a}-\frac{\cos a\cos b}{\sin b\sin a} \qquad 6.3\text{-}9$$

képlethez, és ebből a

$$\cos c=\cos a\,\cos b+\sin a\sin b\cos\gamma$$

összefüggéshez jutunk. *A 6.3-9 egyenlet a gömbháromszög oldalaira vonatkozó cosinustételt fejezi ki.*

6.4. A polár-gömbháromszög

Tekintsük az $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$, $\underline{\mathbf{C}}$ nem komplanáris vektorok által meghatározott gömbháromszöget, amelyre

$$\underline{\mathbf{A}}(\underline{\mathbf{B}}\times\underline{\mathbf{C}})=V>0. \qquad 6.4\text{-}1$$

(Ha $V\neq 0$, az $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$, $\underline{\mathbf{C}}$ vektorok sorrendjének megfelelő választásával mindig elérhetjük, hogy $V>0$ teljesüljön.)

Rendeljünk hozzá az $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$, $\underline{\mathbf{C}}$ vektorok által kifeszített gömbháromszöghöz egy újabb háromszöget, amelynek csúcsait az

$$\underline{\mathbf{A}}^*=\underline{\mathbf{B}}\times\underline{\mathbf{C}},\quad \underline{\mathbf{B}}^*=\underline{\mathbf{C}}\times\underline{\mathbf{A}}\quad\text{és}\quad \underline{\mathbf{C}}^*=\underline{\mathbf{A}}\times\underline{\mathbf{B}} \qquad 6.4\text{-}2$$

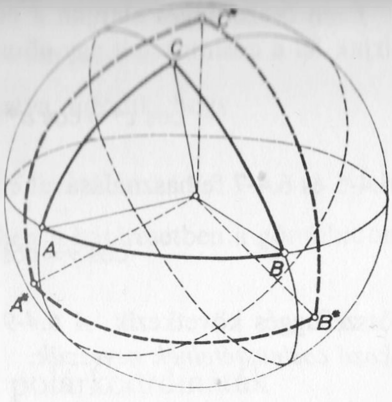
vektorok egyenesei határozzák meg. Az $\underline{A}^*$, $\underline{B}^*$, $\underline{C}^*$ vektorokkal meghatározott háromszöget az eredeti háromszög polár-gömbháromszögének nevezzük *(6.2. ábra)*.

A polár-gömbháromszöget tehát az eredeti gömbháromszöget kimetsző főkörök síkjainak normálvektorai határozzák meg.

Amennyiben 6.4-1 teljesül, akkor a 6.4-2 összefüggések a következőképpen fordíthatók meg:

$$\underline{A}=\frac{1}{V}(\underline{B}^*\times\underline{C}^*),\quad \underline{B}=\frac{1}{V}(\underline{C}^*\times\underline{A}^*),$$

$$\underline{C}=\frac{1}{V}(\underline{A}^*\times\underline{B}^*). \qquad 6.4\text{-}3$$



6.2. ábra

Minthogy a gömbháromszög oldalait jellemző szögek és a gömbháromszög szögei függetlenek a gömb sugarától, s így a csúcsokat kijelölő vektorok hosszától is, 6.4-3-ból következik, hogy az $\underline{A}^*$, $\underline{B}^*$, $\underline{C}^*$ vektorok által kijelölt gömbháromszög polárháromszögét az $\underline{A}$, $\underline{B}$, $\underline{C}$ vektorok jelölik ki.

A két háromszög adatai között fennállnak a következő összefüggések:

$$\cos c^*=\frac{(\underline{A}^*\times\underline{C}^*)(\underline{B}^*\times\underline{C}^*)}{|\underline{A}^*\times\underline{C}^*||\underline{B}^*\times\underline{C}^*|},$$

és így 6.4-3 miatt

$$\cos c^*=\frac{\underline{B}\underline{A}}{BA}=\cos\gamma. \qquad 6.4\text{-}4$$

Tehát

$$\gamma=c^*,$$

és hasonló módon

$$\left.\begin{aligned}\gamma&=c^*,\\ \alpha&=a^*,\ \beta=b^*.\end{aligned}\right\} \qquad 6.4\text{-}5$$

Az inverz transzformációból kiindulva pedig

$$\cos c=\frac{(\underline{A}\times\underline{C})(\underline{B}\times\underline{C})}{|\underline{A}\times\underline{C}||\underline{B}\times\underline{C}|}=\frac{\underline{A}^*\underline{B}^*}{A^*B^*}\cos\gamma^*. \qquad 6.4\text{-}6$$

Tehát

$$\alpha^*=a,\quad \beta^*=b,\quad \gamma^*=c. \qquad 6.4\text{-}7$$

A polár-gömbháromszög oldalai és szögei tehát rendre egyenlőek az eredeti háromszög szögeivel és oldalaival.

Ezen összefüggés alapján a cosinustétel egy fontos változatához juthatunk. Írjuk fel a cosinustételt egy gömbháromszög polárháromszögére:

$$\cos c^* = \cos a^* \cos b^* - \sin a^* \sin b^* \cos \gamma^*. \qquad 6.4\text{-}8$$

6.4-5 és 6.4-7 felhasználásával ebből az eredeti háromszögre a

$$\cos \gamma = \cos \alpha \cos \beta - \sin \alpha \sin \beta \cos c \qquad 6.4\text{-}9$$

összefüggés következik. *A 6.4-9 kapcsolatot a gömbháromszög szögeire vonatkozó cosinustételnek nevezzük.*

6.5. Egy határeset

A gömbháromszögeket síkháromszögekkel közelítjük, ha oldalaik sokkal kisebbek, mint a gömb sugara. Vizsgáljuk meg az

$$a = \frac{\widehat{a}}{R}, \quad b = \frac{\widehat{b}}{R}, \quad c = \frac{\widehat{c}}{R} \qquad 6.5\text{-}1$$

oldalakkal meghatározott gömbháromszöget az $R \to \infty$ határátmenet esetén.

A kis szögek szögfüggvényeit ez esetben helyettesíthetjük az alábbi közelítésekkel:

$$\sin \delta \sim \delta, \quad \cos \delta \sim \sqrt{1 - \delta^2} \sim 1 - \frac{\delta^2}{2}. \qquad 6.5\text{-}2$$

A 6.5-2 közelítések felhasználásával a gömbháromszög oldalaira vonatkozó cosinustétel az

$$1 - \frac{1}{2}\frac{\widehat{c}^2}{R^2} = \left(1 - \frac{1}{2}\frac{\widehat{a}^2}{R^2}\right)\left(1 - \frac{1}{2}\frac{\widehat{b}^2}{R^2}\right) + \frac{\widehat{a}}{R}\frac{\widehat{b}}{R}\cos \gamma \qquad 6.5\text{-}3$$

alakot ölti. Elvégezve a kijelölt műveleteket, és a negyedrendben kicsiny tagokat elhanyagolva, a

$$\widehat{c}^2 = \widehat{a}^2 + \widehat{b}^2 - 2\widehat{a}\widehat{b} \cos \gamma \qquad 6.5\text{-}4$$

formulához jutunk. A 6.5-4 összefüggés a síkháromszögekre vonatkozó cosinustételt adja. A gömb sugarához képest elhanyagolhatóan kicsiny oldalakkal rendelkező gömbháromszögekre tehát jó közelítéssel alkalmazhatjuk a síkháromszögekre vonatkozó cosinustételt.

$$\cos \gamma = \cos \alpha \sin \beta - \sin \alpha \sin \beta \cos c \qquad 6.5\text{-}5$$

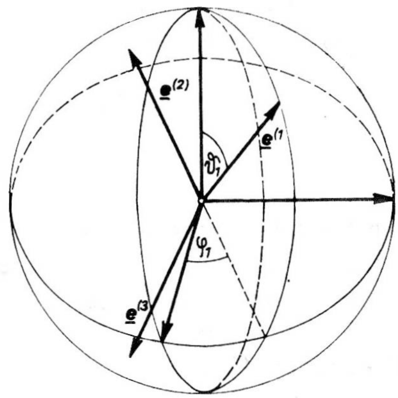
tételt felírva, és a $\cos c \sim 1$ közelítést alkalmazva, adódik, hogy

$$\cos \gamma \sim \cos (\alpha + \beta).$$

Következésképpen $\gamma \sim 180° - (\alpha + \beta)$, azaz ebben a határesetben a gömbháromszög szögeinek összege jó közelítéssel 180°.

6.6. Alkalmazás. A térbeli polárkoordináták egy tulajdonsága

A *6.3. ábráról* leolvasható, hogy egy tetszőleges vektort jellemezhetünk egy szakasz és két szög megadásával is. Az r, ϑ, φ értékeket az $\underline{\mathbf{r}}$ vektor polárkoordinátáinak nevezzük.



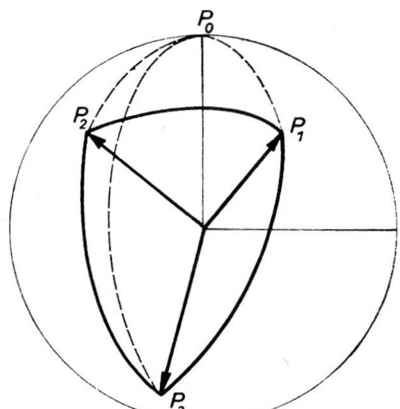
6.3. ábra

Legyenek egy $\mathcal{K}$ koordináta-rendszer alapvektorai az $\underline{\mathbf{e}}^{(1)}$, $\underline{\mathbf{e}}^{(2)}$, $\underline{\mathbf{e}}^{(3)}$ egységvektorok. Jelöljük ezen vektorok polárkoordinátáit $\vartheta^{(k)}$, $\varphi^{(k)}$, r-rel, $k = 1, 2, 3$, r értéke mindhárom vektor esetén egységhosszúságú.

A *6.3. ábrán* bejelölt példa alapján látható, hogy az $\underline{\mathbf{e}}^{(k)}$ vektorok $\mathcal{K}_0$ koordináta-rendszerben vett reprezentációi

$$\mathcal{K}_0(\underline{\mathbf{e}}^{(k)}) = \sin \vartheta_k \cos \varphi_k, \sin \vartheta_k \sin \varphi_k, \cos \varphi_k$$

$$(k = 1, 2, 3). \qquad 6.6\text{-}1$$



6.4. ábra

Az egységsugarú körből az $\underline{\mathbf{e}}^{(k)}$ vektorok egy $P_1P_2P_3$ gömbháromszöget metszenek ki *(6.4. ábra)*. Vegyük hozzá a fenti pontokhoz a $\vartheta_0 = 0$ koordinátával jellemzett P_0 pontot. Ekkor különböző $P_0P_kP_l$ gömbháromszögek keletkeznek. Ezeknek oldalai

$$\widehat{P_0P_k} = \vartheta_k, \quad \widehat{P_0P_l} = \vartheta_l \quad \text{és} \quad \widehat{P_kP_l} = \frac{\pi}{2} \quad (k \neq l)$$

és

$$(P_k P_0 P_l)\sphericalangle = \varphi_l - \varphi_k.$$

Alkalmazva erre a háromszögre a cosinustételt, azt kapjuk, hogy

$$0 = \cos\vartheta_k \cos\vartheta_l + \sin\vartheta_k \sin\vartheta_l \cos(\varphi_l - \varphi_k),$$

tehát

$$\operatorname{ctg}\vartheta_k \operatorname{ctg}\vartheta_l = -\cos(\varphi_l - \varphi_k).$$

A $k \neq l$ értékpárokra három egyenletet kapunk. Ezekből következik, hogy

$$\left.\begin{aligned}
\operatorname{tg}\vartheta_1 &= \sqrt{-\frac{\cos(\varphi_2-\varphi_3)}{\cos(\varphi_1-\varphi_2)\cos(\varphi_1-\varphi_3)}}\\
\operatorname{tg}\vartheta_2 &= \sqrt{-\frac{\cos(\varphi_3-\varphi_1)}{\cos(\varphi_2-\varphi_1)\cos(\varphi_2-\varphi_3)}}\\
\operatorname{tg}\vartheta_3 &= \sqrt{-\frac{\cos(\varphi_2-\varphi_1)}{\cos(\varphi_3-\varphi_1)\cos(\varphi_3-\varphi_2)}}
\end{aligned}\right\} \qquad 6.6\text{-}2$$

Adott φ_1, φ_2 és φ_3 szögekhez tehát a ϑ_1, ϑ_2, ϑ_3 értékek meghatározhatók. Így 6.6-1 és 6.6-2 segítségével három, páronként egymásra merőleges egységvektort három paraméterrel állíthatunk elő.

Megjegyezzük még, hogy a három paraméter annyiban nem független, hogy 6.6-2-nek csak akkor van értelme, ha a gyökjel alatt pozitív szám áll. A gyök mindkét előjele esetén egységvektort kapunk.

II. OPERÁTOROK

7. Lineáris operátorok

A vizsgálat tárgya sok esetben egy fizikai rendszer változása. Egy változást szimbolikusan operátorral fejezhetünk ki. Ha egy $\underline{Q}$ objektumot $\underline{Q}^*$-ra változtatunk, tehát a

$$\underline{Q} \to \underline{Q}^*$$

átmenetet vizsgáljuk, azt is írhatjuk, hogy

$$\underline{Q}^* = \underline{A}\underline{Q},$$

ahol $\underline{A}$ a változást jellemző operátor.

7.1. Forgatási operátorok

Egy példa a változásra az elforgatás. Legyen például $\underline{Q}$ egy P_0, P_1, ... pontokból álló merev test. $\underline{Q}$-t elforgatva, jelöljük az elforgatott testet $\underline{Q}^*$-gal, pontjait pedig P_0^*, P_1^*, ..., P_n^*-gal. A $\underline{Q} \to \underline{Q}^*$ változást egy „forgatási operátor" jellemzi. Az elforgatás tényét az

$$\underline{O}\underline{Q} = \underline{Q}^* \qquad 7.1\text{-}1$$

formulával fejezhetjük ki.

A 7.1-1 formulában $\underline{O}$ az elforgatást létesítő operátort jelöli. Az $\underline{O}$ operátor tehát a test méreteitől és helyzetétől függetlenül az elforgatást jellemzi. Az $\underline{O}$ operátort tetszőleges merev testre is alkalmazhatjuk. A jelölésben alkalmazott aláhúzást annak kiemelésére használjuk, hogy $\underline{O}$ a reprezentációtól és a testtől függetlenül egy objektíve létező műveletet jelent.

Amennyiben $\underline{R}$ egy $\underline{Q}$-tól különböző, adott helyzetű merev test, akkor az $\underline{O}$ operáció ezt a testet az

$$\underline{R}^* = \underline{O}\underline{R} \qquad 7.1\text{-}2$$

helyzetbe viszi át.

A továbbiakban a forgatási operátorok tulajdonságait vizsgáljuk. Álljon a $\underline{Q}$ test a $P_1, P_2 \dots P_n$ pontokból, amelyek helyvektorai rendre

$$\underline{r}_{P_1} = \underline{r}_1, \quad \underline{r}_{P_2} = \underline{r}_2, \dots, \underline{r}_{P_n} = \underline{r}_n. \qquad 7.1\text{-}3$$

Az elforgatott $\underline{Q}^*$ rendszer pontjai pedig legyenek

$$P_1^*, \dots, P_n^*,$$

helyvektoraik pedig

$$\underline{r}_{P1}^* = \underline{r}_1^*, \dots, \underline{r}_{Pn}^* = \underline{r}_n^*. \qquad 7.1\text{-}4$$

Amennyiben a forgatás az $\underline{r}_0$ helyvektorú pont körül történik, akkor

$$\underline{r}_0 = \underline{r}_0^*. \qquad 7.1\text{-}5$$

A további pontok helyvektoraira pedig fennáll az

$$\underline{r}_k^* = \underline{O}\underline{r}_k \qquad 7.1\text{-}6$$

összefüggés.

7.2. Az ortogonális transzformáció

Elemezzük először a forgatás néhány tulajdonságát. Az elforgatást jellemezhetjük anélkül is, hogy az $\underline{r}_0$ forgáspontot megadjuk. Legyen ugyanis

$$\underline{r}_{kl} = \underline{r}_l - \underline{r}_k$$

a P_k-t P_l-lel összekötő vektor.

Az elforgatás ezeket a vektorokat az

$$\underline{r}_{kl}^* = \underline{O}\underline{r}_{kl}$$

formulának megfelelően megváltoztatja. $\underline{r}_{kl}^*$ az elforgatott test P_k^* és P_l^* pontját köti össze.

Az elforgatás tulajdonsága, hogy

$$a = a^*,$$

vagyis

$$|\underline{\mathbf{O}}\underline{\mathbf{a}}| = |\underline{\mathbf{a}}|, \qquad 7.2\text{-}1$$

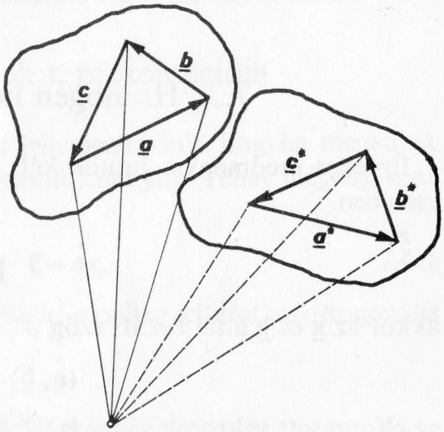
azaz a forgatás a merev test pontjai közötti távolságokat nem változtatja meg.[1] A 7.2-1 tulajdonsággal rendelkező operátorokat ortogonális operátoroknak nevezzük.

A forgatási operáció további tulajdonságait a 7.2-1 képletből állapíthatjuk meg. Legyen ABC egy merev testen kijelölt háromszög, melynek oldalai az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok. Megfelelő irányítás esetén írhatjuk, hogy

$$\underline{\mathbf{a}} + \underline{\mathbf{b}} = \underline{\mathbf{c}}. \qquad 7.2\text{-}2$$

A testtel együtt elforgatott háromszög oldalai legyenek az $\underline{\mathbf{a}}^*$, $\underline{\mathbf{b}}^*$, $\underline{\mathbf{c}}^*$ vektorok (*7.1. ábra*). Az elforgatott idom oldalaira definíció szerint érvényes, hogy

$$\underline{\mathbf{a}}^* + \underline{\mathbf{b}}^* = \underline{\mathbf{c}}^*. \qquad 7.2\text{-}3$$



7.1. ábra

Minthogy

$$\underline{\mathbf{a}}^* = \underline{\mathbf{O}}\underline{\mathbf{a}}, \quad \underline{\mathbf{b}}^* = \underline{\mathbf{O}}\underline{\mathbf{b}}, \quad \underline{\mathbf{c}}^* = \underline{\mathbf{O}}\underline{\mathbf{c}},$$

7.2-3 helyett írhatjuk, hogy

$$\underline{\mathbf{O}}(\underline{\mathbf{a}} + \underline{\mathbf{b}}) = \underline{\mathbf{O}}\underline{\mathbf{a}} + \underline{\mathbf{O}}\underline{\mathbf{b}}. \qquad 7.2\text{-}4$$

Mivel az ABC és az elforgatott $A^*B^*C^*$ háromszögek oldalai egyenlőek:

$$a = a^*, \quad b = b^*, \quad c = c^*, \qquad 7.2\text{-}5$$

a háromszögek egybevágóak, és így szögeik is egyenlőek. Ebből következik, hogy

$$\underline{\mathbf{a}}^*\underline{\mathbf{b}}^* = ab, \qquad 7.2\text{-}6$$

tehát két vektor skaláris szorzata nem változik az elforgatás következtében.

A skaláris szorzat értéke a vektorok tükrözése esetén sem változik. A 7.2-6 összefüggés tehát általánosan jellemzi az ortogonális operátorokat. Könnyen találhatunk olyan jellemzőt is, amelynek segítségével megkülönböztethetjük a tükrözéseket és forgatásokat.

Legyen $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ három, nem egy síkba eső vektor, és legyen

$$\underline{\mathbf{a}}^* = \underline{\mathbf{O}}\underline{\mathbf{a}}, \quad \underline{\mathbf{b}}^* = \underline{\mathbf{O}}\underline{\mathbf{b}}, \quad \underline{\mathbf{c}}^* = \underline{\mathbf{O}}\underline{\mathbf{c}},$$

[1] Egy pontrendszer pontjai közötti távolságok akkor sem változnak, ha a pontokat tetszőleges síkra *tükrözzük*. Így 7.2-1 a forgatási operátorokra vonatkozó szükséges feltételt jelent. Elégségessé azonban csak akkor válik, ha a forgatásokon kívül tükrözéseket is megengedünk.

ahol $\underline{\mathbf{O}}$ egy ortogonális transzformációt jelöl. Amennyiben $\underline{\mathbf{O}}$ forgatást jelent, akkor

$$(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}}) = (\underline{\mathbf{a}}^*, \underline{\mathbf{b}}^*, \underline{\mathbf{c}}^*). \qquad 7.2\text{-}7$$

A tükrözés a három vektor által definiált sodrást megfordítja, ezért ha $\underline{\mathbf{O}}$ tükrözés, akkor

$$(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}}) = -(\underline{\mathbf{a}}^*, \underline{\mathbf{b}}^*, \underline{\mathbf{c}}^*). \qquad 7.2\text{-}8$$

A hármas vegyes szorzat a tükrözés hatására előjelet vált. Ennek segítségével tehát különválaszthatjuk az ortogonális transzformációk közül a tükrözéseket.

7.3. Homogén lineáris transzformációk

Érdekes eredményre jutunk két, egy egyenesbe eső vektor transzformálása esetében.

Ha

$$\underline{\mathbf{b}} = \alpha \underline{\mathbf{a}},$$

akkor az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ által bezárt szög 0°, ha $\alpha > 0$, 180°, ha $\alpha < 0$; minthogy

$$(\underline{\mathbf{a}}, \underline{\mathbf{b}})\sphericalangle = (\underline{\mathbf{a}}^*, \underline{\mathbf{b}}^*)\sphericalangle,$$

az elforgatott vektorok szöge is 0°, illetve 180°, és így 7.1-1 segítségével

$$\underline{\mathbf{a}}^* = \alpha \underline{\mathbf{b}}^*, \quad \text{ha} \quad \underline{\mathbf{a}} = \alpha \underline{\mathbf{b}}.$$

Operátorokkal kifejezve:

$$\underline{\mathbf{O}}(\alpha \underline{\mathbf{a}}) = \alpha \underline{\mathbf{O}} \underline{\mathbf{a}}. \qquad 7.3\text{-}1$$

7.2-2-t és 7.3-1-et összefogva, azt kapjuk, hogy ha

$$\alpha \underline{\mathbf{a}} + \beta \underline{\mathbf{b}} = \underline{\mathbf{c}},$$

akkor

$$\alpha \underline{\mathbf{a}}^* + \beta \underline{\mathbf{b}}^* = \underline{\mathbf{c}}^*.$$

Operátorral kifejezve:

$$\underline{\mathbf{O}}(\alpha \underline{\mathbf{a}} + \beta \underline{\mathbf{b}}) = \alpha \underline{\mathbf{O}} \underline{\mathbf{a}} + \beta \underline{\mathbf{O}} \underline{\mathbf{b}}. \qquad 7.3\text{-}2$$

Több vektorra általánosítva a meggondolást, azt kapjuk, hogy

$$\underline{\mathbf{O}} \sum_{\nu} \alpha^{(\nu)} \underline{\mathbf{a}}^{(\nu)} = \sum_{\nu} \alpha^{(\nu)} (\underline{\mathbf{O}} \underline{\mathbf{a}}^{(\nu)}), \qquad 7.3\text{-}3$$

ahol az $\alpha^{(\nu)}$-k tetszőleges számok, az $\underline{\mathbf{a}}^{(\nu)}$-k pedig tetszőleges vektorok.

A 7.3-3 összefüggés az $\underline{\mathbf{O}}$ operátorra jellemző. Azokat az $\underline{\mathbf{A}}$ operátorokat, amelyek a 7.3-3 tulajdonsággal rendelkeznek, tehát amelyekre fennáll az

$$\underline{\mathbf{A}}\left(\sum_{\nu}\alpha^{(\nu)}\underline{\mathbf{a}}^{(\nu)}\right)=\sum_{\nu}\alpha^{(\nu)}\left(\underline{\mathbf{A}}\underline{\mathbf{a}}^{(\nu)}\right) \qquad 7.3\text{-}4$$

összefüggés, *lineáris* és *homogén (vektor-) operátornak* nevezzük. (Rövidebben lineáris operátorról is beszélhetünk.) Ezek szerint az ortogonális operátorok *lineáris* (pontosabban: lineáris homogén vektor-) operátorok. Az ortogonális operátorok különleges jellemzője a 7.2-5 és 7.2-6 által leírt sajátság, tehát az *ortogonális operátorok olyan lineáris operátorok, amelyek két vektor skaláris szorzatát nem változtatják.*

7.4. A lineáris operátorok reprezentációja

Általában egy $\underline{\mathbf{A}}$ lineáris operátort azzal jellemezhetünk, hogyha megadjuk három nem komplanáris $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektor transzformáltjait. Tehát megadjuk az $\bar{\underline{\mathbf{a}}}$, $\bar{\underline{\mathbf{b}}}$, $\bar{\underline{\mathbf{c}}}$ vektorokat, amelyekre

$$\bar{\underline{\mathbf{a}}}=\underline{\mathbf{A}}\underline{\mathbf{a}}, \quad \bar{\underline{\mathbf{b}}}=\underline{\mathbf{A}}\underline{\mathbf{b}}, \quad \bar{\underline{\mathbf{c}}}=\underline{\mathbf{A}}\underline{\mathbf{c}}. \qquad 7.4\text{-}1$$

A transzformáltakat most felülhúzással jelöljük, a csillag jelölést az ortogonális transzformációk esetére tartjuk fenn.

Ha

$$(\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}})\neq 0,$$

akkor tetszőleges vektort előállíthatunk az

$$\underline{\mathbf{x}}=\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}+\gamma\underline{\mathbf{c}} \qquad 7.4\text{-}2$$

alakban.

Az α, β, γ értékeket 4.6-8 szerint határozhatjuk meg. Alkalmazzuk $\underline{\mathbf{A}}$-t 7.4-2 két oldalára. 7.3-4 segítségével azt kapjuk, hogy

$$\bar{\underline{\mathbf{x}}}=\underline{\mathbf{A}}\underline{\mathbf{x}}=\alpha\bar{\underline{\mathbf{a}}}+\beta\bar{\underline{\mathbf{b}}}+\gamma\bar{\underline{\mathbf{c}}}, \qquad 7.4\text{-}3$$

tehát minden $\underline{\mathbf{x}}$ vektor transzformáltját elő tudjuk állítani, ha az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ vektorok és $\bar{\underline{\mathbf{a}}}$, $\bar{\underline{\mathbf{b}}}$, $\bar{\underline{\mathbf{c}}}$ transzformáltjaik ismeretesek.

A transzformáció előállítására egyszerű kifejezéseket kapunk, ha egy $\mathcal{K}$ koordináta-rendszer alapvektoraiból indulunk ki. Legyen tehát

$$\underline{\mathbf{a}}=\underline{\mathbf{e}}^{(1)}, \quad \underline{\mathbf{b}}=\underline{\mathbf{e}}^{(2)}, \quad \underline{\mathbf{c}}=\underline{\mathbf{e}}^{(3)}.$$

Megadva az

$$\underline{\mathbf{A}}\underline{\mathbf{e}}^{(k)}=\underline{\mathbf{f}}^{(k)} \qquad 7.4\text{-}4$$

vektorokat, azt találjuk, hogy

$$\underline{\mathbf{x}}=\sum x_k\underline{\mathbf{e}}^{(k)}, \qquad 7.4\text{-}5$$

ahol

$$x_1, x_2, x_3=\mathbf{x}=\mathcal{K}(\underline{\mathbf{x}}), \qquad 7.4\text{-}6$$

így 7.4-3 helyett írhatjuk, hogy

$$\bar{\mathbf{x}} = \sum x_l \mathbf{f}^{(l)}. \qquad 7.4\text{-}7$$

Szorozzuk 7.4-7-et $\mathbf{e}^{(k)}$-val. Azt kapjuk, hogy

$$\bar{x}_k = \bar{\mathbf{x}}\mathbf{e}^{(k)} = \sum_l x_l \mathbf{f}^{(l)}\mathbf{e}^{(k)}. \qquad 7.4\text{-}8$$

Legyen

$$\mathbf{e}^{(k)}\mathbf{f}^{(l)} = A_{kl} \quad (k, l = 1, 2, 3), \qquad 7.4\text{-}9$$

akkor

$$\bar{x}_k = \sum_l A_{kl} x_l. \qquad 7.4\text{-}10$$

Az $\bar{\mathbf{x}}$ vektor $\mathcal{K}$-beli reprezentációja

$$\mathcal{K}(\bar{\mathbf{x}}) = \bar{\mathbf{x}} = \bar{x}_1, \bar{x}_2, \bar{x}_3,$$

tehát előállítható az $\mathbf{x}$ vektor $\mathcal{K}$-beli komponenseinek lineáris kombinációjaként.

7.4-10 tehát megadja azt az eljárást, amelynek segítségével $\mathbf{x}$ adott reprezentációjából $\bar{\mathbf{x}}$ reprezentációját kiszámíthatjuk. Az A_{kl} együtthatók számértékei 7.4-9 segítségével állapíthatók meg. Az $\mathbf{x} \to \bar{\mathbf{x}}$ transzformáció ezután bármely $\mathcal{K}$ rendszerben 7.4-10 szerint hajtható végre.

A fenti meggondolást másképpen is megfogalmazhatjuk. Legyen az $\mathbf{A}$ operátor $\mathcal{K}$-beli reprezentációja

$$\mathcal{K}(\mathbf{A}) = A_{kl} \quad (k, l = 1, 2, 3).$$

Ez kilenc számérték, éspedig úgy, hogy ha

$$\bar{\mathbf{a}} = \mathbf{A}\mathbf{a},$$

akkor a $\mathcal{K}$ reprezentációban

$$\bar{a}_k = \sum_{l=1}^{3} A_{kl} a_l.$$

Az A_{kl} értékeket szimmetrikusan rendezve mátrixnak is nevezzük, és az

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

sémával jelöljük.

Az A_{kl} értékeket az A mátrix elemeinek nevezzük. A fenti sémának megfelelően

$$A_{11}, A_{12}, A_{13}$$

az $\mathbf{A}$ mátrix első,

$$A_{k1}, A_{k2}, A_{k3} \quad (k=1, 2, 3)$$

pedig a k-adik sora. Hasonló módon

$$A_{1l}, A_{2l}, A_{3l} \quad (l=1, 2, 3)$$

a mátrix l-edik oszlopa. A mátrixjelölést felhasználva,

$$\mathcal{K}(\underline{\mathbf{A}}) = \mathbf{A},$$

ahol az $\mathbf{A}$ mátrix elemei 7.4-9 szerint

$$A_{kl} = \underline{\mathbf{e}}^{(k)}\underline{\mathbf{f}}^{(l)} \quad (k, l=1, 2, 3).$$

7.5. Az ortogonális transzformációk reprezentációja, ortogonalitási relációk

Az

$$\underline{\mathbf{a}}^* = \underline{\underline{\mathbf{O}}}\underline{\mathbf{a}} \tag{7.5-1}$$

ortogonális transzformációt az általános lineáris transzformációhoz hasonlóan jellemezhetjük. Tehát megadjuk, hogy egy $\mathcal{K}$ rendszer $\underline{\mathbf{e}}^{(l)}$ alapvektorai hogyan változnak az $\underline{\underline{\mathbf{O}}}$ transzformáció folytán, azaz megadjuk az

$$\underline{\mathbf{e}}^{(k)*} = \underline{\underline{\mathbf{O}}}\underline{\mathbf{e}}^{(k)} \tag{7.5-2}$$

vektorokat. Minthogy a forgatás olyan lineáris operáció, amely a vektorok hosszát és szögét nem változtatja, 7.2-6-ból következik, hogy

$$\underline{\mathbf{e}}^{(k)*}\underline{\mathbf{e}}^{(l)*} = \delta_{kl}. \tag{7.5-3}$$

Keressünk még valamilyen egyszerű kritériumot a tükrözések és elforgatások megkülönböztetésére is. Alkossanak az eredeti $\underline{\mathbf{e}}^{(k)}$ alapvektorok jobbrendszert. Az elforgatások ezen a tulajdonságon nem változtatnak, tehát

$$\underline{\mathbf{e}}^{(k)*} \times \underline{\mathbf{e}}^{(l)*} = \underline{\mathbf{e}}^{(m)*}, \tag{7.5-4}$$

ha

$$\varepsilon_{klm} = +1.$$

A tükrözés az alapvektorok sodrását megváltoztatja, így ez esetben

$$\underline{\mathbf{e}}^{(k)*} \times \underline{\mathbf{e}}^{(l)*} = -\underline{\mathbf{e}}^{(m)*}, \tag{7.5-5}$$

ha

$$\varepsilon_{klm} = +1.$$

Legyen a $\mathcal{K}$ reprezentációban

$$\mathcal{K}(\underline{\mathbf{O}}) = \mathbf{O}, \quad \mathcal{K}(\underline{\mathbf{a}}) = \mathbf{a}, \quad \mathcal{K}(\underline{\mathbf{a}}^*) = \mathbf{a}^* \qquad 7.5\text{-}6$$

és

$$a_k^* = \sum_l O_{kl} a_l, \qquad 7.5\text{-}7$$

ahol

$$\mathbf{O} = \begin{pmatrix} O_{11} & O_{12} & O_{13} \\ O_{21} & O_{22} & O_{23} \\ O_{31} & O_{32} & O_{33} \end{pmatrix}.$$

Az O_{kl} értékek nem függetlenek, hiszen az $\underline{\mathbf{O}}$ operátor két vektor skaláris szorzatán nem változtat.

Legyen

$$\underline{\mathbf{b}}^* = \underline{\mathbf{O}}\underline{\mathbf{b}} \quad \text{és} \quad \underline{\mathbf{a}}^* = \underline{\mathbf{O}}\underline{\mathbf{a}}, \qquad 7.5\text{-}8$$

$$b_k^* = \sum_l O_{kl} b_l, \quad a_k^* = \sum_l O_{kl} a_l.$$

Innen 7.5-7 és 7.5-8 segítségével azt kapjuk, hogy

$$\underline{\mathbf{a}}^* \underline{\mathbf{b}}^* = \sum_{k=1}^{3} a_k^* b_k^* = \sum_{m,n} a_m b_n O_{km} O_{kn},$$

minthogy azonban

$$\underline{\mathbf{a}}^* \underline{\mathbf{b}}^* = \underline{\mathbf{a}}\underline{\mathbf{b}},$$

ezért

$$\sum_{m,n} a_m b_n O_{km} O_{kn} = \sum_K a_K b_K. \qquad 7.5\text{-}9$$

A fenti összefüggés akkor és csak akkor áll fenn minden $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorra, ha

$$\sum_k O_{km} O_{kn} = \delta_{mn} \quad (m, n = 1, 2, 3). \qquad 7.5\text{-}10$$

Valóban, minthogy 7.5-9 minden vektorra érvényes kell hogy legyen, vehetjük pl., hogy

$$a_m = \delta_{km}, \quad a_n = \delta_{kn} \quad (m, n = 1, 2, 3).$$

Ezeket az értékeket 7.5-9-be helyettesítve 7.5-10-et kapjuk. A 7.5-10 formulák az úgynevezett ortogonalitási relációk. E relációknak kell teljesülniük egy lineáris transzformáció együtthatóira, hogy a transzformáció ortogonális legyen.

a) Foglalkozzunk először a síkbeli transzformációk leírásával. A sík vektorait két komponenssel írhatjuk le, és így az

$$\underline{\mathbf{a}}^* = \underline{\underline{\mathbf{O}}}\,\underline{\mathbf{a}}$$

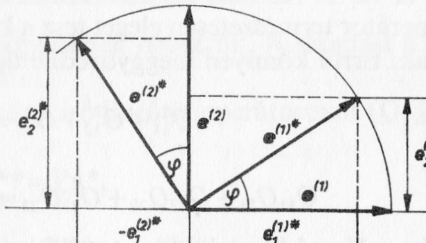
síkbeli forgatást az

$$a_k^* = \sum_{l=1}^{2} O_{kl} a_l \qquad 7.5\text{-}11$$

formulával reprezentálhatjuk, ahol

$$O_{kl} = \underline{\mathbf{e}}^{(k)} \underline{\mathbf{e}}^{(l)*} \quad (k, l = 1, 2).$$

A *7.2. ábrából* láthatjuk, hogy ha a síkbeli $\mathcal{K}$ rendszer egységvektorait φ szöggel elforgatjuk, akkor az



7.2. ábra

$$\underline{\mathbf{e}}^{(1)*} = \cos\varphi,\ \sin\varphi,$$

$$\underline{\mathbf{e}}^{(2)*} = -\sin\varphi,\ \cos\varphi \qquad 7.5\text{-}12$$

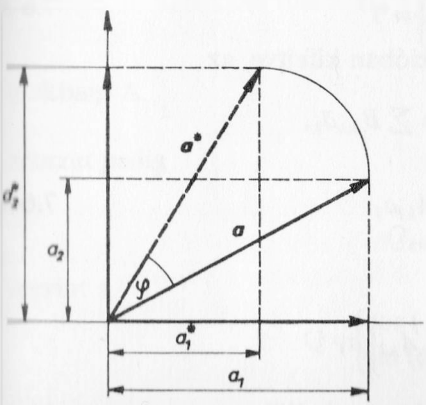
síkbeli vektorokat kapjuk. Eszerint az $\underline{\underline{\mathbf{O}}}$ operátor reprezentációja

$$\mathbf{O} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \qquad 7.5\text{-}13$$

Ez az elforgatás a $\mathcal{K}$ rendszerben így írható fel *(7.3. ábra):*

$$a_1^* = a_1 \cos\varphi - a_2 \sin\varphi,$$

$$a_2^* = a_1 \sin\varphi + a_2 \cos\varphi. \qquad 7.5\text{-}14$$



7.3. ábra

b) A háromdimenziós esetben az $\underline{\mathbf{e}}^{(k)*}$ vektorokat a $\mathcal{K}$ reprezentációban polárkoordinátákkal jellemezhetjük, tehát

$$\mathbf{e}^{(k)*} = \sin\vartheta_k \cos\varphi_k,\quad \sin\vartheta_k \sin\varphi_k,\quad \cos\vartheta_k, \qquad 7.5\text{-}15$$

ahol 6.6-2 szerint

$$\operatorname{tg}\vartheta_k = \sqrt{-\frac{\cos(\varphi_l - \varphi_m)}{\cos(\varphi_k - \varphi_l)\cos(\varphi_k - \varphi_m)}}. \qquad 7.5\text{-}16$$

Megjegyzés: A kétdimenziós forgatást a háromdimenziós forgatás speciális esetének is tekinthetjük, ha a 7.5-12 formulákat kiegészítjük az $\underline{\mathbf{e}}^{(3)}=\underline{\mathbf{e}}^{(3)\prime}$ összefüggéssel. Az $\underline{\mathbf{O}}$ operátor $\mathbf{O}=\mathcal{K}(\underline{\mathbf{O}})$ reprezentációja ez esetben

$$\mathbf{O}=\begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.5-17}$$

A többi formulák is érvényben maradnak az $a_3=a_3^*$ kiegészítéssel.

A 7.5-17 formula explicit formában adja a síkbeli forgatás operátorát. Ez az operátor természetesen eleget tesz a korábban megadott ortogonalitási relációknak. Erről könnyen meggyőződhetünk, hiszen például

$$O_{11}^2+O_{12}^2+O_{13}^2=\cos^2\varphi+\sin^2\varphi+0=1$$

és

$$O_{11}O_{21}+O_{12}O_{22}+O_{13}O_{23}=\cos\varphi\sin\varphi-\sin\varphi\cos\varphi+0=0.$$

Hasonló módon teljesül a további négy reláció is.

7.6. Lineáris transzformációk egymás utáni alkalmazása

Legyen $\underline{\mathbf{A}}$ és $\underline{\mathbf{B}}$ két lineáris operátor. Transzformáljunk először egy $\underline{\mathbf{a}}$ vektort $\underline{\mathbf{A}}$ szerint. Legyen

$$\bar{\underline{\mathbf{a}}}=\underline{\mathbf{A}}\underline{\mathbf{a}}, \tag{7.6-1}$$

majd transzformáljuk $\bar{\underline{\mathbf{a}}}$-t $\underline{\mathbf{B}}$ szerint, és legyen

$$\bar{\bar{\underline{\mathbf{a}}}}=\underline{\mathbf{B}}\bar{\underline{\mathbf{a}}}. \tag{7.6-2}$$

7.6-2 helyett azt is írhatjuk, hogy

$$\bar{\bar{\underline{\mathbf{a}}}}=\underline{\mathbf{B}}(\underline{\mathbf{A}}\underline{\mathbf{a}}). \tag{7.6-3}$$

A 7.6-2 és 7.6-3 formulát egy $\mathcal{K}$ reprezentációban kifejtve, az

$$\bar{a}_k=\sum_l A_{kl}a_l, \quad \bar{\bar{a}}_m=\sum_k B_{mk}\bar{a}_k,$$

ill.

$$\bar{\bar{a}}_m=\sum_{k,\,l} B_{mk}A_{kl}a_l \tag{7.6-4}$$

összefüggésekhez jutunk. 7.6-4 az

$$\bar{\bar{a}}_m=\sum_l\left(\sum_k B_{mk}A_{kl}\right)a_l$$

alakban is felírható. Ennek alapján viszont bevezethetjük a

$$\sum_k B_{mk}A_{kl}=C_{ml} \qquad 7.6\text{-}5$$

jelölést.

Ez annyit jelent, hogy az $\underline{\mathbf{A}}$ és $\underline{\mathbf{B}}$ lineáris transzformációk egymásutánja egyetlen lineáris transzformációval helyettesíthető. Így

$$\bar{\bar{\underline{\mathbf{a}}}}=\underline{\mathbf{C}}\underline{\mathbf{a}}, \qquad 7.6\text{-}6$$

ahol $\underline{\mathbf{C}}$ lineáris operátor. Ennek $\mathcal{K}$-beli reprezentációja

$$\mathcal{K}(\underline{\mathbf{C}})=\mathbf{C},$$

amelynek elemeit 7.6-5 definiálja.

A $\underline{\mathbf{C}}$ operátort az $\underline{\mathbf{A}}$ és $\underline{\mathbf{B}}$ operátorok szorzatának nevezzük. A szorzást a $\mathcal{K}$ reprezentációban 7.6-5 szerint végezhetjük el.

Válasszunk most speciálisan forgatásokat leíró lineáris operátorokat ($\underline{\mathbf{O}}$, $\underline{\mathbf{P}}$), és legyen

$$\underline{\mathbf{a}}^*=\underline{\mathbf{O}}\underline{\mathbf{a}}, \quad \underline{\mathbf{a}}^{**}=\underline{\mathbf{P}}\underline{\mathbf{a}}^*,$$

akkor

$$\underline{\mathbf{a}}^{**}=\underline{\mathbf{P}}(\underline{\mathbf{O}}\underline{\mathbf{a}})=\underline{\mathbf{Q}}\underline{\mathbf{a}},$$

ahol $\underline{\mathbf{Q}}$ reprezentációja

$$\mathcal{K}(\underline{\mathbf{Q}})=\mathbf{Q},$$

amelynek elemei

$$Q_{km}=\sum_{l=1}^{3}P_{kl}O_{lm} \quad (k, m=1, 2, 3).$$

Amennyiben $\underline{\mathbf{O}}$ és $\underline{\mathbf{P}}$ kétdimenziós forgatások, akkor $\mathbf{O}$ és $\mathbf{P}$ felírható az

$$\mathbf{O}=\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

$$\mathbf{P}=\begin{pmatrix}\cos\psi & -\sin\psi\\ \sin\psi & \cos\psi\end{pmatrix}$$

alakban. A

$$\mathbf{Q}=\mathbf{P}\mathbf{O}$$

szorzat pedig

$$Q_{km}=\sum_{l=1}^{2}P_{kl}O_{lm}$$

szerint a

$$\mathbf{Q}=\begin{pmatrix}\cos(\varphi+\psi) & -\sin(\varphi+\psi)\\ \sin(\varphi+\psi) & \cos(\varphi+\psi)\end{pmatrix}$$

alakot ölti.

A φ és ψ szöggel történő síkbeli elforgatások egymásutánja tehát, amint az várható volt, egyetlen $\varphi+\psi$ szöggel történő elforgatással helyettesíthető.

Ez az eredmény is mutatja, hogy eddigi meggondolásaink a forgatások helyes leírásához vezettek.

8. Mátrixalgebra

A forgatási operátorokat *mátrixokkal* reprezentáltuk. A következőkben a fogalmat kibővítjük, és a mátrixműveleteket szabályokba foglaljuk. A mátrixműveleteket úgy vezetjük be, hogy a forgatások leírása áttekinthetőbbé váljék, de a mátrixszámítás egyéb területeken is használható legyen.

8.1. A mátrix fogalma

Mátrixnak — mint már említettük — egy megfelelő módon rendezett számhalmazt nevezünk. A számhalmaz tagjai a mátrix elemei. A mátrix elemeit egy- vagy több indexes betűkifejezéssel jelöljük. Az indexek sorrendje adja a számhalmaz rendezésének módját.

Az

$$a_1 \ldots a_N \tag{8.1-1}$$

számokból álló szám- N-est egydimenziós mátrixnak nevezzük. Mivel a vektorok egydimenziós mátrixokkal reprezentálhatók, az egydimenziós mátrixot gyakran vektornak is nevezik. Ez az elnevezés azonban megtévesztő, mert — mint a későbbiek során látni fogjuk — nem minden szám- N-es tekinthető vektornak. Vektoron egy speciális tulajdonságokkal rendelkező valódi objektumot, illetve annak reprezentációját értjük, s tetszőleges szám-N-es nem mindig rendelkezik ezekkel a tulajdonságokkal. (Képezhetünk például Budapest férfi, nő és gyermek lakóinak számából egy számhármast, ez a számhármas azonban semmiképpen sem tekinthető egy vektor reprezentációjának.)

Ezért a mátrixszámításban használt szám-N-esek elnevezésére az egydimenziós mátrix kifejezést használjuk. A vektor elnevezésre csak akkor térünk át, amikor ez már nem adhat okot félreértésekre. A fenti esetben, mivel a számhalmaz N elemből áll, N komponensű mátrixról beszélünk.

A kétindexes mennyiségek a forgatások leírásakor már használt kétdimenziós sémában írhatók fel. Az A_{kl} elemeket tartalmazó $\mathbf{A}$ mátrix elemei tehát az

$$\mathbf{A} = \begin{bmatrix} A_{11} & \dots & A_{1N} \\ \vdots & & \\ \vdots & & \\ A_{M1} & & A_{MN} \end{bmatrix} \qquad 8.1\text{-}2$$

formába rendezhetők.

Az **A** mátrixot kétdimenziósnak nevezzük, mert kétdimenziós séma segítségével állítható elő. A 8.1-2 séma alapján az **A** mátrixot N oszlopú és M sorú vagy $N \times M$-es mátrixnak is nevezzük. Amennyiben $N = M$, akkor kvadratikus mátrixról beszélünk.

Általában legyenek az $A_{\underbrace{kl\dots s}_{n}}$ mennyiségek egy n indexszel jellemezhető számhalmaz elemei, ekkor azt mondjuk, hogy az $A_{kl\dots s}$ értékek egy n dimenziós mátrix elemei. Az egyes indexek különböző értékeket vehetnek fel. Általában tehát meg kell adni, hogy

$$\begin{aligned} k &= 1\dots N_1, \\ l &= 1\dots N_2, \\ &\vdots \\ s &= 1\dots N_n. \end{aligned}$$

A legegyszerűbb esetben $N_1 = N_2 = \dots = N_n = N$, vagyis minden index ugyanazon értékeken futhat végig. Ezeket a mátrixokat N-ed rendű mátrixoknak nevezzük.

Amennyiben egy mátrix dimenzióját is kifejezésre akarjuk juttatni, akkor az

$$\overset{(n)}{\mathbf{A}}$$

szimbólumot használjuk. Ennek alapján tehát

$\overset{(1)}{\mathbf{A}}$ egydimenziós mátrix,

$\overset{(2)}{\mathbf{A}}$ kétdimenziós mátrix, „szokásos mátrix”

$\overset{(3)}{\mathbf{A}}$ háromdimenziós mátrix.

A továbbiakban ezt a jelölésmódot csak akkor használjuk, ha formuláink egyébként félreérthetővé válnának.

Amennyiben az indexek értelmezési tartománya is lényeges, akkor az alábbi jelölésmódot használjuk:

$\underset{N}{\overset{(1)}{A}}$ N komponensű mátrix;

$\dfrac{\overset{(2)}{A}}{MN}$ N oszloppal és M sorral rendelkező mátrix;

$\dfrac{\overset{(3)}{A}}{N_1 N_2 N_3}$ háromdimenziós mátrix, melynek első indexe $1 \ldots N_1$,
második indexe az $1 \ldots N_2$,
harmadik indexe az $1 \ldots N_3$ értéket veszi fel.

Hasonló jelölést használhatunk a több dimenziós mátrixok esetén.

Az n dimenziós mátrixot tehát $\dfrac{\overset{(n)}{A}}{N_1 N_2 \ldots N_n}$ szimbólummal jelölhetjük. Az egyszerűség kedvéért azonban — félreérthető esetektől eltekintve — az $\overset{(n)}{\mathbf{A}}$ jelölést használjuk.

Végezetül néhány elfajult esettel foglalkozunk. Az $n=0$ esetén, azaz a nulldimenziós mátrixon index nélküli mennyiséget értünk. A nulldimenziós mátrix egy számértéket jelent. A nulldimenziós mátrixot gyakran — nem egészen pontosan — skalárnak is nevezik.

Az $\dfrac{\overset{(1)}{A}}{1}$ egykomponensű mátrixot egyetlen számértékkel jellemezhetjük.

Kétdimenziós mátrixok esetén az

$$\overset{(2)}{\mathbf{A}} = (A_{11}, A_{12}, \ldots, A_{1N})$$

mátrixot sorvektornak, a

$$\overset{(2)}{\mathbf{B}} = \begin{vmatrix} B_{11} \\ B_{21} \\ \vdots \\ B_{N1} \end{vmatrix}$$

mátrixot pedig oszlopvektornak is szokás nevezni.

A sor, illetve oszlop kifejezésnek két dimenzió esetén szemléletes tartalma van. Nem célszerű azonban ezeket a fogalmakat az $n \neq 2$ esetekre is kiterjeszteni. $n=1$ esetén pl. a sor- és oszlopvektorok közötti különbségtétel megduplázná a vektorfogalmat anélkül, hogy a fogalom tartalma bővülne.[1]

[1] Ha egy tetszőleges n dimenziós mátrix valamely $k \leqq n$ indexe csak az 1 értéket veheti fel, akkor a mátrix tulajdonképpen egy $n-k$ dimenziójú mátrixszá egyszerűsödik. Ily módon $\binom{n}{k}$ elfajulási lehetőség van, s ennek megfelelően $\binom{n}{k}$ féle $n-k$ dimenziós mátrixot kellene bevezetnünk.

8.2. Mátrixműveletek

Két mátrix akkor egyenlő, ha azonos dimenziójúak, indexeik értelmezési tartománya rendre megegyezik, és az azonos indexű elemek egyenlőek egymással.

8.2.1. Összeadás és kivonás

Legyen $\overset{(n)}{\mathbf{A}}$ egy $A_{kl\ldots s}$ elemű, $\overset{(n)}{\mathbf{B}}$ pedig egy $B_{kl\ldots s}$ elemű mátrix, amelyek azonos dimenziójúak, és indexeik értelmezési tartománya rendre megegyezik. Ez esetben a

$$\overset{(n)}{\mathbf{C}}=\overset{(n)}{\mathbf{A}}+\overset{(n)}{\mathbf{B}}$$

összegmátrix elemeit definíció szerint a

$$\overset{(n)}{C}_{kl\ldots s}=\overset{(n)}{A}_{kl\ldots s}+\overset{(n)}{B}_{kl\ldots s} \tag{8.2-2}$$

összefüggés határozza meg. A 8.2-2 definíciós egyenlet az $n=0$, 1 esetben a szokásos skalár-, illetve vektorösszeadási szabályt adja vissza. A kétdimenziós eset tehát

$$C_{kl}=A_{kl}+B_{kl}.$$

Összefoglalva: Két mátrix összegének elemei megegyeznek az egyes mátrixok megfelelő indexű elemeinek összegével. A fenti összeadási szabály alapján meghatározhatjuk a nullmátrix elemeit. Nullmátrixon definíció szerint azt a mátrixot értjük, amelyre tetszőleges **A** mátrix esetén fennáll az

$$\mathbf{A}+\mathbf{0}=\mathbf{A}$$

összefüggés. A 8.2-2 összegzési szabályból következik, hogy a nullmátrix minden eleme zérus.

8.2.2. Mátrix szorzása számmal

Legyen

$$\overset{(n)}{\mathbf{A}}=A_{kl\ldots s}.$$

A

$$\overset{(n)}{\mathbf{B}}=\lambda\mathbf{A}=\mathbf{A}\lambda$$

mátrixon azt a mátrixot értjük, amelynek minden eleme az A mátrix elemeinek λ-szorosa, tehát

$$B_{kl\ldots s} = \lambda A_{kl\ldots s} \qquad 8.2\text{-}3$$

Mátrix és skalár szorzatára definíció szerint érvényes a kommutativitás és az asszociativitás, valamint a két műveletre vonatkozóan a

$$\left.\begin{aligned}(\lambda_1+\lambda_2)\mathbf{A} &= \lambda_1\mathbf{A}+\lambda_2\mathbf{A}\\ \lambda(\mathbf{A}+\mathbf{B}) &= \lambda\mathbf{A}+\lambda\mathbf{B}\end{aligned}\right\} \qquad 8.2\text{-}4$$

egyenlőségekkel kifejezett disztributivitás is.

8.2.3. Kétdimenziós mátrixok szorzási szabályai

A kétdimenziós mátrixok szorzását a két egymás utáni forgatás leírásakor kapott kifejezés alapján definiáljuk. Láttuk, hogy az $\underline{O}$, $\underline{P}$ forgatások egymásutánja helyett egyetlen $\underline{Q}$ elforgatással is ugyanarra az eredményre juthatunk. Az elforgatások **O** és **P** mátrixreprezentációinak felhasználásával azt kapjuk, hogy a **Q** mátrix elemeit a

$$Q_{kl} = \sum_m P_{km} O_{ml} \qquad 8.2\text{-}5$$

formulával határozhatjuk meg.

A 8.2-5 kifejezést általánosítva, az **A** és **B** kétdimenziós mátrixok

$$\mathbf{C} = \mathbf{AB}$$

szorzatát a

$$C_{kl} = \sum_m A_{km} B_{ml} \qquad 8.2\text{-}6$$

egyenlettel definiáljuk.

Hangsúlyozzuk, hogy a szorzási szabály a két mátrix közelebbi indexeire történő összegezést ír elő.

$$\underset{\uparrow}{km}\;\underbrace{\qquad\qquad}_{\substack{\downarrow\\ \text{összegző index}}}\;\underset{\uparrow}{ml}\,.$$

Mivel az első index a mátrix sorait, a második index a mátrix oszlopait jelöli, a 8.2-6 definíciót röviden sor—oszlop szorzásnak szokás nevezni.

Az **AB** szorzási művelet csak abban az esetben végezhető el, ha **A** második indexe ugyanannyi értéket vesz fel, mint **B** első indexe.

A kétdimenziós mátrixok szorzása általában nem kommutatív művelet. Ez a definíció közvetlen következménye, mivel az egyik mátrix sorait a másik oszlopaival kell szoroznunk, hiszen a

$$C_{kl}=\sum_{m=1} A_{km}B_{ml}$$

kifejezésben a közeli indexre kell összegeznünk. A szorzási műveletet definiálhatnánk úgy is, hogy a kommutativitás fennálljon. A

$$C^1_{kl}=\sum_{m=1} A_{km}B_{lm}$$

oszlop—oszlop szorzási definíció már kommutatív műveletekre vezetne. Az általunk elfogadott szorzási szabály azonban alkalmasabb objektív összefüggések leírására. A bevezetésben tárgyalt forgatások esetén például két forgatás egymásutánja általában függ az egyes forgatások sorrendjétől. Ezért ez a művelet éppen a nem kommutatív szorzási szabály segítségével írható le helyesen.

A szorzási szabály közvetlen következménye, hogy a nullmátrix és tetszőleges mátrix szorzatának eredménye ismét a nullmátrix

$$\mathbf{0A}=\mathbf{0}.$$

A kvadratikus mátrixok körében célszerű bevezetni az

$$E_{kl}=\delta_{kl} \tag{8.2-7}$$

elemekkel rendelkező **E** egységmátrixot. Az így definiált mátrixra ugyanis fennáll, hogy tetszőleges **A** mátrixszal szorozva:

$$\mathbf{EA}=\mathbf{AE}=\mathbf{A}. \tag{8.2-8}$$

A mátrixszorzásra vonatkozóan érvényes az asszociatív, valamint az összeadásra és szorzásra vonatkozóan a disztributív törvény.

Az asszociatív törvény érvényességét a következő módon láthatjuk be. Legyen

$$\mathbf{P}=(\mathbf{AB})\mathbf{C}.$$

A szorzási szabály ismételt alkalmazásával a **P** mátrix elemeit a

$$P_{kl}=\sum_{o}(\mathbf{AB})_{ko}C_{ol}=\sum_{o}\left(\sum_{m} A_{km}B_{mo}\right)C_{ol}$$

formula határozza meg. A szummázások sorrendje azonban felcserélhető, így

$$P_{kl}=\sum_{m} A_{km}\left(\sum_{o} B_{mo}C_{ol}\right)=\sum_{m} A_{km}(\mathbf{BC})_{ml},$$

ami viszont az A(BC) sorrendű szorzás eredményével azonos. Tehát írhatjuk, hogy

$$\mathbf{P}=(\mathbf{AB})\mathbf{C}=\mathbf{A}(\mathbf{BC})=\mathbf{ABC}, \qquad 8.2\text{-}9$$

vagyis a szorzat felírásakor a zárójel használatától eltekinthetünk.

Hasonló módon látható be az

$$\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC} \qquad 8.2\text{-}10$$

disztributív törvény érvényessége is, ezt azonban az olvasóra bízzuk.

8.2.4. Egy- és kétdimenziós mátrix szorzata

A művelet a később tárgyalásra kerülő általános mátrix—mátrix szorzás speciális esete. Fontossága miatt azonban külön is tárgyaljuk.

Az egy- és kétdimenziós mátrixszorzatot is a forgatások vizsgálatakor talált összefüggések általánosításaként vezetjük be. Megállapítottuk, hogy az **a** vektor **a*** elforgatottjának komponenseit az

$$a_k^*=\sum_{l=1}^{3} O_{kl}a_l \qquad 8.2\text{-}11$$

formula segítségével határozhatjuk meg. A 8.2-11 összefüggés általánosításával az

$$\underset{(2)}{\mathbf{a}}=a_1\ldots a_N$$

N komponensű mátrix és az $\mathbf{A}$ kétdimenziós mátrix

$$\mathbf{b}=\mathbf{Aa}$$

szorzatának elemeit a

$$b_k=\sum_{l} A_{kl}a_l \qquad 8.2\text{-}12$$

összefüggéssel definiáljuk.

Célszerű a definíciót arra az esetre is kiterjeszteni, amikor az egydimenziós mátrix az első tényező.

A

$$\mathbf{b}'=\mathbf{aA}$$

szorzat elemeit a

$$b_k'=\sum_{l=1}^{3} a_lA_{lk} \qquad 8.2\text{-}13$$

kifejezéssel definiáljuk.

Általában $\mathbf{b}\neq\mathbf{b}'$, tehát

$$\mathbf{Aa}\neq\mathbf{aA}, \qquad 8.2\text{-}14$$

Így ez a művelet sem kommutatív.

Fennáll azonban az

$$\mathbf{a}(\mathbf{A}\mathbf{b})=(\mathbf{a}\mathbf{A})\mathbf{b} \qquad 8.2\text{-}15$$

értelemben vett asszociativitás. A bizonyítás céljából írjuk ki a 8.2-15 bal oldalán szereplő szorzatot komponensekben:

$$\sum_{k=1}^{3} a_k \left(\sum_{m=1}^{3} A_{km} b_m\right) = \sum_{k=1}^{3} \sum_{m=1}^{3} a_k A_{km} b_m. \qquad 8.2\text{-}16$$

8.2-16-ban a szummázás sorrendjét felcserélve, a

$$\sum_{m=1}^{3} b_m \left(\sum_{k=1}^{3} a_k A_{km}\right)$$

szorzathoz jutunk, ami éppen 8.2-15 jobb oldalával egyenértékű. A 8.2-15 asszociatív törvény miatt a szorzatot zárójel nélkül is felírhatjuk, tehát

$$\mathbf{a}(\mathbf{A}\mathbf{b})=(\mathbf{a}\mathbf{A})\mathbf{b}=\mathbf{a}\mathbf{A}\mathbf{b}. \qquad 8.2\text{-}17$$

Az egy- és kétdimenziós mátrixok szorzására fennáll az

$$\left.\begin{aligned}\mathbf{A}(\mathbf{a}+\mathbf{b})&=\mathbf{A}\mathbf{a}+\mathbf{A}\mathbf{b},\\ (\mathbf{a}+\mathbf{b})\mathbf{A}&=\mathbf{a}\mathbf{A}+\mathbf{b}\mathbf{A},\end{aligned}\right\} \qquad 8.2\text{-}18$$

ill.

valamint az

$$(\mathbf{A}+\mathbf{B})\mathbf{a}=\mathbf{A}\mathbf{a}+\mathbf{B}\mathbf{a}, \quad \text{ill.} \quad \mathbf{a}(\mathbf{A}+\mathbf{B})=\mathbf{a}\mathbf{A}+\mathbf{a}\mathbf{B} \qquad 8.2\text{-}19$$

értelemben vett *disztributivitás* is.

Ezeknek a törvényeknek az igazolása a 8.2-18 és 8.2-19 összefüggések komponensenként történő kiírásával egyszerűen végrehajtható.

Végül még egy, a mátrixok egyenlőségére vonatkozó tételt bizonyítunk.

Megjegyezzük, hogy két mátrix szorzata egyenlő lehet a nullamátrixszal, tehát

$$\mathbf{A}\mathbf{B}=\mathbf{0}$$

akkor is, ha $\mathbf{A}\neq\mathbf{0}$ és $\mathbf{B}\neq\mathbf{0}$.

Példaként említjük az

$$\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

szorzatot.

Ha azonban

$$\mathbf{A}\mathbf{X}=\mathbf{0},$$

minden olyan $\mathbf{X}$-re, amelyre a fenti szorzat értelmezhető, akkor ebből $\mathbf{A}=\mathbf{0}$

már következik. E tétel helyett egy hasonló tételt bizonyítunk, ti. ha

$$\mathbf{xAy} = 0 \tag{8.2-20}$$

minden x és y-ra, amelyekre a fenti szorzat értelmezhető, akkor

$$\mathbf{A} = \mathbf{0}.$$

Valóban, ha $\mathbf{A}$ egy $N \times M$-es mátrix, akkor $\mathbf{x}$-nek N, $\mathbf{y}$-nak M komponense kell hogy legyen. Minthogy 8.2-20 minden $\mathbf{x}$-re és $\mathbf{y}$-ra érvényes, $\mathbf{x}$-et megválaszthatjuk úgy, hogy az n-edik komponense 1, a többi 0, $\mathbf{y}$-t pedig úgy, hogy m-edik komponense 1, a többi 0. Legyen

$$x_k = \delta_{kn} \quad (k = 1, 2, \ldots, N),$$

$$y_m = \delta_{lm} \quad (l = 1, 2, \ldots, M),$$

tehát

$$\mathbf{xAy} = \sum_{k,\,l} \delta_{kn} A_{kl} \delta_{lm} = A_{nm} = 0.$$

Ha az $n = 1, 2, \ldots, N$ és $m = 1, 2, \ldots, M$ eseteket sorba vesszük, látjuk, hogy minden A_{nm}-nek el kell tűnnie, azaz $\mathbf{A} = \mathbf{0}$.

Hasonló módon belátjuk, hogy ha tetszőleges $\mathbf{x}$, $\mathbf{y}$-ra

$$\mathbf{xAy} = \mathbf{xBy}, \tag{8.2-21}$$

akkor

$$\mathbf{A} = \mathbf{B}.$$

Ugyanis 8.2-21-ből és a disztributív törvényből következik, hogy

$$\mathbf{x(A - B)y} = 0,$$

tehát

$$\mathbf{A} - \mathbf{B} = \mathbf{0}$$

és

$$\mathbf{A} = \mathbf{B}, \tag{8.2-22}$$

vagyis 8.2-21-ből 8.2-22 valóban következik.

8.3. A transzpozíció

A

$$\mathbf{b} = \mathbf{aA} \tag{8.3-1}$$

szorzási műveletet komponensenként kiírva a 8.3-1 formulához jutunk,

$$b_k = \sum a_l A_{lk}. \tag{8.3-2}$$

Tekintsük a $\mathbf{B}$ mátrixot, amelynek elemeit a

$$B_{kl}=A_{lk} \quad (k, l=1, 2, \ldots) \qquad 8.3\text{-}3$$

összefüggés definiálja.

A **B** mátrixot tehát úgy kapjuk, hogy **A** sorait és oszlopait felcseréljük.

8.3-3-at 8.3-2-be beírva, a

$$b_k=\sum_l B_{kl}a_l \qquad 8.3\text{-}4$$

összefüggést kapjuk. Az a_l és B_{kl} értékek itt minden további nélkül felcserélhetők, hiszen ezek már számértékeket jelölnek.

Tehát

$$\mathbf{b}=\mathbf{aA}=\mathbf{Ba}, \qquad 8.3\text{-}5$$

vagyis **a**-ból **b** úgy is előállítható, hogy **a**-t a **B** mátrixszal jobbról szorozzuk.

A **B** mátrixot az **A** mátrix *transzponáltjának* nevezzük, és a transzponált jelölésére a

$$\mathbf{B}=\widetilde{\mathbf{A}} \qquad 8.3\text{-}6$$

jelölést használjuk. A transzponált mátrixra általában érvényes, hogy

$$\mathbf{aA}=\widetilde{\mathbf{A}}\mathbf{a}. \qquad 8.3\text{-}7$$

A definícióból az is következik, hogy ha 8.3-6 fennáll, akkor

$$\widetilde{\mathbf{B}}=\mathbf{A}. \qquad 8.3\text{-}8$$

Az ismételt transzponálás tehát visszavezet az eredeti mátrixhoz, vagyis

$$\widetilde{\widetilde{\mathbf{A}}}=\mathbf{A}.$$

Ez nyilvánvaló, hiszen a transzpozíció a sorok és oszlopok felcserélését jelenti, s az ismételt csere az eredeti állapotot állítja vissza.

8.3.1. Néhány speciális mátrix

Szimmetrikus mátrixoknak nevezzük az olyan mátrixokat, amelyekre

$$A_{kl}=A_{lk} \quad (l, k=1, 2, 3).$$

Szimmetrikus mátrixok esetén tehát

$$\mathbf{A}=\widetilde{\mathbf{A}},$$

és így tetszőleges **a** esetén

$$\mathbf{aA}=\mathbf{Aa}.$$

A szimmetrikus mátrix speciális esete az ún. diagonálmátrix. *Diagonálmátrixnak* nevezzük az olyan mátrixot, amelynek elemei a főátlóbeli elemek kivételével eltűnnek. Tehát

$$D_{kl}=D_k\delta_{kl}, \tag{8.3-9}$$

illetve részletesen

$$\mathbf{D}=\begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & \\ \cdot & & & \\ \vdots & & & \\ 0 & \dots & \dots & D_N \end{pmatrix}.$$

A már megismert **E** egységmátrix a diagonálmátrix speciális esete, hiszen

$$E_{kl}=\delta_{kl}.$$

8.3.2. A transzpozíció szabályai

A mátrixokkal való számolás szempontjából igen fontos az összeg- és szorzatmátrixok transzponáltjának kifejezése az eredeti mátrixok elemeivel.

Legyenek A_{kl}, B_{kl}, C_{kl} rendre az **A**, **B**, **C** mátrixok elemei. A $\mathbf{C}=\mathbf{A}+\mathbf{B}$ mátrix $\tilde{\mathbf{C}}$ transzponáltja

$$\tilde{\mathbf{C}}=\widetilde{\mathbf{A}+\mathbf{B}}=\tilde{\mathbf{A}}+\tilde{\mathbf{B}},$$

hiszen komponensekben kifejezve:

$$C_{kl}=A_{kl}+B_{kl},$$

$$\tilde{C}_{kl}=C_{lk}=A_{lk}+B_{lk}=\tilde{A}_{kl}+\tilde{B}_{kl}. \tag{8.3-10}$$

Az **A** és **B** mátrixok $\mathbf{C}=\mathbf{AB}$ szorzatának transzponáltjára pedig fennáll, hogy

$$\tilde{\mathbf{C}}=(\widetilde{\mathbf{AB}})=\tilde{\mathbf{B}}\tilde{\mathbf{A}}. \tag{8.3-11}$$

8.3-11 bizonyítására írjuk fel **C**-t komponensekben

$$\left.\begin{aligned} C_{kl}&=\sum_m A_{km}B_{ml}, \\ \tilde{C}_{kl}&=C_{lk}=\sum_m A_{lm}B_{mk}. \end{aligned}\right\} \tag{8.3-12}$$

Ha a 8.3-12 szummában transzponáltakra térünk át,

$$\tilde{C}_{kl}=\sum_m \tilde{A}_{ml}\tilde{B}_{km}=\sum_m \tilde{B}_{km}\tilde{A}_{ml}. \tag{8.3-13}$$

8.3-13 egyenértékű 8.3-11-gyel. Szorzat transzponáltját tehát úgy kapjuk, hogy a tényezők transzponáltjait fordított sorrendben összeszorozzuk.

E szabályt tetszőleges számú mátrix szorzatára kiterjesztve, azt kapjuk, hogy ha

$$\mathbf{A}=\mathbf{PQ}\ldots\mathbf{S}, \qquad 8.3\text{-}14$$

akkor

$$\tilde{\mathbf{A}}=\tilde{\mathbf{S}}\ldots\tilde{\mathbf{Q}}\tilde{\mathbf{P}}, \qquad 8.3\text{-}15$$

vagyis egy szorzat transzponáltja megegyezik a tényezők transzponáltjainak fordított sorrendben vett szorzatával.

8.4. Szimmetrikus és antiszimmetrikus mátrixok

A szimmetrikus mátrixokat az

$$\mathbf{A}=\tilde{\mathbf{A}} \qquad 8.4\text{-}1$$

összefüggéssel definiáltuk.

Antiszimmetrikusnak az olyan mátrixokat nevezzük, amelyekre fennáll, hogy

$$\mathbf{B}=-\tilde{\mathbf{B}}. \qquad 8.4\text{-}2$$

Definíció szerint tehát

$$B_{kl}=-B_{lk}. \qquad 8.4\text{-}3$$

A definícióból következik, hogy $B_{kk}=0$, tehát az antiszimmetrikus mátrix diagonális elemei eltűnnek. Tetszőleges **A** mátrixból egyszerűen képezhetünk szimmetrikus és antiszimmetrikus mátrixokat. Tekintsük ugyanis a

$$\mathbf{B}=\mathbf{A}+\tilde{\mathbf{A}} \qquad 8.4\text{-}4$$

és

$$\mathbf{C}=\mathbf{A}-\tilde{\mathbf{A}} \qquad 8.4\text{-}5$$

mátrixokat.

A 8.4-4-ben szereplő **B** mátrix B_{kl} elemei a

$$B_{kl}=A_{kl}+\tilde{A}_{kl}=A_{kl}+A_{lk}$$

formulával határozhatók meg. Leolvasható, hogy

$$B_{kl}=B_{lk}.$$

A

$$\mathbf{B}=\mathbf{A}+\tilde{\mathbf{A}}$$

mátrix tehát szimmetrikus mátrix. Hasonló módon belátható, hogy a **C** mátrix antiszimmetrikus.

A fentiek alapján adódik, hogy tetszőleges **A** mátrix egy szimmetrikus és egy antiszimmetrikus mátrix összegére bontható.

Ugyanis

$$\mathbf{A}=\frac{1}{2}(\mathbf{A}+\tilde{\mathbf{A}})+\frac{1}{2}(\mathbf{A}-\tilde{\mathbf{A}}), \tag{8.4-6}$$

ahol a jobb oldal első tagja szimmetrikus, a második tagja pedig antiszimmetrikus mátrix.

8.5. A diadikus szorzat

Legyen **a** egy N komponensű, és **b** egy M komponensű mátrix. **a** és **b** diadikus szorzatának nevezzük és

$$\mathbf{A}=\mathbf{a}\circ\mathbf{b}\text{-vel} \tag{8.5-1}$$

jelöljük azt a mátrixot, amelynek komponensei az

$$A_{kl}=a_k b_l \quad (k=1, 2, \ldots, N; l=1, 2, \ldots, M) \tag{8.5-2}$$

mennyiségek. Itt a_k és b_l **a** és **b** komponenseit jelentik.

A definícióból következik, hogy egy **A** mátrix és az **a** ∘ **b** mátrix szorzatára fennállnak az

$$\mathbf{A}(\mathbf{a}\circ\mathbf{b})=(\mathbf{A}\mathbf{a})\circ\mathbf{b} \tag{8.5-3}$$

és az

$$(\mathbf{a}\circ\mathbf{b})\mathbf{A}=\mathbf{a}\circ(\mathbf{b}\mathbf{A})$$

összefüggések. A **c** egydimenziós mátrix és az **a** ∘ **b** mátrix szorzatára pedig az

$$(\mathbf{a}\circ\mathbf{b})\mathbf{c}=\mathbf{a}(\mathbf{b}\mathbf{c}) \tag{8.5-4}$$

és

$$\mathbf{c}(\mathbf{a}\circ\mathbf{b})=(\mathbf{c}\mathbf{a})\mathbf{b}$$

összefüggések, hiszen pl.

$$[(\mathbf{a}\circ\mathbf{b})\mathbf{c}]_k=\sum_l a_k b_l c_l=a_k\sum_l b_l c_l=[\mathbf{a}(\mathbf{b}\mathbf{c})].$$

8.6. Több dimenziós mátrixok szorzása

A szorzási művelet általánosítása céljából először a diadikus szorzást általánosítjuk. Ezt a műveletet a *mátrixok direkt* szorzatának nevezzük. Az $\overset{(n)}{\mathbf{A}}$ és $\overset{(m)}{\mathbf{B}}$ mátrixok

$$\overset{(n+m)}{\mathbf{C}}=\overset{(n)}{\mathbf{A}}\circ\overset{(m)}{\mathbf{B}} \tag{8.6-1}$$

direkt szorzatán azt a $\overset{(n+m)}{\mathbf{C}}$ mátrixot értjük, amelynek elemeit a

$$C_{\underbrace{kl\dots s}_{n}\,\underbrace{pq\dots z}_{m}} = A_{kl\dots s}\, B_{pq\dots z} \qquad 8.6\text{-}2$$

összefüggés határozza meg. A definícióból látható, hogy a direkt szorzás tetszőleges dimenziójú és az indexekben tetszőleges értelmezési tartományú mátrixok körében elvégezhető. A direkt szorzás eredményeként mindig olyan mátrixhoz jutunk, amelynek dimenziója a tényezőmátrixok dimenzióinak összege.

Látható az is, hogy az $n=1$, $m=1$ eset megegyezik az egydimenziós mátrixok 8.5. fejezetben definiált diadikus szorzatával.

Vizsgáljuk meg most, hogy milyen módon általánosíthatnánk a többi mátrixszorzási műveletet. Az eddig bevezetett szorzási szabályok a következők:

$$\overset{(1)}{\mathbf{a}}\,\overset{(1)}{\mathbf{b}} = \sum_k a_k b_k,$$

$$(\overset{(1)}{\mathbf{a}}\,\overset{(2)}{\mathbf{A}})_l = \sum_k a_k A_{kl},$$

$$(\overset{(2)}{\mathbf{A}}\,\overset{(1)}{\mathbf{a}})_k = \sum_l A_{kl} a_l, \qquad 8.6\text{-}3$$

$$(\overset{(2)}{\mathbf{A}}\,\overset{(2)}{\mathbf{B}})_{kl} = \sum_m A_{km} B_{ml}.$$

A szorzási szabályok közös vonása az, hogy a tényezőket mindig az egymáshoz legközelebb fekvő indexpárra kell összegezni.

Ennek a szabálynak közvetlen általánosításaként például az

$$\overset{(3)}{\mathbf{A}}\,\overset{(3)}{\mathbf{B}} = \overset{(4)}{\mathbf{C}}$$

szorzat elemeit a

$$\overset{(4)}{C}_{klmn} = (\overset{(3)}{\mathbf{A}}\,\overset{(3)}{\mathbf{B}})_{klmn} = \sum_s A_{kls} B_{smn} \qquad 8.6\text{-}4$$

összefüggéssel definiáljuk.

Általában egy m és n dimenziós mátrix szorzata definiálható úgy, hogy egy $m+n-2$ dimenziós mátrixra vezessen. Az eredményben az egyes elemek indexeit a szorzó első $m-1$ indexe és a szorzandó utolsó $n-1$ indexe adja meg.

A

$$\overset{(m+n-2)}{\mathbf{C}} = \overset{(m)(n)}{\mathbf{A}\,\mathbf{B}}$$

szorzat elemei a

$$C_{\underbrace{kl\ldots n}_{m-1}\underbrace{r\ldots z}_{n-1}} = \sum_p A_{\underbrace{kl\ldots n}_{m-1}p} B_{p\underbrace{r\ldots z}_{n-1}}$$

számok, ha **A** elemei $A_{kl\ldots np}$, **B** elemei pedig $B_{pr\ldots z}$ voltak.

Az így definiált mátrixszorzás és a direkt szorzat között érdekes kapcsolat állapítható meg. A mátrixszorzat ugyanis a direkt szorzatból úgy kapható meg, hogy a szorzó utolsó indexét és a szorzandó első indexét egybeejtjük, majd összegezünk az egybeejtett indexpárra. Tehát az **A**, **B** mátrixok szorzatának elemeire fennáll a

$$C_{kl\ldots z} = \sum_s (\overset{(n)}{\mathbf{A}} \circ \overset{(m)}{\mathbf{B}})_{kl\ldots ssq\ldots z} = \sum_s A_{kl\ldots s} B_{sq\ldots z} \qquad 8.6\text{-}5$$

összefüggés. A művelet természetesen csak akkor végezhető el, ha az egybeejtett indexek értelmezési-tartománya megegyezik.

Több indexpár egybeejtésével, majd szummázással további szorzási műveleteket is bevezethetünk. Háromindexes mátrixok esetén például az

$$((\overset{(3)}{\mathbf{A}}\overset{(3)}{\mathbf{B}})) = \overset{(2)}{\mathbf{C}} \qquad 8.6\text{-}6$$

szorzat elemeit a

$$C_{kl} = \sum_{st} A_{kst} B_{tsl} \qquad 8.6\text{-}7$$

összefüggéssel definiálhatjuk. A szummázást először a legközelebb fekvő indexpárra, majd a megmaradó legközelebbi indexpárra kell elvégezni. A jelölésben a kettős zárójel arra utal, hogy két indexpárra kell összegezni.

Általában tetszőleges mátrix esetén két indexet egybeejthetünk, majd összegezhetünk rá, ha az indexek értelmezési tartománya megegyezik. Ezt a műveletet *kontrakciónak* nevezzük. Tehát

$$\underbrace{\overset{(n-2)}{A}_{kl\ldots z}}_{n-2\text{ index}} = \sum_{s=s'=1}^{N} \underbrace{\overset{(n)}{A}_{kl\ldots s\ldots s'\ldots z}}_{n\text{ index}}\,.$$

Kétdimenziós kvadratikus mátrixok esetén például a kontrakció eredményeként a főátlóban álló elemek összegét kapjuk. Ezt az összeget a mátrix *spurjának* nevezzük.

$$\sum_{k=1}^{N} A_{kk} = \text{spur } \mathbf{A}. \qquad 8.6\text{-}8$$

A 8.6-3-ban felsorolt különböző típusú szorzatok mindegyike a direkt szorzat kontrakciójaként fogható fel.

9. Homogén lineáris transzformációk mátrixreprezentációja

A 8. pontban kifejtett mátrixalgebra alkalmas a lineáris operátorok — ezek között az ortogonális operátorok — tulajdonságainak kifejtésére.

Legyen $\underline{\mathbf{A}}$ egy lineáris operátor. Az

$$\bar{\underline{\mathbf{a}}} = \underline{\mathbf{A}}\underline{\mathbf{a}} \qquad 9.1$$

összefüggést a $\mathcal{K}$ reprezentációban az

$$\bar{a}_k = \sum_l A_{kl} a_l \qquad 9.2$$

formulával, ill. a mátrixszorzási jelölést felhasználva, az

$$\bar{\mathbf{a}} = \mathbf{A}\mathbf{a} \qquad 9.3$$

képlettel fejezhetjük ki, ahol az $\mathbf{A} = \mathcal{K}(\underline{\mathbf{A}})$ az a mátrix, amely az $\underline{\mathbf{A}}$ operátor $\mathcal{K}$-beli reprezentációja. Két lineáris operátor egymás utáni alkalmazását mátrixszorzat segítségével fejezzük ki.

Legyen

$$\bar{\underline{\mathbf{a}}} = \underline{\mathbf{A}}\underline{\mathbf{a}} \quad \text{és} \quad \bar{\bar{\underline{\mathbf{a}}}} = \underline{\mathbf{B}}\bar{\underline{\mathbf{a}}}.$$

Akkor a $\mathcal{K}$ reprezentációban

$$\bar{\mathbf{a}} = \mathbf{A}\mathbf{a} \quad \text{és} \quad \bar{\bar{\mathbf{a}}} = \mathbf{B}\bar{\mathbf{a}} = \mathbf{B}(\mathbf{A}\mathbf{a}),$$

illetve

$$\mathbf{B}(\mathbf{A}\mathbf{a}) = (\mathbf{B}\mathbf{A})\mathbf{a} = \mathbf{C}\mathbf{a}, \qquad 9.4$$

ahol

$$\mathbf{C} = \mathbf{B}\mathbf{A}.$$

A mátrixszorzási szabály értelmében tehát

$$C_{kl} = \sum_m B_{km} A_{ml},$$

vagyis az $\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$ transzformációk egymásutánját egy $\underline{\mathbf{C}}$ transzformáció helyettesíti, ahol

$$\mathbf{C} = \mathcal{K}(\underline{\mathbf{C}}),$$

vagyis $\underline{\mathbf{C}}$ $\mathcal{K}$-beli reprezentációja a $\underline{\mathbf{B}}$ és $\underline{\mathbf{A}}$ operátorok $\mathbf{B}$, $\mathbf{A}$ reprezentációjának mátrixszorzata.

Megállapitható még, hogy az egy- és kétdimenziós mátrixok szorzatára fennáll az

$$\mathbf{A}\mathbf{a} = \mathbf{a}\tilde{\mathbf{A}} \qquad 9.5$$

összefüggés, hiszen

$$(\mathbf{Aa})_k = \sum_l A_{kl} a_l = \sum_l a_l \tilde{A}_{lk} = (a\tilde{A})_k.$$

Amennyiben **A** és **a** lineáris operátor, ill. vektor reprezentációja, akkor $\tilde{\mathbf{A}}$ annak az $\underline{\tilde{\mathbf{A}}}$ operátornak a reprezentációja, amelyre teljesül a

$$\mathbf{b}(\mathbf{Aa}) = \mathbf{a}(\tilde{\mathbf{A}}\mathbf{b}) \qquad 9.6$$

összefüggés. 9.6-ban a zárójelezés az operátor és a két vektor szorzásának sorrendjét szabja meg. Látható, hogy ez a sorrend nem változtatható meg.

9.1. Az ortogonális transzformációk reprezentációja

Az ortogonális transzformációk a lineáris transzformációk különleges esetét képviselik. Leírásukra tehát alkalmas a fenti módszer. A $\underline{\mathbf{O}}$ ortogonális operátor reprezentációja legyen

$$\mathbf{O} = \mathcal{K}(\underline{\mathbf{O}}),$$

ahol

$$O_{kl} = \underline{\mathbf{e}}^{(k)} \underline{\mathbf{e}}^{(l)*} \quad (k, l = 1, 2, 3).$$

Itt $\underline{\mathbf{e}}^{(l)*}$ az $\underline{\mathbf{e}}^{(l)}$ alapvektor $\underline{\mathbf{O}}$ szerinti elforgatottja.

9.1.1. Az ortogonalitási relációk

9.5-ben $\underline{\mathbf{A}}$ helyett $\underline{\mathbf{O}}$-t írva, a $\mathcal{K}$ reprezentációban

$$\mathbf{a}^* = \mathbf{Oa} = \mathbf{a}\tilde{\mathbf{O}}, \qquad 9.1\text{-}1$$

vagy $\underline{\mathbf{O}}$-t egy $\underline{\mathbf{b}}$ vektorra alkalmazva,

$$\mathbf{b}^* = \mathbf{Ob} = \mathbf{b}\tilde{\mathbf{O}}. \qquad 9.1\text{-}2$$

Az $\underline{\mathbf{O}}$ operátornak sajátossága azonban, hogy

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \underline{\mathbf{a}}^* \underline{\mathbf{b}}^*.$$

A $\mathcal{K}$ reprezentációban ez a szorzat 9.1-1 és 9.1-2 felhasználásával az

$$\mathbf{a}^*\mathbf{b}^* = \mathbf{a}\tilde{\mathbf{O}}\mathbf{O}\mathbf{b} \qquad 9.1\text{-}3$$

alakot ölti.

Minthogy 9.1-3 tetszőleges $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ vektorra érvényes, 8.2-21 szerint

$$\tilde{\mathbf{O}}\mathbf{O} = \mathbf{E}. \qquad 9.1\text{-}4$$

Ezzel az ortogonalitási reláció 7.5-10 által kifejezett formájához jutottunk.

9.1.2. Az inverz transzformáció

Az ortogonalitási reláció segítségével egy $\underline{\mathbf{O}}$ operátor által definiált elforgatás megfordítottját is leírhatjuk. Szorozzuk 9.1-1-et $\tilde{\mathbf{O}}$-val balról, azt kapjuk, hogy

$$\tilde{\mathbf{O}}\mathbf{a}^* = \tilde{\mathbf{O}}\mathbf{O}\mathbf{a},$$

tehát 9.1-4 segítségével

$$\mathbf{a} = \tilde{\mathbf{O}}\mathbf{a}^* \qquad 9.1\text{-}5$$

vagy

$$\mathbf{a} = \mathbf{a}^*\mathbf{O}. \qquad 9.1\text{-}6$$

Tehát amennyiben $\underline{\mathbf{O}}^{-1}$-nek nevezzük azt a forgatási operátort, amely az $\underline{\mathbf{O}}$ forgatást visszaforgatja, akkor

$$\mathcal{K}(\underline{\mathbf{O}}^{-1}) = \tilde{\mathbf{O}}, \quad \text{ha} \quad \mathcal{K}(\underline{\mathbf{O}}) = \mathbf{O}. \qquad 9.1\text{-}7$$

Az inverz forgatást tehát a $\mathcal{K}$ rendszerben az egyes forgatási operátorok $\mathbf{O}$ reprezentációinak $\tilde{\mathbf{O}}$ transzponált mátrixa reprezentálja.

Szorozzuk 9.1-5-öt $\mathbf{O}$-val balról, azt kapjuk, hogy

$$\mathbf{O}\mathbf{a} = \mathbf{O}\tilde{\mathbf{O}}\mathbf{a}^*, \qquad 9.1\text{-}8$$

tehát 9.1-1 segítségével

$$\mathbf{a}^* = \mathbf{O}\tilde{\mathbf{O}}\mathbf{a}^*. \qquad 9.1\text{-}9$$

9.1-9 minden $\underline{\mathbf{a}}^*$ vektor $\mathbf{a}^*$ reprezentációjára érvényes (hiszen minden $\underline{\mathbf{a}}^*$ vektor egy megfelelő $\underline{\mathbf{a}} = \underline{\mathbf{O}}^{-1}\underline{\mathbf{a}}^*$ vektor elforgatottja), így tehát 9.1-9-ből következik, hogy

$$\mathbf{O}\tilde{\mathbf{O}} = \mathbf{E}. \qquad 9.1\text{-}10$$

9.1-10 hasonlít 9.1-4-hez, de valójában más összefüggéseket reprezentál. Erről meggyőződhetünk, ha 9.1-4-et, illetve 9.1-10-et komponensekben írjuk fel. Ugyanis 9.1-4 szerint

$$\sum_m O_{mk} O_{ml} = \delta_{kl},$$

illetve 9.1-10 szerint

$$\sum_m O_{km} O_{lm} = \delta_{kl},$$

tehát az egyik esetben az első, a másik esetben a második indexpár szerint kell összegezni. A 9.1-4 és 9.1-10 összefüggések azonban nem függetlenek — hiszen az utóbbiakat az elsőkből levezettük.

Az ortogonalitási relációk csak hat független összefüggést adnak a kilenc O_{kl} érték között. Kézenfekvő tehát, hogy az ortogonális mátrixokat három független paraméter segítségével határozhatjuk meg. Egy ilyen paraméteres előállítást korábban már megadtunk.

Megjegyezzük még, hogy a fentiek szerint éppen a 9.1-4, illetve 9.1-10 összefüggések azok a szükséges és elégséges feltételek, amelyeket egy lineáris operátor reprezentációjának ki kell elégítenie ahhoz, hogy ortogonális operátort írjon le.

Továbbá, minthogy a skaláris szorzatból indultunk ki, a 9.1-4, illetve 9.1-10-nek eleget tevő operátorok forgatást vagy tükrözést jellemeznek.

9.2. Két elforgatás egymásutánja

Legyen

$$\underline{\mathbf{a}}^* = \underline{\mathbf{O}}\,\underline{\mathbf{a}}, \quad \underline{\mathbf{a}}^{**} = \underline{\mathbf{P}}\,\underline{\mathbf{a}}^*, \qquad 9.2\text{-}1$$

tehát az $\underline{\mathbf{a}}$-ból $\underline{\mathbf{a}}^*$-hoz jutunk, ha először $\underline{\mathbf{O}}$ szerint, azután $\underline{\mathbf{P}}$ szerint forgatunk. $\mathcal{K}$-beli reprezentációban tehát

$$\mathbf{a}^* = \mathbf{O}\mathbf{a}, \quad \mathbf{a}^{**} = \mathbf{Q}\mathbf{a}, \qquad 9.2\text{-}2$$

ahol

$$\mathbf{Q} = \mathbf{P}\mathbf{O}$$

és

$$\mathbf{O} = \mathcal{K}(\underline{\mathbf{O}}), \quad \mathbf{P} = \mathcal{K}(\underline{\mathbf{P}})$$

az $\underline{\mathbf{O}}$ és $\underline{\mathbf{P}}$ operátorok reprezentációja. Minthogy $\mathbf{O}$, $\mathbf{P}$ ortogonális operátorokat reprezentál,

$$\mathbf{P}\tilde{\mathbf{P}} = \mathbf{O}\tilde{\mathbf{O}} = \mathbf{E}, \qquad 9.2\text{-}3$$

viszont 9.2-2 szerint

$$\tilde{\mathbf{Q}} = \widetilde{(\mathbf{P}\mathbf{O})} = \tilde{\mathbf{O}}\tilde{\mathbf{P}},$$

így

$$\mathbf{Q}\tilde{\mathbf{Q}} = \mathbf{P}\mathbf{O}\tilde{\mathbf{O}}\tilde{\mathbf{P}} = \mathbf{P}\tilde{\mathbf{P}} = \mathbf{E},$$

tehát $\underline{\mathbf{Q}}$ is ortogonális operátor. Két forgatás egymásutánja tehát ismét forgatást hoz létre. Ez az eredmény magától értetődő, így megfontolásaink ellentmondásmentességét támasztja alá.

10. Permutációs operátorok

10.1. A csoport fogalma

Az előző pontban az ortogonális mátrixok tulajdonságaival foglalkoztunk. Megállapítottuk, hogy

a) az ortogonális mátrixok szorzata ismét ortogonális mátrix;
b) az ortogonális mátrixok szorzása asszociatív művelet;
c) az $\mathbf{E}$ egységmátrix is ortogonális mátrix;
d) minden ortogonális mátrixnak van inverze, úgyhogy

$$\mathbf{O}\tilde{\mathbf{O}}=\mathbf{E}, \quad \tilde{\mathbf{O}}\mathbf{O}=\mathbf{E}.$$

A fenti *a*), *b*), *c*), *d*) tulajdonságokkal szokás a *csoport* fogalmát definiálni. Az általános definíciót így fogalmazzuk:

Legyenek a, b, c, ... egy G halmaz elemei, amelyek között értelmeztünk egy műveletet. A műveletet a továbbiakban szorzásnak nevezzük, és pl. az a és b elemek szorzatát ab-vel jelöljük. A G halmazt csoportnak nevezzük, ha teljesülnek a következő axiómák:

1. Amennyiben a és b G eleme, akkor

$$ab=c \tag{10.1-1}$$

G-nek egyértelműen meghatározott eleme.

2. A művelet asszociatív, azaz

$$a(bc)=(ab)c. \tag{10.1-2}$$

3. Létezik egységelem, azaz G-nek van olyan e eleme, amellyel tetszőleges x elemet szorozva, az x elemet kapjuk eredményül:

$$ex=x. \tag{10.1-3}$$

4. G tartalmazza tetszőleges a elemének inverz elemét. Az a elem b inverzét a

$$ba=ab=e \tag{10.1-4}$$

összefüggés definiálja.

A 10.1-1, 10.1-2, 10.1-3, 10.1-4 összefüggéseket csoportaxiómáknak szokás nevezni.

Hangsúlyozzuk, hogy a szorzási művelet kommutativitását nem kötöttük ki. Amennyiben még ez is teljesül, akkor a csoportot Abel-féle vagy kommutatív csoportnak nevezzük. Láttuk, hogy az ortogonális mátrixok, illetve transzformációk csoportot alkotnak.

10.2. A permutációs csoport

A permutáció fogalma

Legyen A, B, C, ..., S n db elem, amelyet n helyre helyezünk el úgy, hogy minden helyre csak egy elem kerüljön. Az A, B, C, ..., S elemekből ily módon $N!$ számú különböző elemsorozatot hozhatunk létre.

Induljunk ki az elemek egy megadott sorrendjéből:

A	B	.	.	.	S
1	2	3	.	.	n

10.2-1

Permutálásnak nevezzük azt a műveletet, amikor egy adott sorrendű elemsorozatot *átrendezünk*, s így az eredeti elemek egy másik sorrendjéhez jutunk. Így a permutálás elemsorozatokra értelmezett operáció. A permutációs operátort úgy jellemezhetjük, hogy megadjuk, hogy az egyes helyeken levő elemeket milyen más helyekre viszi át.

Egy permutációs operátor megadható például az alábbi téglalapmátrixszal:

$$\mathbf{P}=\begin{pmatrix} 1 & 2 & \dots & n \\ 1' & 2' & \dots & n' \end{pmatrix}, \qquad 10.2\text{-}2$$

ahol az egymás alatti számok azt jelentik, hogy ez az operáció az első helyen álló elemet az $1'$, a második helyen álló elemet a $2'$ stb. helyekre viszi át.

Amennyiben megállapodunk abban, hogy az átrendezést mindig az 1, 2, ..., n sorrendben adjuk meg, akkor a fenti $\mathbf{P}$ permutációs operátort már az alsó

$$1', 2', \dots, n'$$

is jellemzi.

Azonban nem mindig célszerű az a jelölés, ahol az eredeti sorrendből indulunk ki; vegyük például négy elem egy permutációjának

$$\mathbf{P}=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \qquad 10.2\text{-}3$$

átrendezését.

Erre a permutációra jellemző, hogy az első helyen levő elemet a második helyre, a második helyen levőt a harmadik helyre visszük stb. E permutációt azonban például így is jelölhetjük:

$$\mathbf{P}=\begin{pmatrix} 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{pmatrix}; \qquad 10.2\text{-}4$$

a 10.2-2 és 10.2-4 jelölések nyilván ugyanazt a permutációt jelzik. Így 10.2-3 helyett írhatjuk azt is, hogy

$$\mathbf{P}=\begin{pmatrix} a & b & \dots & s \\ a' & b' & \dots & s' \end{pmatrix}, \qquad 10.2\text{-}5$$

ahol az $a\ a'$, $b\ b'$ párok az $1\ 1'$, $2\ 2'$ párokkal azonosak, tehát csak az oszlopok sorrendjét változtattuk meg. A permutáció 10.2-5 szerinti jelölése n utasítást jelent; ezeket az utasításokat különböző sorrendben is megadhatjuk, és így jutunk a 10.2-4 jelöléshez.

N elemen értelmezett permutációs operátorok $N!$ elemből álló csoportot alkotnak. A 10.1-1, 2, 3, 4 csoporttulajdonságok a következő módon teljesülnek:

1. Legyen P és Q két átrendezés.[1] Ha egy adott sorrendet először P szerint, azután az így átrendezett sort Q szerint rendezzük át, egy új átrendezést kapunk, amelyet R-rel jelölhetünk.

Az eredetileg az n-edik helyen levő elem először az n'-edik helyre jut. A második átrendezésben pedig ez az elem az n'-edik helyről az n''-edik helyre kerül. A két átrendezést egy lépésben is végrehajthatjuk úgy, hogy az n-edik helyen levő elemet rögtön az n''-edik helyre visszük.

Írhatjuk tehát, hogy

$$\mathbf{P}=\begin{pmatrix}1 & \dots & n & \dots & N\\ 1' & \dots & n' & \dots & N'\end{pmatrix}, \quad \mathbf{Q}=\begin{pmatrix}1' & 2' & \dots & n' & \dots & N'\\ 1'' & 2'' & \dots & n'' & \dots & N''\end{pmatrix}$$

és

$$\mathbf{R}=\mathbf{QP}=\begin{pmatrix}1 & 2 & \dots & n & \dots & N\\ 1'' & 2'' & \dots & n'' & \dots & N''\end{pmatrix}.$$

Az R permutációt a Q és P permutációk szorzatának tekinthetjük.

2. A permutációk ilyenfajta egymásutánjára érvényes az asszociatív törvény. Három egymás utáni átrendezést (P, Q, R-et) ugyanis kétféle módon lehet két lépésben végrehajtani. A P-t és Q-t összevonva, egy

$$\mathbf{S}=\mathbf{QR}$$

átrendezést kapunk; ezt végrehajtva, majd a harmadik átrendezést alkalmazva, a végleges

$$\mathbf{T}=\mathbf{RS}=\mathbf{R}(\mathbf{QP})$$

sorrendhez jutunk.

A T sorrendet úgy is elérhetjük, hogy először a P-nek megfelelő átrendezést hajtjuk végre, azután pedig az $(\mathbf{RQ})=\mathbf{U}$ kombinált átrendezést, tehát

$$\mathbf{T}=\mathbf{PU}=(\mathbf{RQ})\mathbf{P},$$

vagyis

$$\mathbf{T}=\mathbf{R}(\mathbf{QP})=(\mathbf{RQ})\mathbf{P}.$$

3. Egységelemnek azt az átrendezést vesszük, ahol semmi sem változik, tehát

$$\mathbf{E}=\begin{pmatrix}1 & \dots & N\\ 1 & \dots & N\end{pmatrix},$$

így $\mathbf{PE}=\mathbf{EP}=\mathbf{P}$, hiszen E nem változtat a sorrenden.

4. Minden permutációnak létezik inverze, hiszen minden átrendezést vissza is rendezhetünk. A visszarendezést az eredeti permutáció reciprokának nevezzük.

[1] P és Q ... ebben a fejezetben a $\underline{\mathbf{P}}$ és $\underline{\mathbf{Q}}$ ... permutációs operátorok egy reprezentációját jelenti, ami téglalapmátrix is lehet. A reprezentációk szorzatát azonban most nem a szokásos mátrixszorzással kell képeznünk.

Vagyis, ha $\mathbf{P}$ $n \to n'$ átmeneteket tartalmaz, akkor $\mathbf{P}^{-1}$ az $n' \to n$ átmenetekből áll. Tehát

$$\mathbf{P}=\begin{pmatrix} 1 & 2 & \dots & N \\ 1' & 2' & \dots & N' \end{pmatrix} \quad \text{és} \quad \mathbf{P}^{-1}=\begin{pmatrix} 1' & 2' & \dots & N' \\ 1 & 2 & \dots & N \end{pmatrix}.$$

Látjuk tehát, hogy a csoporttulajdonságok értelmezhetők és érvényesülnek.

Az asszociatív törvényből következik az alábbi fontos tulajdonság.

Legyen $\mathbf{P}^{-1}$ és $\mathbf{Q}^{-1}$ a $\mathbf{P}$, ill. $\mathbf{Q}$ permutációk inverze. Tehát

$$\mathbf{P}\mathbf{P}^{-1}=\mathbf{P}^{-1}\mathbf{P}=\mathbf{E},$$

$$\mathbf{Q}\mathbf{Q}^{-1}=\mathbf{Q}^{-1}\mathbf{Q}=\mathbf{E}.$$

Akkor

$$\mathbf{Q}\mathbf{P}=\mathbf{R}$$

-ből következik, hogy

$$\mathbf{R}^{-1}=\mathbf{P}^{-1}\mathbf{Q}^{-1},$$

hiszen jobbról szorozva $\mathbf{R}$-rel,

$$\mathbf{R}^{-1}\mathbf{R}=(\mathbf{P}^{-1}\mathbf{Q}^{-1})(\mathbf{Q}\mathbf{P})=\mathbf{P}^{-1}(\mathbf{Q}^{-1}\mathbf{Q})\mathbf{P}=\mathbf{P}^{-1}\mathbf{P}=\mathbf{E}.$$

Több tényezőre kiterjesztve, azt is kapjuk, hogy

$$(\mathbf{P}\mathbf{Q}\dots\mathbf{S})^{-1}=\mathbf{S}^{-1}\dots\mathbf{Q}^{-1}\mathbf{P}^{-1}.$$

Példaként az $N=2$, 3 esetet tárgyaljuk részletesebben.

10.2.1. Kételemű elemsorozatokon értelmezett operátorcsoport

Az A, B elemekből két különböző sorrendű elemsorozat képezhető:

$$A, B \quad \text{és} \quad B, A.$$

Ennek megfelelően két permutációs operátort értelmezhetünk: Ezek, a fenti mátrixjelölést alkalmazva,

$$\mathbf{P}_1=\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix},$$

$$\mathbf{P}_2=\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

$\mathbf{P}_1$ az elemek eredeti sorrendjén nem változtat, $\mathbf{P}_2$ pedig a sorrendet felcseréli, tehát

$$P_1(A, B) = (A, B),$$
$$P_2(A, B) = (B, A).$$

A permutációk szorzótáblája tehát a következő:

	P_1	P_2
P_1	P_1	P_2
P_2	P_2	P_1

Az egységelem az $E = P_1$, hiszen nem változtatja meg az adott sorrendet.

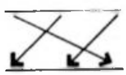
Mind P_1-nek, mind P_2-nek önmaga az inverze, mert a fenti tábla értelmében

$$P_1P_1 = E, \quad \text{és} \quad P_2P_2 = E.$$

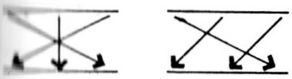
10.2.2. Háromelemű elemsorozatokon értelmezett operátorcsoport

A következőkben az A, B, C elemekből alkotott elemsorozatokon értelmezett permutációs operátorok tulajdonságaival foglalkozunk. Minthogy három elem hat permutációt tesz lehetővé, ebben az esetben már hatféle különböző operátort definiálhatunk.

Ezeket az operátorokat a már használt téglalapmátrixokkal egyenértékű, de sokkal szemléletesebb geometriai ábrázolásmód segítségével állítjuk elő. Egy operátort pl. a *10.1. ábrán* látható sémával reprezentálhatunk. Ez azt jelenti, hogy a kérdéses operátor egy tetszőleges háromelemű elemsorozatot transzformál úgy, hogy az első helyen álló elemet a harmadik helyre, a második helyen állót az első helyre, a harmadik helyen állót pedig a második helyre viszi át. A nyilak jól szemléltetik a helycseréket. A geometriai ábrázolásmód segítségével egyszerűen meghatározhatjuk két operátor szorzatát is, tehát két egymás utáni elemcsere eredményét *(10.2. ábra)*.



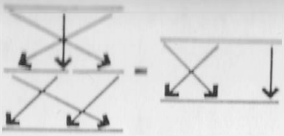
10.1. ábra



10.2. ábra

Még jobban látható az elemek átrendezése, ha reprezentációikat egymás alá rajzoljuk. Így az operátorok szorzata a *10.3. ábra* segítségével szemlélhető. A lehetséges hatféle operátort a *10.4. ábra* szerinti sémák szemléltetik.

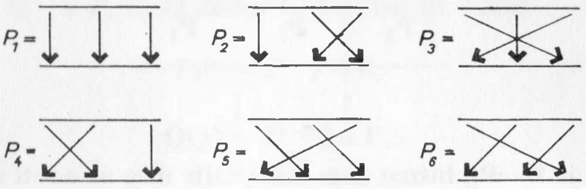
Az operátorok ábráiból látható, hogy a P_1 operátor nem változtat egy adott elemsor sorrendjén, P_2, P_3 és P_4 pedig egy elemet a helyén hagy, a másik kettő sorrendjét pedig felcseréli.



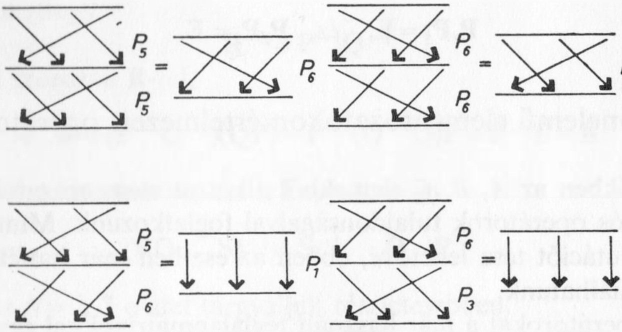
10.3. ábra

Ezeket az operátorokat, amelyek csak két elem sorrendjét cserélik fel, *transzpozíciós* operátoroknak nevezzük. A $\mathbf{P}_5$ és $\mathbf{P}_6$ operátorok az elemek sorrendjét ciklikusan változtatják.

Az operátorok szorzata a fentiekben ismertetett grafikus eljárás segítségével egyszerűen meghatározható. Példaként még a $\mathbf{P}_5$ és $\mathbf{P}_6$ operátorok négyzeteit és szorzatát határozzuk meg *(10.5. ábra)*.



10.4. ábra



10.5. ábra

Hasonló módon határozhatjuk meg a többi operátorok szorzatát is, s ennek eredményeként a

$$\mathbf{P}_k\mathbf{P}_l$$

k \ l	1	2	3	4	5	6
1	$\mathbf{E}=\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$	$\mathbf{P}_4$	$\mathbf{P}_5$	$\mathbf{P}_6$
2	$\mathbf{P}_2$	$\mathbf{E}=\mathbf{P}_1$	$\mathbf{P}_5$	$\mathbf{P}_6$	$\mathbf{P}_3$	$\mathbf{P}_4$
3	$\mathbf{P}_3$	$\mathbf{P}_6$	$\mathbf{E}=\mathbf{P}_1$	$\mathbf{P}_5$	$\mathbf{P}_4$	$\mathbf{P}_2$
4	$\mathbf{P}_4$	$\mathbf{P}_5$	$\mathbf{P}_6$	$\mathbf{E}=\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$
5	$\mathbf{P}_5$	$\mathbf{P}_4$	$\mathbf{P}_2$	$\mathbf{P}_3$	$\mathbf{P}_6$	$\mathbf{E}=\mathbf{P}_1$
6	$\mathbf{P}_6$	$\mathbf{P}_3$	$\mathbf{P}_4$	$\mathbf{P}_2$	$\mathbf{E}=\mathbf{P}_1$	$\mathbf{P}_5$

táblázathoz jutunk.

A táblázat alapján megállapítható, hogy ezek az operátorok valóban csoportot alkotnak.

Vegyük észre még, hogy a $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_4$ operátorok kihagyásával az eredeti csoport egy ($\mathbf{P}_1 = \mathbf{E}$, $\mathbf{P}_5$, $\mathbf{P}_6$)

k \ l	$\mathbf{P}_1 = \mathbf{E}$	$\mathbf{P}_5$	$\mathbf{P}_6$
$\mathbf{P}_1 = \mathbf{E}$	$\mathbf{E}$	$\mathbf{P}_5$	$\mathbf{P}_6$
$\mathbf{P}_5$	$\mathbf{P}_5$	$\mathbf{P}_6$	$\mathbf{E}$
$\mathbf{P}_6$	$\mathbf{P}_6$	$\mathbf{E}$	$\mathbf{P}_5$

alcsoportjához jutunk.

10.3. Az N elemű permutációk néhány tulajdonsága

A permutációk gazdag elméletéből csak egy, a továbbiakban szükséges tételre térünk ki. Kimutatjuk, hogy N elem $N!$ permutációját egyértelműen két, egyaránt $\frac{1}{2}N!$ elemet tartalmazó részre választhatjuk szét. Az egyes részekhez tartozó elemeket az elem *páros*, ill. *páratlan* permutációinak nevezzük.

10.3.1. Transzpozíció és szomszédcsere

Egy permutáció során nem szükséges, hogy a cserékben minden hely részt vegyen. Az operáció egyértelműen megadható akkor is, ha csupán a megváltozó helyeket tüntetjük fel. A

$$\mathbf{P} = \begin{Bmatrix} 1 & 2 \ldots k & k+1 \ldots N \\ 1' & 2' \ldots k' & k+1 \ldots N \end{Bmatrix}$$

operátor a rövidebb

$$\mathbf{P} = \begin{pmatrix} 1 & 2 & \ldots & k \\ 1' & 2' & \ldots & k' \end{pmatrix} \qquad 10.3\text{-}1$$

szimbólummal is egyértelműen jellemezhető. (Természetesen nem szükséges, hogy a megváltozó helyek az eredeti sorrendben elöl legyenek.)

Igen fontos szerepet játszanak azok az operátorok, amelyek két helyet felcserélnek. Ezek az operátorok a 10.3-1 egyszerűsített jelölés segítségével a

$$\mathbf{T}=\begin{pmatrix} k, & l \\ l, & k \end{pmatrix} \qquad 10.3\text{-}2$$

szimbólummal jelölhetők. 10.3-2 tehát azt fejezi ki, hogy a k-adik helyen elhelyezett elem az l-edik helyre kerül, az l-edik helyen elhelyezett elem a k-adik helyre kerül, a többi elem helye pedig változatlan marad. A 10.3-2 alakú permutációkat *transzpozícióknak* nevezzük. A 10.3-2 jelölés a transzpozíciók esetén tovább egyszerűsíthető, a transzpozíció jellemzésére elegendő ugyanis a két felcserélődő hely megadása. Ezért a továbbiakban a

$$\mathbf{T}=(k, l) \qquad 10.3\text{-}3$$

jelölést használjuk. Bizonyításainkban különösen jól felhasználhatók azok a transzpozíciók, amelyekben két szomszédos hely cserélődik. Ezeket szomszédcseréknek nevezzük és

$$\mathbf{S}=(k, k+1) \qquad 10.3\text{-}4$$

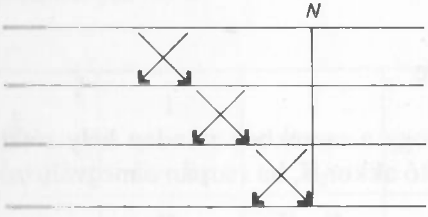
-gyel jelöljük.

A következőkben megmutatjuk, hogy minden permutáció szomszédcserék sorozatára bontható, tehát

$$\mathbf{P}=\mathbf{S}_k\mathbf{S}_{k-1}\ldots\mathbf{S}_1, \qquad 10.3\text{-}5$$

ahol S_k szomszédcseréket, $\mathbf{P}$ pedig egy adott permutációt jelent.

Ahhoz, hogy 10.3-5 érvényességét belássuk, kezdjük a cseréket azzal az $\bar{n}$ hellyel, amely a $\mathbf{P}$ permutáció alkalmazása után az utolsóvá válik. Tehát $\mathbf{P}$ alkalmazása után $\bar{n}\to N$-be kerül. Ha $\bar{n}<N$, akkor ezt a helyet a jobb szomszéddal cseréljük, azután a következő jobb szomszéddal, mindaddig, míg az utolsó helyre jut *(10.6. ábra)*.



10.6. ábra

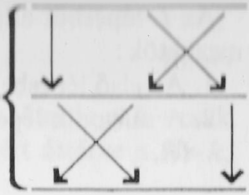
Ezek után ehhez hasonló eljárással az $N-1$-edik helyre kerülő elemet visszük lépésenként az utolsó előtti helyre. Így folytatva, végeredményben — ha nagyon sok lépésben is — elérjük a $\mathbf{P}$ permutáció által rögzített sorrendet.

Megjegyezzük, hogy a szomszédcserék helyett transzpozíciókat alkalmazva is elérhetjük a $\mathbf{P}$ által előírt elrendezést, és ez általában kevesebb lépést igényel, mintha szomszédcserékre szorítkozunk.

Valóban, legyen

$$\mathbf{P}=\mathbf{T}_l\mathbf{T}_{l-1}\ldots\mathbf{T}_1.$$

$\mathbf{T}_1$-nek választhatjuk azt a transzpozíciót, amellyel az $\bar{n}$-edik helyet az utolsó N-edik hellyel cseréljük fel; $\mathbf{T}_2$ az a csere, amellyel az így átrendezett sorrend megfelelő helyét a végső sorrend $N-1$-edik helyére visszük *(10.7. ábra)*. Így lépésről lépésre haladva, ismét előállíthatjuk a végső sorrendet (I. ábra).



10.7. ábra

10.3.2. Páros és páratlan permutációk

A páros és páratlan permutációk bevezetésére első lépésként megmutatjuk, hogy az egységpermutációt

$$\mathbf{E} = \mathbf{S}_L \mathbf{S}_{L-1} \ldots \mathbf{S}_1 \qquad 10.3\text{-}6$$

csak páros számú egymás utáni szomszédcserével állíthatjuk elő. Világosabban kifejezve: alkalmazzuk egy N elemű sorra egymás után az $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_L$ szomszédcseréket. E cserék megváltoztatják az eredeti sorrendet, azonban előfordulhat, hogy egy bizonyos lépés után újra az eredeti sorrend áll elő. Bebizonyítjuk, hogy ez csak páros számú lépés után történhet meg.

Vizsgáljuk ehhez a k-adik hely mozgását a permutáció szomszédcserékkel történő előállításában az egyes lépések során.

Az $\mathbf{S}_1$ szomszédcserében olyan $k \to k'$ átmenet jön létre, amelyben k' a $k+1$, $k-1$, k értékek valamelyikével egyenlő, attól függően, hogy k jobb vagy bal oldali szomszédjával cserél helyet, ill. nem vesz részt cserében.

Hasonló módon a második lépésben $\mathbf{S}_2$-ben olyan $k' \to k''$ átmenet következik be, amelyben k'' értéke

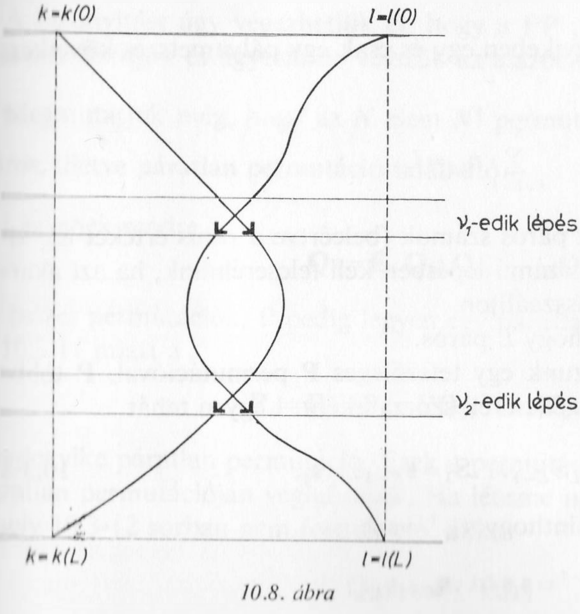
$$k'-1, \quad k'+1 \quad \text{vagy} \quad k'$$

lehet.

A cseréket feltérképezve, megrajzolhatjuk a k-adik hely „pályáját” *(10.8. ábra)*. Az eredetileg k-adik helyen levő elem az n-edik lépés után a $k(n)$-edik helyre kerül. Amennyiben azonban 10.3-6 fennáll, akkor a k-adik hely változásaira felírhatjuk, hogy

$$k(0)=k, \quad k(1),$$
$$k(2), \ldots, k(L)=k,$$

vagyis az L-edik lépés után ismét az eredeti k-adik helyre jutunk vissza.



10.8. ábra

Az L lépésből álló operációsorozat egyes lépései a következő sémával jellemezhetők:

1. Az első lépésben az eredetileg $k(1)$-edik és $l(1)$-edik helyek cserélődnek fel.
2. A második lépésben az eredetileg $k(2)$-edik és $l(2)$-edik helyek cserélődnek fel.

.

.

.

A ν-edik lépésben az eredetileg $k(\nu)$-edik és $l(\nu)$-edik helyek cserélődnek fel.

.

.

.

Az L-edik lépésben az eredetileg $k(L)$-edik és $l(L)$-edik helyek cserélődnek fel.

A $k(1), \ldots, k(L)$ és $l(1), \ldots, l(L)$ helyek a permutációsorozat ábrázolása alapján határozhatók meg. $k(\nu)$ és $l(\nu)$ $(\nu=1\ldots L)$ többször is felveheti a

$$k(\nu)=k$$

és

$$l(\nu)=l \qquad 10.3\text{-}7$$

értékeket. Legyenek $\nu=\nu_1, \nu_2, \ldots, \nu_n$ azon lépések, amelyekben 10.3-6 bekövetkezik. A 10.3-7-nek megfelelő ν_i értékek n_{kl} száma a k és l értékektől függ. Az n_{kl} értéket a permutáció ábrázolása alapján meghatározhatjuk, hiszen n_{kl} éppen a k-adik és l-edik helypályák metszéspontjainak számával egyenlő. (A *10.8. ábrán* látható esetben $n_{kl}=2$.)

Minthogy az L lépés mindegyikében egy és csak egy pályametszés következhet be, kell, hogy

$$\sum_{k,\, l=1}^{N} n_{kl}=L$$

legyen. Az egyes n_{kl}-ek viszont páros számok (beleértve a zérus értéket is), hiszen tetszőleges k, l párt páros számú lépésben kell felcserélnünk, ha azt akarjuk, hogy eredeti sorrendjük visszaálljon.

Ebből azonban következik, hogy L páros.

A következőkben foglalkozzunk egy tetszőleges $\mathbf{P}$ permutációval. $\mathbf{P}$ többféle módon állítható elő szomszédcserék szorzataként. Legyen tehát

$$\mathbf{P}=\mathbf{S}_L\mathbf{S}_{L-1}\ldots\mathbf{S}_1=\mathbf{s}_{K-1}\ldots\mathbf{s}_1, \qquad 10.3\text{-}8$$

10.3-8-ból következik, hogy (minthogy $\mathbf{s}_K^{-1}=\mathbf{s}_K$)

$$\mathbf{P}^{-1}=\mathbf{s}_1\mathbf{s}_2\ldots\mathbf{s}_{K-1}\mathbf{s}_K.$$

Tehát

$$\mathbf{E}=\mathbf{PP}^{-1}=\mathbf{S}_L\mathbf{S}_{L-1}\dots\mathbf{S}_1\mathbf{s}_1\dots\mathbf{s}_{K-1}\mathbf{s}_K. \qquad 10.3\text{-}9$$

10.3-9 az **E** operátor $K+L$ szomszédcseréből álló előállítása. Minthogy **E**-t csak páros számú szomszédcsere szorzatával lehet előállítani, ezért áttérve a $K=k$, $L=l$ jelölésre,

$$k+l=\text{páros szám};$$

így vagy k és l is páratlan, 10.3-10
vagy k és l is páros.

Tehát, ha egy **P** permutáció előállítása a páros számú szomszédcsere szorzata, akkor a szomszédcserékből történő *összes* előállítása páros számú tényezőt tartalmaz. Ezeket a permutációkat *páros permutációknak* nevezzük.

Hasonló módon következik, hogy azon permutációknak, amelyeknek van egy páratlan számú szomszédcseréből történő előállítása, minden előállítása páratlan tényezőt tartalmaz.

Ezeket *páratlan permutációknak* nevezzük.

Látjuk tehát, hogy az N elem $N!$ permutációja egyértelműen páros és páratlan permutációkra osztható.

Belátjuk azt is, hogy ha **P** és **P**′ páratlan permutációk, és **Q** és **Q**′ páros permutációk, akkor

$$\mathbf{PP}' \quad \text{és} \quad \mathbf{QQ}' \quad \text{páros permutáció,}$$

$$\mathbf{PQ} \quad \text{és} \quad \mathbf{P}'\mathbf{Q}' \quad \text{páratlan permutáció.} \qquad 10.3\text{-}11$$

A bizonyítást úgy végezhetjük el, hogy a **PP**′; **QQ**′ szorzatokat szomszédcserékre bontjuk és figyelembe vesszük az asszociatív törvényt.

Megmutatjuk még, hogy az N elem $N!$ permutációja között pontosan $\frac{1}{2}N!$ páros, illetve páratlan permutáció található.

Legyenek rendre

$$\mathbf{Q}_0=\mathbf{E},\ \mathbf{Q}_1,\ \mathbf{Q}_2,\ \dots,\ \mathbf{Q}_n \qquad 10.3\text{-}12$$

az összes permutációk, **P** pedig legyen egy páratlan permutáció.
10.3-11 miatt a

$$\mathbf{P}_0=\mathbf{P},\quad \mathbf{P}_1=\mathbf{PQ}_1,\ \dots,\ \mathbf{P}_n=\mathbf{PQ}_n \qquad 10.3\text{-}13$$

mindegyike páratlan permutáció. Ezek a permutációk az $N!$ permutáció összes páratlan permutációján végigfutnak. Ha létezne ugyanis egy $\mathbf{P}_{n+1}$ permutáció, amely 10.3-12 sorban nem fordult elő, akkor

$$\mathbf{Q}_{n+1}=\mathbf{P}^{-1}\mathbf{P}_{n+1}$$

olyan páros permutáció lenne, amely 10.3-12-ben nem fordult elő — ellentétben a kiindulással, amely szerint 10.3-12 az összes páros permutációt tartalmazza. Látjuk tehát, hogy a páros, illetve páratlan permutációk száma azonos.

Megjegyzés: Az $N=2$ esetben

$$\mathbf{P}_1=\mathbf{E} \quad \text{páros},$$

$$\mathbf{P}_2=(1, 2) \quad \text{páratlan}.$$

Az $N=3$ esetben az eredeti sorrend ciklikus permutációi:

$$(1\ 2\ 3) \quad (2\ 3\ 1) \quad (3\ 1\ 2) \quad \text{páros permutációk;}$$

a nem ciklikus permutációk:

$$(2\ 1\ 3) \quad (1\ 3\ 2) \quad (3\ 2\ 1) \quad \text{páratlanok.}$$

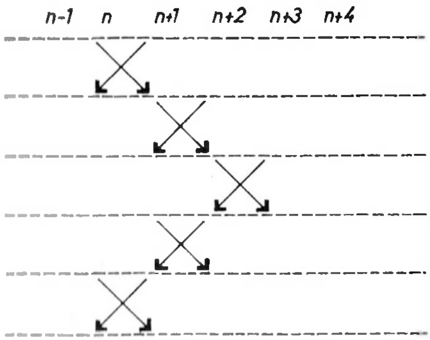
Ebből az is következik, hogy két ciklikus permutáció szorzata ismét ciklikus permutáció!

10.3.3. Permutációk előállítása transzpozíciókkal

Egy transzpozíció pl. ha a k-adik helyet felcseréljük a $k+n$-edik hellyel:

$$\mathbf{T}=(k, n+k)=\begin{pmatrix} 1 & 2\ldots k\ldots n+k\ldots n \\ 1 & 2\ldots n+k\ldots k\ldots n \end{pmatrix},$$

páratlan permutáció. Ezt a következőképpen látjuk be: $\mathbf{P}$-t úgy állíthatjuk elő szomszédcserékkel, hogy k-t $k+1$-gyel, $k+1$-et $k+2$-vel stb. végül $k+n-1$-et $k+n$-nel cseréljük fel. Így n lépésben k-t $n+k$ helyére hoztuk. A $k+n$-edik helyet, amely az utolsó csere folytán a $k+n-1$ helyre jutott, bal felé cserélve $n-1$ szomszédcserével hozhatjuk a k-adik helyre. Ezek a cserék a közbenső helyeket visszahozzák eredeti helyzetükbe. Így $n+(n-1)=2n-1$ lépésben, tehát páratlan szomszédcsere segítségével állítottuk elő a transzpozíciót.



10.9. ábra

Látjuk tehát, hogy a *transzpozíciók* valóban páratlan permutációk. A fenti eredményt $n=3$ esetén az ábrán szemléltetjük *(10.9. ábra)*.

Tehát $(n, n+3)$ öt szomszédcserével előállítható.

Kifejtettük már, hogy tetszőleges permutáció *transzpozíciók* szorzataként is előállítható. Minthogy a transzpozíciók páratlan permutációk, egy páratlan permutáció páratlan, egy páros permutáció csak páros számú transzpozíciók szorzataként állítható elő.

Ezért egy permutáció páros, illetve páratlan voltát úgy is fogalmazhatjuk, hogy a

$$\mathbf{P}=\mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_l$$

permutáció páros, ha l páros; páratlan, ha l páratlan. Ez a definíció egybevág a szomszédcseréknél alkotott definícióval, és minden permutációt egyértelműen páros vagy páratlan permutációnak minősít.

11. Lineáris egyenletrendszerek

11.1. Lineáris egyenletrendszerek felírása mátrixokkal

A 4.6. fejezetben már foglalkoztunk a háromismeretlenes egyenletrendszerek megoldásával. Általában egy lineáris egyenletrendszert az

$$\begin{aligned} A_{11}x_1+A_{12}x_2+\ldots+A_{1N}x_N&=y_1\\ A_{21}x_1+A_{22}x_2+\ldots+A_{2N}x_N&=y_2\\ \vdots\qquad\qquad\qquad \vdots\quad &\quad \vdots\\ A_{M1}x_1+A_{M2}x_2+\ldots+A_{MN}x_N&=y_M \end{aligned} \qquad 11.1\text{-}1$$

alakban írhatunk fel.

A 11.1-1 rendszert többek között eliminációs módszerrel oldhatjuk meg, azaz az egyes egyenleteket megfelelő faktorokkal megszorozzuk, majd az így kapott egyenleteket összeadjuk. Ha a faktorokat megfelelően választjuk, elérhetjük, hogy az új egyenletek nem tartalmazzák az összes ismeretlent. Ily módon lépésről lépésre haladva, egyismeretlenes egyenlethez juthatunk. Az eliminációs módszert a mátrixalgebra segítségével rendszeresebbé (és így egyszerűbbé) tehetjük.

A 11.1-1 egyenletet az

$$\mathbf{A}\mathbf{x}=\mathbf{y} \qquad 11.1\text{-}2$$

formában is felírhatjuk úgy, hogy

$$\mathbf{A}=\begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1N}\\ \vdots & & & \\ A_{M1} & & \ldots & A_{MN}\end{pmatrix} \qquad 11.1\text{-}3$$

egy $N \times M$-es mátrix, és

$$\mathbf{x} = x_1, \ldots, x_N,$$

$$\mathbf{y} = y_1, \ldots, y_M \qquad 11.1\text{-}4$$

egydimenziós mátrixok.

Abban az esetben, ha $N = M$, az ismeretlenek száma egyenlő az egyenletek számával, s ekkor **A** négyzetes mátrix. A következőkben csak ezzel az esettel foglalkozunk.

Szorozzuk a 11.1-2-t egy $N \times N$-es **B** mátrixszal. Azt kapjuk, hogy

$$\mathbf{BAx} = \mathbf{By}. \qquad 11.1\text{-}5$$

11.1-5 teljesülése szükséges (de nem mindig elégséges) feltétel a 11.1-2-t kielégítő **x** értékekre vonatkozóan. Egyszerűen látható, hogy 11.1-5 nem elégséges feltétel, hiszen a $\mathbf{B} = \mathbf{0}$ esetben 11.1-5 semmitmondó, $\mathbf{0} = \mathbf{0}$ egyenletté válik.

Amennyiben sikerül olyan **B** mátrixot találni, amelyre

$$\mathbf{BA} = \mathbf{E}, \qquad 11.1\text{-}6$$

akkor 11.1-5-ből és 11.1-6-ból következik, hogy

$$\mathbf{x} = \mathbf{By}, \qquad 11.1\text{-}7$$

minthogy $\mathbf{Ex} = \mathbf{x}$.

11.1-7 szükséges feltételt ad **x**-re. Azt azonban, hogy 11.1-7 valóban 11.1-1, illetve 11.1-2 megoldása-e, behelyettesítéssel kell eldönteni. 11.1-1 és 11.1-7-ből következik, hogy

$$\mathbf{Ax} = \mathbf{ABy}.$$

Tehát 11.1-7 akkor megoldása a 11.1-2-nek, ha

$$\mathbf{ABy} = \mathbf{y}.$$

Ez csak akkor igaz, ha

$$\mathbf{AB} = \mathbf{E}. \qquad 11.1\text{-}8$$

A megoldás szükséges és elégséges feltétele tehát egy **B** mátrix létezése, amelyre

$$\mathbf{AB} = \mathbf{BA} = \mathbf{E}. \qquad 11.1\text{-}9$$

A **B** mátrixot, amely 11.1-9-nek eleget tesz,

$$\mathbf{B} = \mathbf{A}^{-1}$$

-gyel jelöljük és **A** reciprok mátrixának nevezzük.

Felhívjuk a figyelmet arra, hogy a

$$\mathbf{BA} = \mathbf{E} \qquad 11.1\text{-}10$$

egyenlőségből nem következik automatikusan az

$$\mathbf{AB}=\mathbf{E} \qquad 11.1\text{-}11$$

összefüggés. A 11.1-10-nek eleget tevő mátrixokat **A** bal oldali reciprokainak, a 11.1-11-et kielégítő mátrixokat pedig **A** jobb oldali reciprokainak nevezzük.

A reciprok mátrix explicit előállításával a későbbiek során foglalkozunk. Most csak annyit bizonyítunk, hogy amennyiben az **A** mátrixhoz létezik olyan **B** mátrix, amelyre

$$\mathbf{AB}=\mathbf{E} \quad \text{és} \quad \mathbf{BA}=\mathbf{E}, \qquad 11.1\text{-}12$$

akkor **B** az **A** mátrix egyetlen reciprok mátrixa. Tegyük fel ugyanis, hogy $\mathbf{B}_1$ az **A** mátrix **B**-től különböző bal oldali reciproka, tehát

$$\mathbf{B}_1\mathbf{A}=\mathbf{E}. \qquad 11.1\text{-}13$$

Szorozzuk be 11.1-13-at jobbról **B**-vel. A

$$\mathbf{B}_1\mathbf{AB}=\mathbf{EB}$$

egyenlethez jutunk. Az asszociativitás és 11.1-12 első egyenletének felhasználásával a

$$(\mathbf{B}_1\mathbf{A})\mathbf{B}=\mathbf{B}_1(\mathbf{AB})=\mathbf{B}_1=\mathbf{B}. \qquad 11.1\text{-}14$$

Tehát

$$\mathbf{B}_1=\mathbf{B}.$$

Hasonló módon láthatjuk be, hogy **A** tetszőleges jobb oldali reciproka is **B** kell legyen. Tehát, ha **B** egyszerre bal és jobb oldali reciprok, akkor **A**-nak ez az egyetlen reciproka létezik.

11.2. A determináns fogalma

Az $N=M=3$ esetben megmutattuk, hogy a 11.1-1 típusú egyenletrendszer akkor és csak akkor rendelkezik egyértelmű megoldással, ha

$$\det \mathbf{A}\neq 0, \qquad 11.2\text{-}1$$

ahol

$$\det \mathbf{A}=\sum_{i,k,l} \varepsilon_{ikl}A_{1i}A_{2k}A_{3l}. \qquad 11.2\text{-}2$$

Megmutatjuk, hogy az $N=3$ esetre vonatkozó eredmény tetszőleges N értékre általánosítható. A 11.2-2 definíciót az $N\neq 3$ esetekre megfelelő módon kiterjesztve adódik, hogy az

$$\mathbf{Ax}=\mathbf{y}$$

egyenlet egy és csak egy megoldásrendszerrel rendelkezik, ha

$$\det \mathbf{A}\neq 0.$$

Beláthjuk ugyanis, hogy ha $\det \mathbf{A} \neq 0$ teljesül, akkor $\mathbf{A}$-nak egy és csak egy reciprok mátrixa létezik, tehát ebben az esetben

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \qquad 11.2\text{-}3$$

egyértelműen adja a 11.2-2 megoldásrendszerét.

E tétel bizonyításakor mindenekelőtt $\det \mathbf{A}$-t kell az $N \neq 3$ esetekre definiálni. Tetszőleges $\mathbf{A}$ mátrix determinánsának definíciója a következő:

$$\det \mathbf{A} = \sum_{k,\, l,\, \dots,\, s} \varepsilon_{kl\dots s} A_{1k} A_{2l} \dots A_{Ns}, \qquad 11.2\text{-}4$$

ahol k, l, ..., s olyan indexeket jelent, amelyek mindegyike az 1, ..., N értékeket veszi fel.

$\varepsilon_{kl\dots s}$ az N indexű Levi—Civita-szimbólum, amelynek definíciója:

$$\varepsilon_{kl\dots s} = \begin{cases} 1, & \text{ha } kl\dots s \text{ az } 1\ 2\dots N \text{ páros permutációja;} \\ -1, & \text{ha } kl\dots s \text{ az } 1\ 2\dots N \text{ páratlan permutációja;} \\ 0, & \text{ha } kl\dots s \text{ az } 1\ 2\dots N \text{ nem permutációja.} \end{cases} \qquad 11.2\text{-}5$$

A fenti definíció az $N = 3$ esetén érvényes definícióval egybeesik, hiszen három elem ciklikus permutációi páros permutációk. A következőkben megmutatjuk, hogy a 11.2-4 általánosítást felhasználva, adódik a 11.2-3 formula.

Az $N = 1$ esetben érvényes, hogy

$$\varepsilon_1 = 1$$

és

$$\det \mathbf{A} = \mathbf{A}_{11}.$$

$N = 2$ esetén

$$\varepsilon_{11} = \varepsilon_{22} = 0, \quad \varepsilon_{12} = 1, \quad \varepsilon_{21} = -1,$$

tehát

$$\det \mathbf{A} = A_{11}A_{22} - A_{12}A_{21}.$$

Az $N \times N$-es $\mathbf{A}$ mátrix determinánsa N-ed rendű homogén függvénye a mátrix elemeinek. Jelölésként a

$$\det \mathbf{A} = \begin{vmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ \vdots & & & \\ A_{N1} & \dots & \dots & A_{NN} \end{vmatrix}$$

sémát is használjuk. Ha tehát a mátrix elemeit egyenes vonalak közé fogjuk, akkor a mátrix determinánsát kívánjuk jelölni.

Megjegyezzük, hogy a determináns fenti definíciója nem „logikai szükségszerűség", hanem célszerűségi szempontok következménye. Az így definiált mennyiség segítségével sikerül pl. az N egyenletből álló lineáris egyenletrendszer megoldását tömör formában előállítani (amint ezt a következőkben részletezzük).

11.2.1. A Levi—Civita-szimbólum tulajdonságai

A 11.2-5 definícióból a Levi—Civita-szimbólum következő tulajdonságaihoz juthatunk:

1. Ha két indexet felcserélünk, akkor a szimbólum értéke -1-szeresére változik. Tehát pl.

$$\varepsilon_{klm\ldots s} = -\varepsilon_{lkm\ldots s}. \qquad 11.2\text{-}6$$

Valóban, ha $kl\ldots s$ páros (páratlan) permutációja az $1\ldots N$-nek, akkor $lk\ldots s$ a transzpozíció miatt páratlan (páros) permutációvá válik, és így a szimbólum értéke 1-ről -1-re, illetve -1-ről $+1$-re változik. Amennyiben $kl\ldots s$ nem permutációja $1\ldots N$-nek, akkor $lk\ldots s$ sem az, így 11.2-6 mindkét oldala zérus.

2. Általában fennáll, hogy

$$\varepsilon_{kl\ldots s} = p\varepsilon_{k'l'm'\ldots s'}, \qquad 11.2\text{-}7$$

ahol

$$k'l'm'\ldots s' = P(k, l, m, \ldots, s),$$

tehát a vesszős indexek az eredeti indexek valamilyen permutációját képezik. Ekkor

$$p = \begin{cases} +1, & \text{ha } P \text{ páros permutáció,} \\ -1, & \text{ha } P \text{ páratlan permutáció.} \end{cases}$$

11.3. A determináns néhány tulajdonsága

A determináns néhány olyan tulajdonságát elemezzük, amelynek ismerete szükséges a reciprok mátrix előállításához.

Legyen **A** egy A_{kl} elemekkel rendelkező mátrix. Származtassunk ebből az első két sor cseréjével egy **B** mátrixot. Tehát

$$B_{1k} = A_{2k}, \quad B_{2k} = A_{1k},$$

$$A_{lk} = B_{lk}, \quad \text{ha} \quad l > 2. \qquad 11.3\text{-}1$$

Definíció szerint

$$\det \mathbf{B} = \sum \varepsilon_{kl\ldots s} B_{1k} B_{2l} \ldots B_{Ns}. \qquad 11.3\text{-}2$$

11.3-1 felhasználásával

$$\det \mathbf{B} = \sum \varepsilon_{kl\ldots s} A_{2k} A_{1l} \ldots A_{Ns}. \qquad 11.3\text{-}3$$

11.3-3 helyett 11.2-7 segítségével azt írhatjuk, hogy

$$\det \mathbf{B} = -\sum \varepsilon_{lk\ldots s} A_{1l} A_{2k} \ldots A_{Ns}.$$

Ha most k helyett l-et és l helyett k-t írunk (ez a csere nem jelent semmit, hiszen l és k összegező indexek, amelyek ugyanazokat az értékeket veszik fel), akkor azt kapjuk, hogy

$$\det \mathbf{B} = -\det \mathbf{A}.$$

Vagyis, ha az **A** mátrix első és második sorát felcseréljük, akkor determinánsának értéke előjelet vált. Az eredmény általában érvényes marad tetszőleges két sor cseréje esetén is, tehát

$$\det \mathbf{B} = -\det \mathbf{A}, \qquad 11.3\text{-}4$$

ha **B** az **A**-ból úgy keletkezik, hogy **A** K-adik és L-edik sorát felcseréljük.

11.3-4-ből az is következik, hogy ha egy mátrix két sora megegyezik, akkor determinánsának értéke zérus. Ugyanis ha a két egyforma sort felcseréljük, a determináns előjelet vált, másrészt azonban a két egyforma sor cseréje nem változtatja a determináns értékét. Így tehát

$$\det \mathbf{A} = -\det \mathbf{A},$$

ha **A** két egyforma sorral rendelkezik; következésképpen

$$\det \mathbf{A} = 0. \qquad 11.3\text{-}5$$

Végül permutáljuk egy mátrix sorait, és legyen

$$P(1, 2, \ldots, N) = 1'2' \ldots N'.$$

Ez a permutáció P paritásának megfelelően páros, illetve páratlan számú sorcserével valósítható meg. Minthogy minden sorcsere előjelváltást jelent,

$$\det \mathbf{B} = p \det \mathbf{A}, \qquad 11.3\text{-}6$$

ahol

$$B_{k'l} = A_{kl} \quad (l = 1, 2, \ldots, N),$$

tehát **B** sorait **A** sorainak P szerinti permutációja képezi, és

$$p = \begin{cases} +1, & \text{ha} \quad P \text{ páros} \\ -1, & \text{ha} \quad P \text{ páratlan} \end{cases} \text{permutáció.}$$

A Levi—Civita-szimbólum segítségével 11.3-6 a

$$\det \mathbf{B} = \varepsilon_{1'2'\ldots N'} \det \mathbf{B} \qquad 11.3\text{-}7$$

alakban írható fel, ahol

$$B_{k'l} = A_{kl}. \qquad 11.3\text{-}8$$

Ha $1'2' \ldots N'$ az $1\,2 \ldots N$ permutációja, akkor 11.3-7 11.3-6-nak felel meg. Ha viszont $1'\,2' \ldots N'$ az $1\,2 \ldots N$-nek nem permutációja, akkor $1'2' \ldots N'$-ben az egyik vagy másik érték több mint egyszer szerepel, és ekkor

$$\det \mathbf{B}=0.$$

11.3-7 segítségével általánosíthatjuk a determináns definícióját. 11.3-8 szerint

$$\det \mathbf{B}=\sum \varepsilon_{kl\ldots s}A_{1'k}A_{2'l}\ldots A_{N's}.$$

Szorozzuk meg ezt $\varepsilon_{1'2'\ldots N'}$-vel, és használjuk fel 11.3-7-et. Az

$$\varepsilon_{1'2'\ldots N'}\sum \varepsilon_{kl\ldots s}A_{1'k}A_{2'l}\ldots A_{N's}=\varepsilon^2_{1'2'\ldots N'}\det \mathbf{A}$$

összefüggéshez jutunk.

Összegezzünk a fenti egyenletben az összes $1'2'\ldots N'$ értékekre, és vegyük figyelembe, hogy

$$\sum_{1',2',\ldots,N'} \varepsilon^2_{1'2'\ldots N'}=N!,$$

hiszen $\varepsilon^2_{1'2'\ldots N'}$ $N!$ esetben 1, a többi esetben pedig zérus.

Vezessük be az egyszerűbb

$$1'=K,\quad 2'=L\ldots N'=S$$

jelölést, és osszunk végig $N!$-sal

$$\det \mathbf{A}=\frac{1}{N!}\sum_{K,L,\ldots,S}\sum_{k,l,\ldots,s}\varepsilon_{KL\ldots S}\varepsilon_{kl\ldots s}A_{Kk}\ldots A_{Ss}. \qquad 11.3\text{-}9$$

Ezt a formulát tekinthetnénk a determináns definíciójának is. 11.3-9 az eredeti 11.2-4 definícióval egyenértékű. 11.3-9-nek egyes esetekben hasznát vehetjük.

11.3-9-ből következik pl., hogy

$$\det \tilde{\mathbf{A}}=\frac{1}{N!}\sum_{K,\ldots,S}\sum_{k,\ldots,s}\varepsilon_{KL\ldots S}\varepsilon_{kl\ldots s}\tilde{A}_{Kk}\tilde{A}_{Ll}\ldots\tilde{A}_{Ss}=$$

$$=\frac{1}{N!}\sum_{K,\ldots,S}\sum_{k,\ldots,s}\varepsilon_{KL\ldots S}\varepsilon_{kl\ldots s}A_{kK}\ldots A_{sS}.$$

Ha most a nagybetűs és kisbetűs összegező indexet felcseréljük, azt kapjuk, hogy

$$\det \mathbf{A}=\det \tilde{\mathbf{A}},$$

azaz *a transzponált mátrix determinánsa egyenlő az eredeti mátrix determinánsával.*

A fenti eredményből többek között az is következik, hogy a sorok cseréjére vonatkozó szabályok oszlopcserékre is érvényesek, hiszen ha egy mátrix sorait cseréljük, akkor a transzponált mátrix oszlopai cserélődnek. Tehát

$$\det \mathbf{A}=p\det \mathbf{B}\quad (p=\pm 1),$$

ahol **B** úgy származott **A**-ból, hogy **A** sorait, illetve oszlopait átrendeztük, $p=+1$, ha páros, $p=-1$, ha páratlan permutációnak felel meg az átrendezés.

A fentiek közvetlen következménye az is, hogy *egy mátrix determinánsa zérus, ha két oszlopa azonos.*

11.4. A mátrixszorzat determinánsa

Ebben a fejezetben megmutatjuk, hogy az **A** és **B** N-ed rendű mátrixok szorzatára fennáll a

$$\det(\mathbf{AB})=\det\mathbf{A}\det\mathbf{B} \tag{11.4-1}$$

összefüggés. Legyen $\mathbf{AB}=\mathbf{C}$, ekkor

$$\det(\mathbf{AB})=\sum_{kl\ldots s}\varepsilon_{kl\ldots s}C_{1k}C_{2l}\ldots C_{Ns}. \tag{11.4-2}$$

A mátrixszorzás 8.2-6 definíciója alapján viszont

$$C_{Kk}=(AB)_{Kk}=\sum_{k'=1}A_{Kk'}B_{k'k}. \tag{11.4-3}$$

11.4-3-at 11.4-2-be beírva, és a tényezőket rendezve:

$$\det(\mathbf{AB})=\sum_{k,l,\ldots,s}\sum_{k',l',\ldots,s'}\varepsilon_{kl\ldots s}A_{1k'}A_{2l'}\ldots A_{Ns'}B_{k'k}B_{l'l}\ldots B_{s's}. \tag{11.4-4}$$

Az összegezést más sorrendben is elvégezhetjük, ezért 11.4-4 átírható a

$$\det(\mathbf{AB})=\sum_{k',l',\ldots,s'}A_{1k'}A_{2l'}\ldots A_{Ns'}\left(\sum_{k,l,\ldots,s}\varepsilon_{kl\ldots s}B_{k'k}B_{l'l}\ldots B_{s's}\right)$$

alakra. 11.3-7 szerint azonban a második szummára fennáll a

$$\sum_{k,l,\ldots,s}\varepsilon_{k\ldots s}B_{k'k}B_{l'l}\ldots B_{s's}=\varepsilon_{k'l'\ldots s'}\det\mathbf{B} \tag{11.4-5}$$

egyenlőség. Ennek segítségével 11.4-4 a

$$\det(\mathbf{AB})=\det\mathbf{B}\sum_{k',l',\ldots,s'}\varepsilon_{k'l'\ldots s'}A_{1k'}A_{2l'}\ldots A_{Ns'} \tag{11.4-6}$$

alakot ölti. A 11.4-6 jobb oldalán álló szumma azonban éppen det **A**-val egyenlő, így

$$\det(\mathbf{AB})=\det\mathbf{A}\det\mathbf{B}, \tag{11.4-7}$$

vagyis két mátrix szorzatának determinánsa megegyezik a tényezők determinánsainak szorzatával.

11.5. A reciprok mátrix létezésének feltétele

11.4-7 segítségével fontos megállapításra juthatunk a reciprok mátrixszal kapcsolatban. Legyen ugyanis

$$\mathbf{AB}=\mathbf{E} \quad \text{vagy} \quad \mathbf{BA}=\mathbf{E}. \qquad 11.5\text{-}1$$

Minthogy definíció szerint

$$\det \mathbf{E}=\sum \varepsilon_{kl\ldots s}\delta_{1k}\delta_{2l}\ldots\delta_{Ns}=1,$$

11.4-7 felhasználásával adódik, hogy

$$\det \mathbf{A} \det \mathbf{B}=1, \qquad 11.5\text{-}2$$

tehát det **B** és így **B** is csak akkor létezhet, ha

$$\det \mathbf{A}\neq 0. \qquad 11.5\text{-}3$$

Reciprok mátrixszal tehát csak olyan mátrix rendelkezhet, amelynek determinánsa nem zérus.

A reciprok mátrix jelölésére a továbbiakban

$$\mathbf{B}=\mathbf{A}^{-1}\text{-t}$$

használjuk.

A következőkben megmutatjuk, hogy ha 11.5-3 érvényes, akkor az adott **A**-hoz egy és csak egy reciprok mátrix létezik.

Megjegyzés: 11.5-1-et az

$$\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{E}$$

formában is felírhatjuk, ezért $\mathbf{B}=\mathbf{A}^{-1}$-t *bal és jobb oldali reciproknak* vagy *kétoldali* reciproknak nevezzük.

11.6. Almátrixok

A reciprok mátrix explicit előállításához célszerű bevezetni az almátrix fogalmát.

Rendeljünk hozzá az $N\times N$-es **A** mátrix A_{KL} eleméhez egy szintén $N\times N$-es $\mathbf{A}^{(KL)}$ mátrixot. Az utóbbi mátrixot úgy definiáljuk, hogy az **A** mátrix K-adik sorát és L-edik oszlopát megváltoztatjuk. Ebben a sorban és oszlopban a K, L-edik hely kivételével zérust helyezünk el, a sor és oszlop kereszteződésébe pedig 1-et írunk.

Legyen pl.

$$\mathbf{A}=\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14}\\ A_{21} & A_{22} & A_{23} & A_{24}\\ A_{31} & A_{32} & A_{33} & A_{34}\\ A_{41} & A_{42} & A_{43} & A_{44}\end{bmatrix},$$

ekkor $K=2$, $L=3$ esetén

$$\mathbf{A}^{(23)}=\begin{bmatrix} A_{11} & A_{12} & 0 & A_{14}\\ 0 & 0 & 1 & 0\\ A_{31} & A_{32} & 0 & A_{34}\\ A_{41} & A_{42} & 0 & A_{44}\end{bmatrix}. \qquad 11.6\text{-}1$$

Általában $\mathbf{A}^{(KL)}$ elemeit tehát az

$$A_{kl}^{(KL)}=\begin{cases} A_{kl}, & \text{ha} \quad k\neq K \quad \text{és} \quad l\neq L,\\ 0, & \text{ha} \quad k=K \quad \text{vagy} \quad l=L,\\ 1, & \text{ha} \quad k=K \quad \text{és} \quad l=L\end{cases} \qquad 11.6\text{-}2$$

formula adja.
11.6-2 tömörebben az

$$A_{kl}^{(KL)}=(1-\delta_{kK})(1-\delta_{lL})A_{kl}+\delta_{kK}\delta_{lL} \quad (k, l=1, 2, \ldots, N) \qquad 11.6\text{-}3$$

egyenlőséggel is kifejezhető.

Az $\mathbf{A}^{(KL)}$ almátrix $\det \mathbf{A}^{(KL)}=A^{(KL)}$ determinánsát az eredeti mátrix A_{KL} eleméhez tartozó aldeterminánsának nevezzük.

Az $\mathbf{A}^{(KL)}$ aldetermináns értékét definíció szerint a

$$\det \mathbf{A}^{(KL)}=\sum_{1',\ldots,N'} \varepsilon_{1'2'\ldots K'\ldots N'} A_{11'}^{(KL)} A_{22'}^{(KL)} \ldots A_{KK'}^{(KL)} \ldots A_{NN'}^{(KL)} \qquad 11.6\text{-}4$$

formulából határozhatjuk meg. 11.6-3 szerint

$$A_{KK'}^{(KL)}=\delta_{K'L}. \qquad 11.6\text{-}5$$

11.6-5 felhasználásával 11.6-4 a

$$\det \mathbf{A}^{(KL)}=\sum_{1',\ldots,N'} \varepsilon_{1'2'\ldots K'\ldots N'} A_{11'}^{(KL)} \ldots \delta_{K'L} \ldots A_{NN'}^{(KL)} \qquad 11.6\text{-}6$$

alakot ölti. Ha a 11.6-6 összegben csupán az el nem tűnő tagokat vesszük figyelembe, akkor egyrészt az $1'2' \dots K' \dots N'$ indexsorban az L érték csak egyszer fordulhat elő, másrészt $K' = L$ kell legyen, mert egyébként $\delta_{K'L} = 0$. Ez azonban azt jelenti, hogy az el nem tűnő tagokban csak azok az $A_{kk'}^{(KL)}$ elemek szerepelnek, amelyekre $k \neq K$ és $k' \neq L$. Ezekre viszont 11.6-2 szerint fennáll, hogy

$$A_{kk'}^{(KL)} = A_{kk'}.$$

A 11.6-4 összeg helyett vehetjük tehát a

$$\det \mathbf{A}^{(KL)} = \sum_{1', \dots, N'} \varepsilon_{1'2' \dots N'} A_{11'} A_{22'} \dots \delta_{K'L} \dots A_{NN'} \qquad 11.6\text{-}7$$

összeget is. Az $\mathbf{A}^{(KL)}$ almátrix determinánsa az $\mathbf{A}$ mátrix determinánsából tehát úgy kapható meg, hogy a det $\mathbf{A}$-t definiáló összegben az $A_{KK'}$ elemet $\delta_{LK'}$-vel helyettesítjük.

11.7. A kifejtési tétel

A következőkben megmutatjuk, hogy egy mátrix determinánsát aldeterminánsok lineáris kombinációjaként fejezhetjük ki.

Induljunk ki az

$$A_{KK'} = \sum_{L=1}^{N} A_{KL} \delta_{K'L} \qquad 11.7\text{-}1$$

azonosságból. Szorozzuk meg 11.6-6-ot A_{KL}-lel, és összegezzünk L-re. A

$$\sum_{L} A_{KL} \det \mathbf{A}^{(KL)} = \sum_{L, 1', \dots, N'} \varepsilon_{1'2' \dots N'} A_{KL} A_{11'} \dots A_{NN'} \qquad 11.7\text{-}2$$

formulához jutunk. A jobb oldalon felcserélve a szummázások sorrendjét, 11.7-2 a

$$\sum_{L} A_{KL} \det \mathbf{A}^{(KL)} = \sum_{1', \dots, N'} \varepsilon_{1'2' \dots N'} A_{11'} \dots \left(\sum_{L} A_{KL} \delta_{K'L} \right) \dots A_{NN'} \qquad 11.7\text{-}3$$

alakot ölti. 11.7-1 felhasználásával 11.7-3-ból adódik, hogy

$$\sum_{L} A_{KL} \det \mathbf{A}^{(KL)} = \sum_{1', \dots, N'} \varepsilon_{1' \dots K' \dots N'} A_{11'} \dots A_{KK'} \dots A_{NN'}. \qquad 11.7\text{-}4$$

A determináns definícióját figyelembe véve 11.7-4-ből következik, hogy

$$\det \mathbf{A} = \sum_{L=1}^{N} A_{KL} \det \mathbf{A}^{(KL)} \quad (K = 1, 2, \dots, N). \qquad 11.7\text{-}5$$

A fenti összefüggés N különböző kifejezést ad det **A** kiszámítására. 11.7-5-öt kifejtési tételnek nevezzük, és szavakban a következőképpen fogalmazzuk:

Egy **A** *mátrix determinánsát megkapjuk, ha egy sorának* A_{KL} $(L=1, \ldots, N)$ *elemeit a megfelelő* $\mathbf{A}^{(KL)}$ *almátrixok determinánsaival szorozzuk és a szorzatokat összeadjuk.*

Hasonló eredményre jutunk az $\tilde{\mathbf{A}}$ transzponált mátrix determinánsának kifejtésekor. Minthogy

$$\det \mathbf{A} = \det \tilde{\mathbf{A}},$$

felírhatjuk, hogy

$$\det \mathbf{A} = \sum \varepsilon_{1' \ldots N'} A_{1'1} \ldots A_{N'N}. \qquad 11.7\text{-}6$$

Helyettesítsük 11.7-6-ba az

$$A_{K'K} = \sum_{L=1}^{N} A_{LK} \delta_{K'L}$$

összefüggést. Így

$$\det \mathbf{A} = \sum_{L=1}^{N} A_{LK} \sum \varepsilon_{1' \ldots N'} A_{1'1} \ldots \delta_{K'L} \ldots A_{N'N}, \qquad 11.7\text{-}7$$

és a fentebb kifejtett gondolatmenet segítségével, minthogy

$$\det \tilde{\mathbf{A}}^{(KL)} = \det \mathbf{A}^{(KL)},$$

következik, hogy

$$\det \mathbf{A} = \sum_{L=1}^{N} A_{LK} \det \mathbf{A}^{(LK)}. \qquad 11.7\text{-}8$$

A fenti összefüggés a determináns ún. oszlop szerinti kifejtését adja.

További fontos eredményt kapunk, ha a kifejtési tételt egy olyan **B** mátrixra alkalmazzuk, amelynek K'-edik sorába az **A** mátrix K-adik sorának elemeit, a többi sorba pedig **A** megfelelő elemeit helyezzük.

Tehát

$$B_{kl} = \begin{cases} A_{kl}, & \text{ha} \quad k \neq K', \\ A_{Kl}, & \text{ha} \quad k = K'. \end{cases} \qquad 11.7\text{-}9$$

A **B** mátrix tehát két egyforma sorral rendelkezik, és így

$$\det \mathbf{B} = 0, \qquad 11.7\text{-}10$$

viszont

$$\mathbf{B}^{(K'L)} = \mathbf{A}^{(K'L)} \quad (L = 1, 2, \ldots, N), \qquad 11.7\text{-}11$$

hiszen ha a K'-edik sorok elemeit 0-ra, ill. 1-re változtatjuk, akkor éppen azokat a tagokat tüntetjük el, amelyben **A** és **B** különböznek.

Így

$$\sum_L B_{K'L} \det \mathbf{B}^{(K'L)} = \det \mathbf{B} = 0,$$

viszont

$$B_{K'L} = A_{KL} \quad \text{és} \quad \det \mathbf{B}^{(K'L)} = \det \mathbf{A}^{(K'L)},$$

tehát

$$\sum_L A_{KL} \det \mathbf{A}^{(K'L)} = 0, \quad \text{ha} \quad K' \neq K. \qquad 11.7\text{-}12$$

11.7-12-t és 11.7-8-at együtt a

$$\sum_L A_{KL} \det \mathbf{A}^{(K'L)} = \delta_{KK'} \det \mathbf{A} \qquad 11.7\text{-}13$$

formulával fejezhetjük ki.

Hasonló módon a $\tilde{\mathbf{B}}$ transzponált mátrix kifejtése segítségével megállapíthatjuk, hogy

$$\sum_K A_{KL} \det \mathbf{A}^{(KL')} = \delta_{LL'} \det \mathbf{A}.$$

Összefoglalva tehát, azt állapítottuk meg, hogy: *ha egy mátrix K-adik sorának elemeit a K'-edik sor aldeterminánsaival szorozzuk és a tagokat összegezzük, illetve, ha az L-edik oszlop elemeit az L'-edik oszlop aldeterminánsaival szorozzuk és az így kapott tagokat összegezzük, akkor a $K=K'$, illetve $L=L'$ esetben* **A** *determinánsát, egyéb esetekben pedig zérust kapunk.*

11.8. Az adjungált mátrix

Egy mátrix elemeinek

$$\tilde{\tilde{A}}_{KL} = \det \mathbf{A}^{(KL)} \quad (K, L = 1, 2, \ldots, N)$$

aldeterminánsai egy $N \times N$-es $\tilde{\tilde{\mathbf{A}}}$ mátrixot alkotnak. Ezen mátrix $\bar{\mathbf{A}}$ transzponáltját **A** *adjungált mátrixának nevezzük.* Az $\bar{\mathbf{A}}$ adjungált mátrix elemei:

$$\bar{A}_{KL} = \det \mathbf{A}^{(LK)}.$$

Az adjungált mátrix segítségével a kifejtési tétel különböző formái egyszerűen kifejezhetők. Ugyanis

$$\bar{\mathbf{A}}\mathbf{A} = \mathbf{A}\bar{\mathbf{A}} = \mathbf{E} \det \mathbf{A}. \qquad 11.8\text{-}1$$

Abban az esetben, ha

$$\det \mathbf{A} \neq 0,$$

végigoszthatjuk a 11.8-1 egyenletet det **A**-val. Ekkor az

$$\frac{\bar{\mathbf{A}}}{\det \mathbf{A}}\mathbf{A}=\mathbf{A}\frac{\bar{\mathbf{A}}}{\det \mathbf{A}}=\mathbf{E} \qquad 11.8\text{-}2$$

egyenlőséghez jutunk. 11.8-2-ből látható, hogy a

$$\mathbf{B}=\frac{\bar{\mathbf{A}}}{\det \mathbf{A}} \qquad 11.8\text{-}3$$

mátrix az **A** mátrix bal és jobb oldali reciproka egyszerre. Így, mint azt már 11.1-ben beláttuk, a 11.8-3 az **A** mátrix egyetlen reciprok mátrixának előállítása.

A reciprok mátrix explicit kifejezését aldeterminánsokkal adtuk meg. Félreértések elkerülése végett megjegyezzük, hogy a matematikai irodalomban az aldeterminánsokat többnyire a 11.6-4 formulától eltérő módon definiálják. Mint látni fogjuk, a kétféle definíció között előjelkülönbségek lépnek fel.

A szokásos definíció a következő:

Hagyjuk ki az eredeti **A** mátrixból a K-adik sort és L-edik oszlopot, és a megmaradt elemeket toljuk össze úgy, hogy egy $(N-1)\times(N-1)$-es mátrix keletkezzen.

Az így alkotott almátrixot $\mathbf{A}^{[KL]}$-lel jelöljük. A felső indexeket közrefogó szögletes zárójel jelzi, hogy az almátrixot az utóbbi definíció szerint alkottuk. A 11.6-1-nek megfelelő almátrixot ezzel a definícióval az

$$\mathbf{A}^{[23]}=\begin{pmatrix} A_{11} & A_{12} & A_{14} \\ A_{31} & A_{32} & A_{34} \\ A_{41} & A_{42} & A_{44} \end{pmatrix} \qquad 11.8\text{-}4$$

formula adja.

Aldeterminánson többnyire az így kapott almátrix determinánsát értik.

Belátjuk, hogy

$$\det \mathbf{A}^{[KL]}=(-1)^{K+L}\det \mathbf{A}^{(KL)}, \qquad 11.8\text{-}5$$

azaz a kétféle definíció szerint vett aldetermináns-értékek legfeljebb előjelben térnek el egymástól.

Induljunk ki a

$$\det \mathbf{A}^{(KL)}=\sum_{k,l,\dots,s}\varepsilon_{kl\dots s}A_{1k}^{(KL)}\dots A_{Ns}^{(KL)}$$

formulából. Rendezzük át az almátrixot úgy, hogy a KL-edik eleme $(A_{KL}^{(KL)}=1)$ az $A_{11}^{(KL)}$ elem helyére kerüljön, azaz a K-adik sort hozzuk az első sor, az L-edik oszlopot az első oszlop helyére. Amennyiben a sor- és oszlopcseréket is úgy hajtjuk végre, hogy a megfelelő sort (oszlopot) mindig csak a közvetlen szomszédjaival cseréljük fel, akkor a csere $K+L$ lépésben hajtható végre. A 11.3. pontban kimutattuk, hogy minden csere egy előjelváltozást jelent. Az ily módon kapott új **B** mátrix determinánsára tehát fennáll, hogy

$$\det \mathbf{B} = (-1)^{K+L} \det \mathbf{A}.$$

A determináns definíciója értelmében viszont

$$\det \mathbf{B} = \sum_{k,l,\dots,s} \varepsilon_{kl\dots s} B_{1k} \dots B_{Ns},$$

ahol

$$\mathbf{B} = \begin{pmatrix} 1 & [0] \\ [0] & \mathbf{A}^{[KL]} \end{pmatrix}.$$

A [0] jelölés az első sorban és első oszlopban álló $N-1$ zérust foglalja össze tömörebb formában.

Így

$$\det \mathbf{B} = \det \mathbf{A}^{[KL]}.$$

Ezzel tehát 11.8-5-öt igazoltuk.

A 11.7-5 kifejtési tétel az aldeterminánsok utóbbi értelmezésének felhasználásával a

$$\sum_L (-1)^{K+L} A_{KL} \det \mathbf{A}^{[KL]} = \det \mathbf{A} \qquad 11.8\text{-}6$$

alakot ölti. Az általunk bevezetett definíció alkalmazása esetén a szummában a $(-1)^{K+L}$ tényező nem szerepel.

11.9. A lineáris egyenletrendszerek megoldása

Térjünk vissza a 11.1-ben megkezdett probléma tárgyalására. Az N egyenletből álló, N ismeretlent tartalmazó egyenletrendszert az

$$\mathbf{A}\mathbf{x} = \mathbf{y} \qquad 11.9\text{-}1$$

alakban írhatjuk fel. Amennyiben $\det \mathbf{A} \neq 0$, létezik $\mathbf{A}^{-1} = \mathbf{B}$, és ennek megfelelően 11.9-1 egyetlen megoldását az

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \qquad 11.9\text{-}2$$

szolgáltatja. A 11.1. fejezetben keresett $\mathbf{B}$ mátrix tehát éppen a 11.8-3 mátrix.

A 11.9-2 megoldást részletesebben kiírva, a keresett ismeretleneket az

$$\begin{aligned} x_1 &= A^+_{11} y_1 + A^+_{12} y_2 + \dots + A^+_{1N} y_N \\ &\vdots \\ x_N &= A^+_{N1} y_1 + A^+_{N2} y_2 + \dots + A^+_{NN} y_N \end{aligned} \qquad 11.9\text{-}3$$

formulák határozzák meg. A formulában $\tilde{\tilde{A}}_{KL}$-lel az $\mathbf{A}^{-1}$ mátrix elemeit jelöljük. Almátrixok segítségével a megoldás az

$$x_K=\left(\sum_{L=1}^{N}\tilde{\tilde{A}}_{KL}y_L\right)\frac{1}{\det\mathbf{A}}\quad(K=1,\ldots,N) \tag{11.9-4}$$

alakban írható fel.

11.9-4-et szokás Cramer-féle szabálynak nevezni. A Cramer-szabályt igen gyakran két determináns hányadosaként írjuk fel. Legyen ugyanis

$$\mathbf{Y}^{(K)}=\begin{bmatrix}A_{11}\ldots A_{1K-1} & y_1 & A_{1K+1}\ldots A_{1N}\\ \vdots & & \\ A_{N1}\ldots A_{NK-1} & y_N & A_{NK+1}\ldots A_{NN}\end{bmatrix}. \tag{11.9-5}$$

Az $\mathbf{Y}^{(K)}$ mátrixot tehát úgy kapjuk az $\mathbf{A}$ mátrixból, hogy a K-adik oszlopot az y_L $(L=1,\ldots,N)$ értékekre cseréljük. Az $\mathbf{Y}^{(K)}$ mátrix segítségével a Cramer-szabály az

$$x_K=\frac{\det\mathbf{Y}^{(K)}}{\det\mathbf{A}} \tag{11.9-6}$$

alakban írható fel, mert 11.9-4 zárójeles kifejezése éppen az $\mathbf{Y}^{(K)}$ mátrix determinánsának K-adik oszlop szerinti kifejtése.

Minthogy 11.9-3 szükséges feltételt jelent az $\mathbf{x}$ értékekre, és csak egyetlen $\mathbf{A}^{-1}=\mathbf{B}$ mátrix létezik, 11.9-2 valóban egyetlen megoldása a 11.9-1 egyenletrendszernek.

Abban a speciális esetben, amikor $\mathbf{y}=\mathbf{0}$, és így

$$\mathbf{Ax}=\mathbf{0}, \tag{11.9-7}$$

a $\det\mathbf{A}\neq 0$ feltétel teljesülése esetén az egyenletrendszert csak az

$$\mathbf{x}=\mathbf{0} \tag{11.9-8}$$

triviális megoldás elégíti ki.

11.9-7-et homogén lineáris egyenletrendszernek nevezzük. A nem elfajuló ($\det\mathbf{A}\neq 0$) *homogén lineáris egyenletrendszer csak a triviális* $\mathbf{x}=\mathbf{0}$ *megoldással rendelkezik.*

11.10. Néhány mátrix determinánsának kiszámítása

Az egyenletrendszerek megoldásakor, de egyéb problémák esetén is szükségünk lehet mátrixok determinánsának meghatározására.

Egy N-ed rendű mátrix determinánsa a definíció alapján mindig kiszámítható $N!$ tag összegeként. Ha N nagy szám, akkor az összegezés igen hosszadalmássá válhat. Emiatt hasznos azoknak a matematikai fogásoknak a megkeresése, amelyek segítségével egy determináns értéke egyszerűbben határozható meg.

Hasznos lehet például, ha a determinánst diagonálmátrix vagy háromszögmátrix determinánsának kiszámítására vezethetjük vissza. Ezeknek a mátrixoknak a determinánsát ugyanis a főátlóban álló elemek szorzata határozza meg:

$$\det \mathbf{A} = A_{11}A_{22}\ldots A_{NN}. \tag{11.10-1}$$

A 11.10-1 összefüggést elegendő háromszögmátrixra belátni, hiszen a diagonálmátrix a háromszögmátrix speciális esetének is tekinthető.

Az $\mathbf{A}$ mátrix determinánsa definíció szerint

$$\det \mathbf{A} = \sum_{1', 2', \ldots, N'} \varepsilon_{1'2'\ldots N'} A_{11'}\ldots A_{NN'}. \tag{11.10-2}$$

Ha $\mathbf{A}$-nak pl. a főátlója felett csupa zéruselem áll, akkor

$$A_{kk'} = 0, \quad \text{ha} \quad k' > k.$$

A 11.10-2 formulában a szumma minden tagjára fennáll, hogy

$$1' + 2' + \ldots + N' = 1 + 2 + \ldots + N,$$

azaz

$$(1 - 1') + (2 - 2') + \ldots + (N - N') = 0.$$

Amennyiben 11.10-2 egyes tagjaiban csupa nem zérus szorzótényezőt akarunk választani, akkor az egyes zárójelekben $k' > k$ miatt nem állhat negatív szám. Az egyenlőség tehát csak

$$1 = 1', 2 = 2', \ldots, N = N'$$

esetén teljesülhet.

Ezzel éppen 11.10-1-et bizonyítottuk.

A mátrixok diagonális vagy háromszögformára hozásához többnyire azt használjuk fel, hogy egy mátrix determinánsának értéke nem változik, ha a mátrix egy sorához hozzáadjuk egy másik sorának többszörösét.

Legyen $\mathbf{A}$ egy $N \times N$-es mátrix. Képezzük azt a $\mathbf{B}$ mátrixot, amelynek első sorát úgy kapjuk, hogy $\mathbf{A}$ első sorához hozzáadjuk $\mathbf{A}$ második sorának λ-szorosát. Tehát

$$\left.\begin{aligned} B_{1j} &= A_{1j} + \lambda A_{2j}, \\ B_{kj} &= A_{kj}, \qquad \text{ha} \quad k \geqq 2 \end{aligned}\right\}.$$

Definíció szerint

$$\det \mathbf{B} = \sum_{k, l, m, \ldots, s} \varepsilon_{klm\ldots s}(A_{1k} + \lambda A_{2k})A_{2l}\ldots A_{Ns}. \tag{11.10-3}$$

Az összegezés tulajdonságainak alapján 11.10-3 jobb oldala szétbontható.

$$\det \mathbf{B} = \sum_{k,l,m,\ldots,s} \varepsilon_{klm\ldots s} A_{1k} A_{2l} \ldots A_{Ns} + \lambda \sum_{k,l,m,\ldots,s} \varepsilon_{kl\ldots s} A_{2k} A_{2l} \ldots A_{Ns} \qquad 11.10\text{-}4$$

11.10-4 első tagja éppen det **A**, a második tag egy olyan mátrix determinánsát adja, amelynek első és második sora megegyezik. Ez utóbbi érték tehát zérus. Tehát

$$\det \mathbf{A} = \det \mathbf{B}. \qquad 11.10\text{-}5$$

Tételünket speciálisan az első és második sor esetére bizonyítottuk be. Teljesen azonos módon végezhető el azonban a bizonyítás a K-adik és L-edik sorra. Ugyancsak érvényes a tétel a determináns oszlopaira is.

11.11. Magasabb rendű almátrixok

A következőkben általánosítjuk a 11.6. fejezetben bevezetett almátrixfogalmat, és ennek megfelelően általánosítjuk a kifejtési tételt is.

11.11.1. Másodrendű almátrixok és aldeterminánsok

Képezzük az $\mathbf{A}^{(K_1L_1)}$ almátrix K_2L_2 eleméhez tartozó almátrixot, és az ily módon keletkezett mátrix jelölésére vezessük be az

$$\left(\mathbf{A}^{(K_1L_1)}\right)^{(K_2L_2)} = \mathbf{A}^{(K_1K_2,\, L_1L_2)} \qquad 11.11\text{-}1$$

szimbólumot. 11.11-1-et az **A** mátrix másodrendű almátrixának nevezzük.

Példaként felírjuk a negyedrendű **A** mátrix $\mathbf{A}^{(12,23)}$ almátrixát:

$$\mathbf{A}^{(12,23)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ A_{31} & 0 & 0 & A_{34} \\ A_{41} & 0 & 0 & A_{44} \end{pmatrix}. \qquad 11.11\text{-}2$$

A másodrendű almátrixokat leggyakrabban a $K_1 \neq K_2$ és $L_1 \neq L_2$ esetben használjuk. Ekkor az almátrixképzés két lépése különböző sorokat és oszlopokat érint, így a végeredmény nem függ a sorok, ill. oszlopok megváltoztatásának sorrendjétől, tehát

$$\left(\mathbf{A}^{(K_1L_1)}\right)^{(K_2L_2)} = \left(\mathbf{A}^{(K_2L_2)}\right)^{(K_1L_1)}, \quad \text{ha} \quad \begin{matrix} K_1 \neq K_2, \\ L_1 \neq L_2; \end{matrix} \qquad 11.11\text{-}3$$

vagy a 11.1-1 jelöléssel

$$\mathbf{A}^{(K_1K_2,\,L_1L_2)}=\mathbf{A}^{(K_2K_1,\,L_2L_1)}.$$

Az alkalmazások során azonban más esetek is előfordulhatnak. Az almátrix-képzés sorrendjét is megadó 11.11-1 definíció természetesen ezekben az esetekben is egyértelmű végeredményre vezet.

A továbbiakban a másodrendű almátrix determinánsának (másodrendű aldetermináns) kiszámításával foglalkozunk.

Először egyszerűsítő jelölésként bevezetjük a

$$\mathbf{K}=K_1,\,K_2 \quad \text{és} \quad \mathbf{L}=L_1,\,L_2 \qquad 11.11\text{-}4$$

kétkomponensű mennyiségeket. A 11.11-3 által definiált almátrixot tehát az

$$\mathbf{A}^{(K_1K_2,\,L_1L_2)}=\mathbf{A}^{(\mathbf{K},\,\mathbf{L})} \qquad 11.11\text{-}5$$

szimbólummal jelölhetjük.

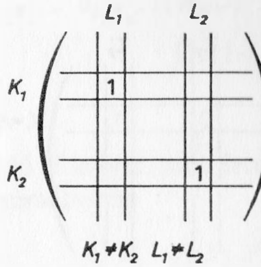
A determináns definíciója szerint

$$\det \mathbf{A}^{(\mathbf{K},\,\mathbf{L})}=\sum_{1',\,2',\,\dots\,N'} \varepsilon_{1'2'\dots N'}A_{11'}^{(\mathbf{KL})}A_{22'}^{(\mathbf{KL})}\dots A_{NN'}^{(\mathbf{K},\,\mathbf{L})}. \qquad 11.11\text{-}6$$

A $K_1\neq K_2$, $L_1\neq L_2$ esetben érvényesek az

$$A_{K_1K_1'}^{(\mathbf{K},\,\mathbf{L})}=\delta_{L_1K'_1},$$
$$A_{K_2K_2'}^{(\mathbf{K},\,\mathbf{L})}=\delta_{L_2K'_2} \qquad 11.11\text{-}7$$

összefüggések *(11.1. ábra)*.



11.1. ábra

Az ábrákon a mátrix jellegzetes sorait és oszlopait tüntettük fel. Azokban a sorokban és oszlopokban, ahol nem tüntettünk fel számértéket, mindenütt zérus áll. A jelöletlen sorok és oszlopok elemei megegyeznek az eredeti mátrix megfelelő elemeivel.

Minthogy azonban

$$A_{kk'}^{(\mathbf{K},\,\mathbf{L})}=A_{kk'}, \quad \text{ha} \quad k\neq K_1,\,K_2 \text{ és } k'\neq L_1,\,L_2, \qquad 11.11\text{-}8$$

a 11.11-6 összeget a

$$\det \mathbf{A}^{(\mathbf{KL})}=\sum_{1',\,2',\,\dots,\,N'} \varepsilon_{1'2'\dots N'}A_{11'}\dots\delta_{L_1K_1}\dots\delta_{L_2K_2}\dots A_{NN'} \qquad 11.11\text{-}9$$

formában is felírhatjuk. Az $\mathbf{A}^{(\mathbf{K},\,\mathbf{L})}$ almátrix determinánsát tehát úgy kapjuk meg az eredeti mátrix determinánsából, hogy az összegezendő tagokban a K_1-edik és K_2-edik tényezők esetén elvégezzük az

$$A_{K_1 K_1'} = \delta_{L_1 K_1'}$$

és

$$A_{K_2 K_2'} = \delta_{L_2 K_2'}$$

helyettesitést.

11.11-9-ből következik, hogy

$$\det \mathbf{A}^{(K_1K_2,\,L_1L_2)} = -\det \mathbf{A}^{(K_1K_2,\,L_2L_1)} = -\det \mathbf{A}^{(K_2K_1,\,L_1L_2)}. \qquad 11.11\text{-}10$$

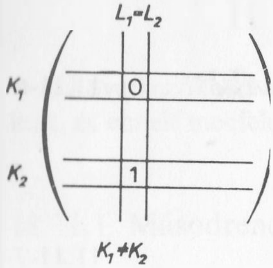
Megállapitottuk, hogy a 11.11-9 összeg a $K_1 \neq K_2$, $L_1 \neq L_2$ esetén azonosan egyenlő a 11.11-6 determinánst definiáló összeggel.

11.11-9 azonban értelmes definiciót jelent a $K_1 \neq K_2$, de $L_1 = L_2$ esetben is. Ekkor azonban

$$\sum_{1',\dots,N'} \varepsilon_{1'2'\dots N'} A_{11'} \dots \delta_{L_1 K_1'} \dots \delta_{L_2 K_2'} \dots A_{NN'} = 0. \qquad 11.11\text{-}11$$

Másrészt tudjuk, hogy

$$\det \mathbf{A}^{(K_1K_2,\,L_1L_1)} = 0,$$

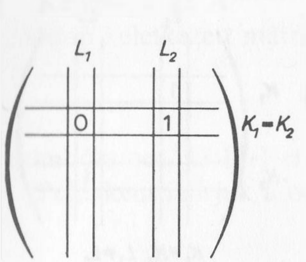


11.2. ábra

azaz a 11.11-9 összeg det $\mathbf{A}^{(KL)}$ értékét a $K_1 \neq K_2$ esetén tetszőleges L_1, L_2-re helyesen adja *(11.2. ábra)*.

Foglalkoznunk kell még a $K_1 = K_2$ esettel. A 11.11-9 átírás ekkor értelmetlenné válik; az eredeti 11.11-1 almátrix vizsgálatából azonban adódik, hogy

$$\det \mathbf{A}^{(KK,\,L_1L_2)} = 0, \quad \text{ha} \quad L_1 \neq L_2,$$



11.3. ábra

mert ebben az esetben az L_1-edik oszlop csupa zéruselemet tartalmaz *(11.3. ábra)*. A $K_1 = K_2$, $L_1 = L_2$ esetben pedig az eredeti elsőrendű almátrixhoz jutunk, tehát

$$\det \mathbf{A}^{(KL)} = \det \mathbf{A}^{(KL)}.$$

11.11.2. Magasabb rendű almátrixok

A másodrendű almátrixokból egy sor és egy oszlop megfelelő változtatásával harmadrendű almátrixhoz juthatunk. Tehát

$$\left(\mathbf{A}^{(K_1K_2,\,L_1L_2)}\right)^{(K_3L_3)} = \mathbf{A}^{(K_1K_2K_3,\,L_1L_2L_3)}. \qquad 11.11\text{-}12$$

Az eljárást folytatjuk, s így minden n-edik lépés után az eredeti mátrix n-ed rendű almátrixaihoz jutunk. Ezeket az almátrixokat a 11.11-12 jelölés alapján az

$$\mathbf{A}^{(K_1 K_2 \ldots K_n,\, L_1 \ldots L_n)}$$

szimbólummal jelöljük. A 11.11-12 almátrix az eredeti mátrixból definíció szerint úgy képezhető, hogy először az $\mathbf{A}$ mátrix K_1-edik sorát és L_1-edik oszlopát cseréljük fel az almátrixképzés 11.6-2 szabályainak megfelelően, majd az így kapott mátrix K_2-edik sorát és L_2-edik oszlopát változtatjuk meg stb. A 11.11-12 jelölés tehát egyértelműen rögzíti a sorok, ill. oszlopok cseréjének sorrendjét. Megjegyezzük, hogy amennyiben egy N-ed rendű mátrix N-ed rendű almátrixát képezzük úgy, hogy

$$K_1, K_2 \ldots K_N = 1, 2 \ldots N$$

és

$$L_1, L_2 \ldots L_N = 1, 2 \ldots N, \qquad 11.11\text{-}13$$

akkor az $\mathbf{E}$ egységmátrixhoz jutunk, tehát

$$\mathbf{A}^{(1\,2 \ldots N,\, 1\,2 \ldots N)} = \mathbf{E}. \qquad 11.11\text{-}14$$

Egy n-ed rendű almátrix determinánsát a 11.11-9 formula általánosításával határozhatjuk meg. A 11.11-9 formulához vezető meggondolás ismétlésével adódik, hogy

$$\det \mathbf{A}^{(K_1 \ldots K_n,\, L_1 \ldots L_n)} = \sum_{1', \ldots, N'} \varepsilon_{1'2' \ldots N'} A_{11'} \ldots \delta_{L_1 K_1'} \ldots \delta_{L_2 K_2'} \ldots \delta_{L_n K_n'} \ldots A_{NN'}, \qquad 11.11\text{-}15$$

ha

$$K_1 \neq K_2 \neq \ldots K_N.$$

Másképpen fogalmazva ez azt jelenti, hogy az n-ed rendű aldetermináns értékét meghatározó összeg az eredeti determinánst adó szummából az

$$A_{K_\nu K_\nu'} \rightarrow \delta_{L_\nu K_\nu'} \quad (\nu = 1, 2, \ldots, n) \qquad 11.11\text{-}16$$

helyettesítés elvégzésével adódik, ha a K_ν sorindexek mind különbözőek.

Az almátrixképzés egyszerűbb jelölésére célszerű a

$$\mathbf{K}(n) = K_1 \ldots K_n$$

és

$$\mathbf{L}(n) = L_1 \ldots L_n \qquad 11.11\text{-}17$$

n komponensű mennyiségek bevezetése, tehát

$$\mathbf{A}^{(K_1 \ldots K_n,\, L_1 \ldots L_n)} = \mathbf{A}^{(\mathbf{K}(n),\, \mathbf{L}(n))}.$$

A továbbiakban a $\mathbf{K}(n)$ és $\mathbf{L}(n)$ n komponensű mennyiségeket az almátrixképzésben nem szereplő indexek közül vett további tetszőleges indexekkel N komponensű mennyiségekké egészítjük ki,

$$\mathbf{K}=K_1\ldots K_nK_{n+1}\ldots K_N,$$

$$\mathbf{L}=L_1\ldots L_nL_{n+1}\ldots L_N \qquad 11.11\text{-}18$$

N komponensű mennyiségek első n eleme tehát megegyezik $\mathbf{K}(n)$, ill. $\mathbf{L}(n)$ komponenseivel. A fenti jelölések segítségével az n-ed rendű almátrixok determinánsa egyszerűen megadható. Amennyiben ugyanis $\varepsilon_{\mathbf{K}}\neq 0$, akkor

$$\det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}=\varepsilon_{\mathbf{K}}\sum_{\mathbf{K}'}\varepsilon_{\mathbf{K}'}\delta_{L_1K_1'}\ldots\delta_{L_nK_n'}A_{K_{n+1}K_{n+1}'}\ldots A_{K_NK_N'}$$

11.11-19

alakban írható fel, hiszen a 11.11-19-ben szereplő szumma egy, az eredeti almátrix sorainak megfelelő cseréjével származtatható mátrix determinánsa, így csak egy előjelben különbözhet az eredeti aldeterminánstól. Ezt az előjelet azonban $\varepsilon_{\mathbf{K}}$ faktorral kompenzáltuk.

11.11.3. A kifejtési tétel általánosítása

Szorozzuk meg 11.11-19-et $A_{K_nL_n}$-nel, és összegezzünk L_n-re. Azt kapjuk, hogy

$$\sum_{L_n}A_{K_nL_n}\det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}=\sum_{L_n}A_{K_nL_n}\varepsilon_{\mathbf{K}}\sum_{\mathbf{K}'}\varepsilon_{\mathbf{K}'}\delta_{L_1K_1'}\ldots\delta_{L_nK_n'}A_{K_{n+1}K_{n+1}'}\ldots A_{K_nK_n'}.$$

11.11-20

11.11-20-ban az összegezések sorrendjét felcserélve és figyelembe véve a

$$\sum_{L_n=1}^{N}\delta_{L_nK_n'}A_{K_nL_n}=A_{K_nK_n'} \qquad 11.11\text{-}21$$

egyenlőséget, a

$$\sum_{L_n=1}^{N}A_{K_nL_n}\det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}=\varepsilon_{\mathbf{K}}\sum_{\mathbf{K}'}\varepsilon_{\mathbf{K}'}\delta_{L_1K_1'}\ldots\delta_{L_{n-1}K_{n-1}'}A_{K_nK_n'}A_{K_{n+1}K_{n+1}'}\ldots A_{K_NK_N'}.$$

11.11-22

A 11.11-22 jobb oldalán álló összeg viszont egy $n-1$-ed rendű almátrix determinánsa, tehát

$$\sum_{L_n=1}^{N}A_{K_nL_n}\det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}=\varepsilon_{\mathbf{K}}\det \mathbf{A}^{(\mathbf{K}(n-1),\,\mathbf{L}(n-1))}, \qquad 11.11\text{-}23$$

ahol $\mathbf{K}(n)$ és $\mathbf{K}(n-1)$, valamint $\mathbf{L}(n)$ és $\mathbf{L}(n-1)$ első n komponense megegyezik.

A 11.11-23 egyenlet tulajdonképpen azt a magától értetődő tényt fejezi ki, hogy az $n-1$-ed rendű aldeterminánsok n-ed rendű aldeterminánsokból előállíthatók.

A fenti eljárást folytatva, 11.11-19-et $A_{K_n L_n} A_{K_{n-1} L_{n-1}}$-gyel szorozva és L_n-re, valamint L_{n-1}-re összegezve és a megfelelő átalakításokat elvégezve, a

$$\sum_{L_n, L_{n-1}}^{N} A_{K_{n-1} L_{n-1}} A_{K_n L_n} \det \mathbf{A}^{(\mathbf{K}(n),\, \mathbf{L}(n))} = \varepsilon_{\mathbf{K}} \det \mathbf{A}^{(\mathbf{K}(n-2),\, \mathbf{L}(n-2))} \qquad 11.11\text{-}24$$

összefüggéshez jutunk. Hasonló módon 11.11-19-et az $A_{K_1 L_1} A_{K_2 L_2} \ldots A_{K_n L_n}$ szorzattal beszorozva és minden L_i $(i=1, \ldots, n)$-re összegezve, a

$$\sum_{L_1, \ldots, L_n = 1}^{N} A_{K_1 L_1} A_{K_2 L_2} \ldots A_{K_n L_n} \det \mathbf{A}^{(\mathbf{K}(n),\, \mathbf{L}(n))} = \varepsilon_{\mathbf{K}} \det \mathbf{A} \qquad 11.11\text{-}25$$

formulát kapjuk. A 11.11-25 formula megadja, hogy egy **A** mátrix determinánsa hogyan adható meg n-ed rendű aldeterminánsai segítségével, így a kifejtési tétel általánosításának tekinthető. Végül még egyszer felhívjuk a figyelmet arra, hogy a fentiek csak az $\varepsilon_{\mathbf{K}} \neq 0$ esetben érvényesek.

11.11.4. Kiegészítő almátrixok

A kifejtési tétel 11.11-25 általánosítását egyszerűbben is megfogalmazhatjuk az ún. kiegészítő almátrixok segítségével. A kiegészítő almátrix fogalmának bevezetéséhez induljunk ki a csupa különböző elemből álló

$$\mathbf{K} = K_1 \ldots K_N,$$

$$\mathbf{L} = L_1 \ldots L_N \qquad 11.11\text{-}26$$

indexsorozatokból. Teljesüljenek tehát az

$$\varepsilon_{\mathbf{K}} \neq 0 \quad \text{és} \quad \varepsilon_{\mathbf{L}} \neq 0$$

feltételek. Konstruáljuk az $N \times N$-es **A** mátrix $\mathbf{K}(n)$ és $\mathbf{L}(n)$-nel jellemzett $\mathbf{A}^{(\mathbf{K}(n),\, \mathbf{L}(n))}$ almátrixát. Az almátrixképzésben a korábbi definícióknak megfelelően a

$$\mathbf{K}(n) = K_1 \ldots K_n,$$

$$\mathbf{L}(n) = L_1 \ldots L_n$$

indexek vesznek részt. Jelöljük a **K** és **L** mennyiségeknek az almátrixképzésben nem szereplő komponenseit az alábbi módon:

$$\mathbf{K}(-n)=K_{n+1}\ldots K_N,$$

$$\mathbf{L}(-n)=L_{n+1}\ldots L_N, \qquad 11.11\text{-}27$$

és képezzük az

$$\mathbf{A}^{(\mathbf{K}(-n),\ \mathbf{L}(-n))}=\mathbf{A}^{(K_{n+1}\ldots K_N,\ L_{n+1}\ldots L_N)} \qquad 11.11\text{-}28$$

almátrixot. A 11.11-28 almátrixot az $\mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}$ almátrix kiegészítő almátrixának nevezzük. Mivel a $\mathbf{K}(n)$, $\mathbf{L}(n)$ indexsorozatok az $\varepsilon_{\mathbf{K}}\neq 0$ feltételek teljesülése mellett is sokféle módon egészíthetők ki N komponensű mennyiségekké, adott $\mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}$ almátrixhoz a fenti definíció alapján több kiegészítő mátrixot rendelhetünk.

A többértelműség megszüntetésére a továbbiakban rendezett indexsorokat is felhasználunk. Általában egy

$$\mathbf{M}=M_1\ldots M_N$$

indexsoron belül az indexek két csoportját nagyság szerint rendezzük. Tehát az $\varepsilon_{\mathbf{M}}\neq 0$ feltétel mellett az

$$M_1<M_2<\ldots<M_n$$

és $\qquad 11.11\text{-}29$

$$M_{n+1}<M_{n+2}<\ldots<M_N$$

egyenlőtlenségek teljesülését is megköveteljük.

A következőkben a kiegészítő almátrix determinánsának meghatározásával foglalkozunk, majd a 11.11-25 kifejtési tételt írjuk fel aldeterminánsok és kiegészítő aldeterminánsaik segítségével.

A 11.11-19 formulához hasonlóan az $\mathbf{A}^{(\mathbf{K}(-n),\ \mathbf{L}(-n))}$ kiegészítő almátrix determinánsa a

$$\det \mathbf{A}^{(\mathbf{K}(-n),\ \mathbf{L}(-n))}=\varepsilon_{\mathbf{K}}\sum_{\mathbf{K}'}\varepsilon_{K'}A_{K_1K_1'}\ldots A_{K_nK_n'}\delta_{L_{n+1}K_{n+1}'}\ldots\delta_{L_NK_N'} \qquad 11.11\text{-}30$$

alakban írható fel.

Alakítsuk át a

$$\sum_{L(n)}^{N}A_{K_1L_1}A_{K_2L_2}\ldots A_{K_nL_n}\det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}=\varepsilon_{\mathbf{K}}\det \mathbf{A} \qquad 11.11\text{-}31$$

kifejtési tételt. Az összegezést hajtsuk végre két lépésben. Induljunk ki egy rendezett **M** indexsorozatból, amelyre $\mathbf{M}(n)=\mathbf{L}(n)$, és összegezzünk először az $L_1\ldots L_n$ értékek permutációira. Ezután futtassuk végig **M**-et az összes lehetséges rendezett indexsorozatokon. Formálisan egészítsük ki még a 11.11-31 összeg tagjait a $\delta_{L_{n+1}M_{n+1}}\ldots\delta_{L_NM_N}$ tényezőkkel. 11.11-31 ezzel a

$$\sum_{\mathbf{M}} \sum_{\mathbf{L}(n)} A_{K_1L_1}A_{K_2L_2}\ldots A_{K_nL_n}\delta_{L_{n+1}M_{n+1}}\ldots\delta_{L_NM_N} \det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))} = \varepsilon_{\mathbf{K}} \det \mathbf{A}$$

11.11-32

alakot ölti, hiszen az új szorzótényezők az összegen nem változtatnak.

A 11.11-32 összeg akkor sem változik, ha minden tagját $\varepsilon_{\mathrm{L}}^2$-tel megszorozzuk. Tehát 11.11-32 az

$$\sum_{\mathbf{M}} \sum_{\mathbf{L}(n)} A_{K_1L_1}A_{K_2L_2}\ldots A_{K_nL_n}\delta_{L_{n+1}M_{n+1}}\delta_{L_nM_n}\varepsilon_{\mathrm{L}}^2 \det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))} = \varepsilon_{\mathbf{K}} \det \mathbf{A}$$

11.11-33

alakban is felírható.

Látható, hogy

$$\varepsilon_{\mathrm{L}} \det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))} = \varepsilon_{\mathrm{M}} \det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}. \tag{11.11-34}$$

Ezt a tényezőt kiemelve a szumma elé, és a kiegészítő aldeterminánsra vonatkozó 11.11-30 formulát felhasználva, adódik, hogy

$$\sum_{\mathbf{L}(n)} \varepsilon_{\mathrm{L}} A_{K_1L_1}\ldots A_{K_nL_n}\delta_{L_{n+1}M_{n+1}}\ldots\delta_{L_NM_N} = \varepsilon_{\mathbf{K}} \det \mathbf{A}^{(\mathbf{K}(-n),\,\mathbf{L}(-n))}.$$

11.11-35

11.11-34 és 11.11-35 felhasználásával a kifejtési tétel

$$\sum_{\mathbf{M}} \varepsilon_{\mathrm{M}} \det \mathbf{A}^{(\mathbf{K}(-n),\,\mathbf{L}(-n))} \det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))} = \det \mathbf{A} \quad (\varepsilon_{\mathbf{K}} \neq 0) \tag{11.11-36}$$

alakjához jutunk. 11.11-36-ot szokás Laplace-tételnek is nevezni. A tétel azt fejezi ki, hogy az **A** mátrix n sorát kiválasztva, és képezve e sorok segítségével **A** összes, az **M** rendezettségének megfelelő $n \times n$-es almátrixait és ezek kiegészítő almátrixait, det **A** előállítható az összetartozó aldeterminánsok szorzatainak összegeként.

Ellenőrzésként érdemes meggondolni, hogy 11.11-36 az $n=1$ esetben speciális esetként visszaadja a kifejtési tétel 11.7-6 alakját. 11.11-36-ban ez esetben

$$\mathbf{K}(1) = K, \quad \mathbf{M}(1) = M,$$

tehát

$$\det \mathbf{A}^{(\mathbf{K}(1),\,\mathbf{L}(1))} = \det \mathbf{A}^{(K,\,M)}.$$

Mivel **M** rendezett indexsorozat, $\mathbf{A}^{(\mathbf{K}(-1),\,\mathbf{L}(-1))}$ almátrix egyetlen megmaradt A_{KM} elemének előjele ε_{M}-szeresére változik, tehát

$$\det \mathbf{A}^{(\mathbf{K}(-1),\,\mathbf{L}(-1))} = \varepsilon_{\mathrm{M}} A_{KM}.$$

Ily módon

$$\varepsilon_{\mathrm{M}}^2 \sum_{M} A_{KM} \det \mathbf{A}^{(KM)} = \det \mathbf{A},$$

ami $\varepsilon_{\mathrm{M}}^2 = 1$ miatt megegyezik a kifejtési tétel 11.7-6 alakjával.

Egy mátrix rangját a mátrix aldeterminánsai segítségével definiáljuk. Az A mátrix rangját

$$\mathcal{R}(\mathbf{A})$$

-val jelöljük. *A mátrix rangja megegyezik a legmagasabb rendű el nem tűnő aldetermináns rendjével.*

Tehát az $N \times N$-es **A** mátrix rangja

$$\mathcal{R}(\mathbf{A}) = N, \qquad 11.11\text{-}37$$

ha

$$\det \mathbf{A} \neq 0;$$

$$\mathcal{R}(\mathbf{A}) < N,$$

ha $\det \mathbf{A} = 0$. Amennyiben $\det \mathbf{A} = 0$, de $\bar{\mathbf{A}} \neq 0$, akkor

$$\mathcal{R}(\mathbf{A}) = N - 1,$$

hiszen az **A** mátrix elsőrendű aldeterminánsai nem lehetnek mind zérusértékűek.

Általában

$$\mathcal{R}(\mathbf{A}) = N - n,$$

ha létezik legalább egy olyan

$$\left.\begin{aligned} \mathbf{K}(n) &= K_1 \ldots K_n, \\ \mathbf{L}(n) &= L_1 \ldots L_n \end{aligned}\right\} \qquad 11.11\text{-}38$$

indexcsoport, hogy

$$\det \mathbf{A}^{(\mathbf{K}(n),\, \mathbf{L}(n))} \neq 0,$$

azonban

$$\det \mathbf{A}^{(\mathbf{K}'(\nu),\, \mathbf{L}'(\nu))}{}_{\nu} = 0$$

tetszőleges $\mathbf{K}'(\nu)$, $\mathbf{L}'(\nu)$ indexcsoportokra, ha $\nu < n$.

11.12. Az elfajult homogén lineáris egyenletrendszer

A 11.9. fejezetben megállapítottuk, hogy az

$$\mathbf{A}\mathbf{x} = \mathbf{0}$$

homogén lineáris egyenletrendszernek a

$$\det \mathbf{A} \neq 0 \qquad 11.12\text{-}1$$

esetben csak az $\mathbf{x} = \mathbf{0}$ triviális megoldása létezik.

A homogén lineáris egyenletrendszert elfajultnak nevezzük, ha

$$\det \mathbf{A}=0. \tag{11.12-2}$$

A következőkben megmutatjuk, hogy 11.12-2 fennállása esetén a 11.12-1 egyenletrendszernek a triviálistól különböző megoldásai is vannak. A nem triviális megoldások meghatározásakor azonban figyelembe kell vennünk az egyenletrendszer elfajultságának mértékét is.

11.12.1. Az első rendben elfajult homogén lineáris egyenletrendszer

Az elfajultság legegyszerűbb esete az, amikor

$$\mathbf{Ax}=\mathbf{0}, \quad \det \mathbf{A}=0, \quad \text{de} \quad \bar{\mathbf{A}}\neq\mathbf{0}, \tag{11.12-3}$$

vagyis az adjungált mátrix nem azonosan zérusmátrix. Ezt az egyenletrendszert egyszeresen elfajultnak nevezzük.

Mivel az adjungált mátrix nem azonosan zérus, létezik legalább egy olyan KL indexpár, amelyre

$$\bar{A}_{KL}=\det \mathbf{A}^{(LK)}\neq 0.$$

Megállapítottuk, hogy $\det \mathbf{A}=0$ esetén

$$\mathbf{A}\bar{\mathbf{A}}=\mathbf{0}. \tag{11.12-4}$$

11.12-4 komponensekben történő kifejtéséből adódik, hogy

$$(\mathbf{A}\bar{\mathbf{A}})_{kK}=\sum_l A_{kl}\det \mathbf{A}^{(kl)}=0. \tag{11.12-5}$$

A 11.12-5 formulát összehasonlítva a 11.12-3 egyenletrendszer

$$\sum_l A_{kl}x_l=0$$

alakú k-adik sorával, adódik, hogy

$$x_l=\alpha \det \mathbf{A}^{(KL)} \quad (l=1, \ldots, N) \tag{11.12-6}$$

$\alpha\neq 0$ esetén a 11.12-3 egyenletrendszer nem triviális megoldása. 11.12-6-ban $\alpha\neq 0$-t tetszőlegesen választhatjuk meg. α megfelelő választásával elérhetjük például, figyelembe véve 11.12-3-at, hogy

$$x_L=a \tag{11.12-7}$$

legyen, ahol $a\neq 0$ egy előzetesen megadott érték. Az egyenletrendszer 11.12-7-nek eleget tevő megoldásrendszerét az

$$x_l = x_l^{(1)} = a \frac{\det \mathbf{A}^{(Kl)}}{\det \mathbf{A}^{(KL)}} \quad (l=1, \dots, N) \qquad 11.12\text{-}8$$

összefüggés szolgáltatja. A 11.12-8 megoldást behelyettesítve az eredeti egyenletrendszerbe, meggyőződhetünk arról, hogy

$$\mathbf{A}\mathbf{x}^{(1)} = \mathbf{0} \quad \text{és} \quad x_L^{(1)} = a.$$

Beláttuk tehát, hogy az egyszeresen elfajult homogén lineáris egyenletrendszer rendelkezik nem triviális megoldással, s hogy a megoldásban egy megfelelő ismeretlen értéke szabadon választható.

11.12.2. A kétszeresen elfajult homogén lineáris egyenletrendszer

Kétszeresen elfajulónak nevezzük a homogén lineáris egyenletrendszert, ha

$$\mathbf{A}\mathbf{x} = \mathbf{0} \qquad 11.12\text{-}9$$

és

$$\det \mathbf{A} = 0, \quad \bar{\mathbf{A}} = \mathbf{0}, \qquad 11.12\text{-}10$$

de létezik két olyan K_1L_1, K_2L_2, $K_1 \neq K_2$ és $L_1 \neq L_2$ indexpár, amelyekre teljesül a

$$\det \mathbf{A}^{(K_1K_2,\, L_1L_2)} \neq 0 \qquad 11.12\text{-}11$$

feltétel. A 11.12-10 és 11.12-11 feltétel azt jelenti, hogy az egyenletrendszer **A** együtthatómátrixának rangja $N-2$, tehát **A** determinánsa, valamint összes elsőrendű aldeterminánsa eltűnő, de van legalább egy nem zérusértékű másodrendű aldeterminánsa.

Belátjuk, hogy 11.12-9 rendelkezik nem triviális megoldással. Határozzuk meg az $\mathbf{A}^{(K_1L_1)}$ almátrix determinánsát a kifejtési tétel segítségével (11.11-36 formula). Ha $K_1 \neq K_2$, akkor

$$\sum_{l=1}^{N} A_{kl} \det \mathbf{A}^{(K_1K_2,\, L_2l)} = \delta_{kK_2} \det \mathbf{A}^{(K_1L_1)} \quad (k \neq K_1). \qquad 11.12\text{-}12$$

Mivel **A** első el nem tűnő aldeterminánsa másodrendű:

$$\det \mathbf{A}^{(K_1L_1)} = 0,$$

tehát

$$\sum_{l=1}^{N} A_{kl} \det \mathbf{A}^{(K_1K_2,\, L_1l)} = 0, \quad \text{ha} \quad k \neq K_1. \qquad 11.12\text{-}13$$

A $k \neq K_1$ feltétel onnan származik, hogy a kifejtési tétel 11.12-12 formája csak $k \neq K_1$ esetén érvényes, hiszen $\mathbf{A}^{(K_1 L_1)}$ elemei csak ebben az esetben egyeznek meg az eredeti mátrix elemeivel.

Belátjuk azonban, hogy 11.12-13 a $k = K_1$ esetben is érvényben marad. Cseréljük fel ugyanis 11.12-13-ban a K_1 és K_2 értékeket. A

$$\sum_{l=1}^{N} A_{kl} \det \mathbf{A}^{(K_2 K_1, L_1 l)} = 0 \quad (k \neq K_2) \qquad 11.12\text{-}14$$

összefüggéshez jutunk. Minthogy azonban $K_1 \neq K_2$ esetén

$$\det \mathbf{A}^{(K_2 K_1, Ll)} = -\det \mathbf{A}^{(K_1 K_2, Ll)}, \qquad 11.12\text{-}15$$

11.12-14 és 11.12-15-ből következik, hogy $K_1 \neq K_2$ esetén

$$\sum_{l} A_{kl} \det \mathbf{A}^{(K_1 K_2, L_1 l)} = 0 \qquad 11.12\text{-}16$$

tetszőleges k-ra fennáll.

A 11.12-16 összefüggés alapján megállapíthatjuk, hogy tetszőleges a érték esetén az

$$x_l^{(1)} = a \frac{\det \mathbf{A}^{(K_1 K_2, L_1 l)}}{\det \mathbf{A}^{(K_1 K_2, L_1 L_2)}} \quad (l = 1, \ldots, N) \qquad 11.12\text{-}17$$

értékek 11.12-19 egyenletrendszer olyan megoldásrendszerét alkotják, amelyben

$$x_{L_1}^{(1)} = 0, \quad \text{és} \quad x_{L_2}^{(1)} = a.$$

Az $a = 0$ választással a triviális $\mathbf{x} = \mathbf{0}$ megoldáshoz jutunk.

L_1 és L_2 szerepének felcserélésével az $\mathbf{Ax} = \mathbf{0}$ egyenletrendszer

$$x_l^{(2)} = b \frac{\det \mathbf{A}^{(K_1 K_2, L_2 l)}}{\det \mathbf{A}^{(K_1 K_2, L_1 L_2)}} \quad (l = 1, \ldots, N) \qquad 11.12\text{-}18$$

megoldásrendszeréhez jutunk, amelyben

$$x_{L_1}^{(2)} = -b \quad \text{és} \quad x_{L_2}^{(2)} = 0. \qquad 11.12\text{-}19$$

Minthogy azonban

$$\mathbf{Ax}^{(1)} = \mathbf{0} \quad \text{és} \quad \mathbf{Ax}^{(2)} = \mathbf{0},$$

$$\mathbf{A}(\mathbf{x}^{(1)} + \mathbf{x}^{(2)}) = \mathbf{0} \qquad 11.12\text{-}20$$

is fennáll, azaz

$$\mathbf{x}=\mathbf{x}^{(1)}+\mathbf{x}^{(2)} \qquad 11.12\text{-}21$$

is megoldása az $\mathbf{Ax}=\mathbf{0}$ egyenletrendszernek.

A 11.12-21 megoldásrendszerben

$$x_{L_1}=-b \quad \text{és} \quad x_{L_2}=a \qquad 11.12\text{-}22$$

szabadon választott értékek.

Összefoglalva a fentieket: *A kétszeresen elfajuló homogén lineáris egyenletrendszernek van nem triviális megoldása, és a megoldás során két megfelelő ismeretlen értékét szabadon megválaszthatjuk.*

11.12.3. Az elfajult homogén lineáris egyenletrendszer általános esete

Foglalkozzunk most az

$$\mathbf{Ax}=0 \qquad 11.12\text{-}23$$

homogén lineáris egyenletrendszer megoldásával abban az esetben, ha

$$\mathcal{R}(\mathbf{A})=N-n. \qquad 11.12\text{-}24$$

Legyen tehát

$$\det \mathbf{A}^{(\mathbf{K}(n-1),\,\mathbf{L}(n-1))}=0 \qquad 11.12\text{-}25$$

tetszőleges $\mathbf{K}(n-1)$ és $\mathbf{L}(n-1)$ indexcsoportra, azonban létezzen legalább egy $\mathbf{K}(n)$, $\mathbf{L}(n)$ indexcsoport, amelyre

$$\det \mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}\neq 0 \qquad \begin{matrix}(\mathbf{K}(n)=K_1\dots K_n) \\ (\mathbf{L}(n)=L_1\dots L_n)\end{matrix} \qquad 11.12\text{-}26$$

Az $n-1$-ed rendű aldeterminánsokra felírva a kifejtési tételt, a

$$\sum_{l=1}^{N} A_{kl} \det \mathbf{A}^{(\mathbf{K}(n),\,L_1\dots L(n-1)l)}=\delta_{kK_n} \det \mathbf{A}^{(\mathbf{K}(n-1),\,\mathbf{L}(n-1))} \quad (k\neq K_1,\dots,K_{n-1}) \qquad 11.12\text{-}27$$

formulához jutunk. 11.12-26 $k=K_n$-re érvényes, a $k\neq K_1,\dots,K_{n-1}$ feltételek pedig onnan származnak, hogy az $n-1$-ed rendű almátrixok elemei csak ebben az esetben egyenlők az eredeti mátrixelemekkel.

A kifejtési tételt a $\mathbf{K}(n)$ elemeiből permutációkkal képezett $\mathbf{K}'(n)$ indexcsoportokra felírva,

$$\sum_{l=1}^{N} A_{kl} \det \mathbf{A}^{(\mathbf{K}'(n),\,\mathbf{L}(n-1)l)}=0 \quad (k\neq K'_1,\dots,K'_{n-1}) \qquad 11.12\text{-}28$$

típusú formulákhoz jutunk. Itt $\mathbf{L}(n-1)l = L_1 \ldots L_{n-1}l$. 11.12-28 a $k = K_n'$ esetben érvényes. Mivel azonban

$$\det \mathbf{A}^{(\mathbf{K}'(n),\ \mathbf{L}(n-1)l)} = \pm \det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n-1)l)}, \qquad 11.12\text{-}29$$

ha $K_1 \neq K_2 \neq \ldots \neq K_n$, 11.12-27-ből és 11.12-28-ból következik, hogy

$$\sum_{l=1}^{N} A_{kl} \det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n-1)l)} = 0 \qquad 11.12\text{-}30$$

tetszőleges k értékre fennáll.

11.12-30-at az eredeti 11.12-23 egyenlettel összevetve, megállapítható, hogy az

$$x_l^{(n)} = a_n \frac{\det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n-1)l)}}{\det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}} \quad (l = 1, \ldots, N)$$

értékek az $\mathbf{Ax} = \mathbf{0}$ egyenletrendszer olyan megoldásrendszerét alkotják, amelyben

$$x_{L_n}^{(n)} = a_n \quad \text{és} \quad x_{L_\nu}^{(n)} = 0 \quad (\nu = 1, \ldots, n-1). \qquad 11.12\text{-}31$$

$a_n \neq 0$ esetén tehát $(n-1)$ ismeretlen értéke zérus, $x_{Ln}^{(n)} = a_n$ azonban nem zérus, azaz a megoldásrendszer nem triviális.

$\mathbf{L}(n)$ komponenseit permutálva, hasonló meggondolással az

$$x_l^{(\nu)} = a_\nu \frac{\det \mathbf{A}^{(\mathbf{K}(n),\ L_1 \ldots L_{\nu-1} l L_{\nu+1} \ldots L_n)}}{\det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}} \quad (l = 1, \ldots, N) \qquad 11.12\text{-}32$$

megoldásrendszerhez jutunk. Itt $L_1 \ldots L_{\nu-1} l L_{\nu+1} \ldots L_n$ azt jelenti, hogy az $\mathbf{L}(n)$ indexcsoport ν-edik eleme változó. Az így kapott kifejezésekre fennáll az

$$x_{L_\mu}^{\nu} = a_\nu \delta_{\nu\mu} \quad (\mu = 1, 2, \ldots, n) \qquad 11.12\text{-}33$$

összefüggés. Az egyes megoldásrendszerekben tehát az $x_{L_1}, \ldots, x_{L_n}$ ismeretlenek közül legfeljebb egy különbözik zérustól.

Mivel

$$\mathbf{Ax}^{(\nu)} = \mathbf{0} \quad (\nu = 1, \ldots, n),$$

az $\mathbf{x}^{(\nu)}$ megoldásrendszerek lineáris kombinációja is az egyenletrendszer megoldása. Az

$$x_l = x_2^{(1)} + \ldots + x_l^{(n)} \quad (l = 1, \ldots, N)$$

összeg azt a megoldást szolgáltatja, amelyben

$$x_{l_\nu}=a_\nu \quad (\nu=1, \ldots, n). \tag{11.12-34}$$

Összefoglalva a fentieket, *megállapítható, hogy az n-ed rendben elfajuló homogén lineáris egyenletrendszernek van nem triviális megoldása, mert a megoldás során n ismeretlen értékét szabadon választhatjuk meg.*

11.12.4. Az elfajuló egyenletrendszer megoldásainak vizsgálata

Az előzőekben beláttuk, hogy az elfajult homogén egyenletrendszereknek mindig van nem triviális megoldása. A 11.12.3. fejezetben általános módszert dolgoztunk ki, amellyel az egyenletrendszer egy megoldását meghatározhatjuk. Ebben a fejezetben megmutatjuk, hogy a 11.12.3-ban kidolgozott módszerrel az egyenletrendszer összes megoldásait meghatározhatjuk.

A 11.12.3-ban megállapítottuk, hogy az $\mathbf{Ax}=\mathbf{0}$ egyenletrendszer megoldásában

$$\mathcal{R}(\mathbf{A})=N-n,$$

tehát

$$\det \mathbf{A}^{(\mathrm{K}(n),\, \mathrm{L}(n))}\neq 0 \tag{11.12-35}$$

esetén megfelelő n ismeretlen értékét szabadon választhatjuk meg.

Legyen

$$x_{L_\nu}=a_\nu \quad (\nu=1, \ldots, n). \tag{11.12-36}$$

Helyettesítsük be az $\mathbf{Ax}=\mathbf{0}$ egyenletrendszerbe a 11.12-35-nek megfelelően választott x_{L_ν} értékeket, és vigyük át a konstans értéket a jobb oldalra. A kapott egyenletrendszer megoldása azonos az eredeti $\mathbf{Ax}=\mathbf{0}$ egyenletrendszer megoldásával.

Figyelembe véve a már megválasztott x_{L_ν} értékeket, az eredeti egyenletrendszerrel egyenértékű az

$$\mathbf{Ax}^{(\mathrm{K}(n),\, \mathrm{L}(n))}=\mathbf{Y} \tag{11.12-37}$$

egyenletrendszer, ahol

$$Y_k=\begin{cases} -\sum\limits_{\nu=1}^{n} A_k L_\nu a_\nu, & \text{ha} \quad k\neq K_\mu \quad (\mu=1, \ldots, n), \\ a_\mu, & \text{ha} \quad k=K_\mu. \end{cases} \tag{11.12-38}$$

11.12-36-ban $N-n$ egyenlet (ahol $k\neq K_\mu$) megegyezik az eredeti egyenletrendszer megfelelő egyenleteivel, és ezek az egyenletek csak $N-n$ számú ismeretlent (ahol $l\neq L_\nu$) tartalmaznak. A fennmaradó n egyenlet pedig az $x_{L_\nu}=a_\nu$ feltételekkel azonos, így mintegy a szabadon választható ismeretleneket határozza meg.

Az $\mathbf{Ax}=\mathbf{0}$ egyenletrendszer 11.12-35 feltételt kielégítő megoldásrendszerére teljesülnie kell a 11.12-36 egyenletrendszernek is. Így 11.12-36 egy szükséges feltételt jelent az eredeti egyenlet megoldásrendszerére. 11.12-35 miatt a 11.12-37 inhomogén egyenletrendszer egyértelműen megoldható. 11.9-2 szerint 11.12-37 megoldása az

$$\mathbf{x}=\left(\mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}\right)^{-1}\mathbf{Y}. \qquad 11.12\text{-}39$$

Az előző fejezetben kimutattuk, hogy az $\mathbf{Ax}=\mathbf{0}$ egyenletrendszernek van megoldása. A megoldásnak ki kell elégítenie a 11.12-37 egyenletrendszert. 11.12-39 viszont a 11.12-37 egyenletrendszer egyértelmű megoldása. Következésképpen 11.12-39 az $\mathbf{Ax}=\mathbf{0}$ egyenletrendszer egyetlen olyan megoldását szolgáltatja, amely eleget tesz a 11.12-36 feltételnek.

Amennyiben tehát a szabadon választható ismeretleneket rögzítjük, az elfajuló egyenletrendszer megoldása egyértelművé válik. Megjegyezzük még, hogy ha

$$a_\nu=0 \quad (\nu=1,\ldots,n) \qquad 11.12\text{-}40$$

értékeket választunk, akkor az egyenletrendszernek csak triviális $\mathbf{x}=\mathbf{0}$ megoldását kapjuk.

Az inhomogén egyenletrendszer megoldását az $\mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}$ almátrix segítségével határoztuk meg. Felmerül a kérdés, hogy nem jutunk-e újabb megoldásokhoz, ha létezik olyan $\mathbf{A}^{(\mathbf{K}(n),\,\mathbf{L}(n))}\neq\mathbf{A}^{(\mathbf{K}'(n),\,\mathbf{L}'(n))}$ almátrix, amelynek determinánsa szintén nem zérus, s az előzőekben ismertetett megoldási eljárást az $\mathbf{A}^{(\mathbf{K}'(n),\,\mathbf{L}'(n))}$ mátrix segítségével hajtjuk végre.

Az $\mathbf{A}^{(\mathbf{K}'(n),\,\mathbf{L}'(n))}$ almátrix determinánsának segítségével az eredeti egyenletrendszer megoldásrendszerét az

$$\mathbf{x}'=\sum_{\nu=1}^{n}\mathbf{x}^{(\nu)'} \qquad 11.12\text{-}41$$

alakban állíthatjuk elő, ahol

$$x_l^{(\nu)'}=a'_\nu\,\frac{\det\mathbf{A}^{(K'_1K'_2\ldots K'_n,\;L'_1L'_{\nu-1}lL'_{\nu+1}\cdots L_n)}}{\det\mathbf{A}^{(\mathbf{K}'\mathbf{L}')}}. \qquad 11.12\text{-}42$$

A 11.12-42 megoldásrendszer is tartalmaz n szabadon választható ismeretlent, amelyek most az

$$x_{L'_\nu}=a'_\nu \qquad 11.12\text{-}43$$

értékek. A megoldás a szabadon választható ismeretlenek értékének rögzítése után egyértelműen meghatározható.

Az így nyert megoldások tartalmazzák az egyenletrendszer 11.12-36 feltételnek eleget tevő megoldását is. Az $\mathbf{A}^{(K(n)L(n))}$ almátrix segítségével nyert 11.12-39 megoldásrendszerből kiválaszthatjuk az x_{L_ν} értékeket. Ezen értékek előzetes megadásával az $\mathbf{A}^{(\mathbf{K}'(n),\ \mathbf{L}'(n))}$ almátrix segítségével egy egyértelmű $\mathbf{x}'$ megoldásrendszerhez juthatunk. Mivel mind a 11.12-39-nek megfelelő $\mathbf{x}$, mind pedig $\mathbf{x}'$ egyértelmű megoldásrendszere az eredeti homogén egyenletrendszernek, kell, hogy

$$\mathbf{x}=\mathbf{x}'$$

legyen.

A különböző nem zérus értékű aldeterminánsok felhasználásával tehát nem juthatunk az egyenletrendszer különböző megoldásaihoz.

11.13. Az elfajult inhomogén egyenletrendszer

Határozzuk meg az

$$\mathbf{A}\mathbf{x}=\mathbf{y} \qquad 11.13\text{-}1$$

inhomogén egyenletrendszer megoldását, ahol az együtthatómátrix rangja

$$\mathcal{R}(\mathbf{A})=N-n, \qquad 11.13\text{-}2$$

és legyen

$$\det \mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}\neq 0. \qquad 11.13\text{-}3$$

Első lépésként keressük a 11.13-1 egyenletrendszer megoldását az

$$x_{L_\nu}=0 \quad (\nu=1,\ldots,n) \qquad 11.13\text{-}4$$

feltételek mellett. A feltétel szerint tehát n ismeretlen értékét önkényesen zérusnak választottuk. Amennyiben 11.13-1-nek van 11.13-4-nek megfelelő megoldása, akkor ez a megoldás kielégíti az

$$\mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}\mathbf{x}=\mathbf{Y} \qquad 11.13\text{-}5$$

egyenletrendszert is, ahol

$$Y_k=\begin{cases} y_k, & \text{ha} \quad k\neq K_\nu, \\ 0, & \text{ha} \quad k=K_\nu \end{cases} \quad (\nu=1,\ldots,n). \qquad 11.13\text{-}6$$

A 11.13-5 egyenletrendszer azonban 11.13-3 miatt egyértelműen megoldható. 11.13-5 megoldását az

$$\mathbf{x}=\left(\mathbf{A}^{(\mathbf{K}(n),\ \mathbf{L}(n))}\right)^{-1}\mathbf{Y} \qquad 11.13\text{-}7$$

összefüggés határozza meg. Amennyiben tehát az inhomogén egyenletrendszernek van 11.13-4-nek megfelelő megoldása, akkor azt 11.13-7 szolgáltatja.

A 11.13-7 megoldás előállításakor azonban nem használtuk fel az eredeti inhomogén egyenletrendszer K_μ-edik ($\mu=1, 2, \ldots, n$) egyenleteit. 11.13-7 tehát csak akkor megoldása az eredeti egyenletrendszernek, ha kielégíti a K_μ-edik egyenleteket is, azaz eleget tesz a

$$\sum A_{kl} x_l = y_k \quad (k = K_\mu;\ \mu = 1, \ldots, n) \qquad 11.13\text{-}8$$

feltételeknek. 11.13-7 felhasználásával 11.13-8 az

$$[\mathbf{A}(\mathbf{A}^{(K(n),\,L(n))^{-1}}\mathbf{y})]_{K_\mu} = y_{K_\mu} \quad (\mu = 1, \ldots, n) \qquad 11.13\text{-}9$$

alakot ölti. 11.13-9 n összefüggést ad az y_k értékek között pusztán az együttható-mátrix elemeinek felhasználásával.

Az y_k értékek közül $N-n$ számot tetszés szerint megválasztva, a fennmaradó n érték a 11.13-9 lineáris kombinációk segítségével meghatározható úgy, hogy a 11.13-1 lineáris egyenletrendszernek legyen 11.13-4-nek megfelelő megoldása.

Keressük most az

$$\mathbf{Ax} = \mathbf{y}$$

egyenletrendszer megoldását az

$$x_{L_\nu} = a_\nu \quad (\nu = 1, \ldots, n) \qquad 11.13\text{-}10$$

feltételek mellett.

Tételezzük fel, hogy az egyenletrendszernek van 11.13-4-nek megfelelő megoldása, és jelöljük ezt a megoldást

$$\mathbf{x} = \mathbf{x}^{(0)} \qquad 11.13\text{-}11$$

val. Az előző fejezetekből tudjuk, hogy az

$$\mathbf{Ax} = \mathbf{0} \qquad 11.13\text{-}12$$

homogén egyenletrendszernek létezik olyan

$$\mathbf{x} = \mathbf{x}^{(a)} \qquad 11.13\text{-}13$$

megoldása, amelyben

$$x_{L_\nu} = a_\nu \quad (\nu = 1, \ldots, n).$$

Azonnal látható, hogy

$$\mathbf{x} = \mathbf{x}^{(0)} + \mathbf{x}^{(a)} \qquad 11.13\text{-}14$$

az inhomogén egyenletrendszer 11.13-10 feltételnek is eleget tevő megoldása, hiszen

$$\mathbf{A}(\mathbf{x}^{(0)} + \mathbf{x}^{(a)}) = \mathbf{Ax}^{(0)} + \mathbf{Ax}^{(a)} = \mathbf{y},$$

mert feltevéseink szerint $\mathbf{Ax}^{(0)} = \mathbf{y}$ és $\mathbf{Ax}^{(a)} = \mathbf{0}$.

A 11.13-14 megoldás mindig létezik, ha a 11.13-11 megoldás létezik. 11.13-11 létezésének feltétele pedig éppen a 11.13-9 egyenletek teljesülése.

Tehát 11.13-9 az elfajuló inhomogén egyenletrendszer megoldásának általános feltétele.

11.13-14 az inhomogén egyenletrendszer egyetlen olyan megoldása, amely a 11.13-10 feltételt kielégíti. Tételezzük fel ugyanis, hogy az inhomogén egyenletrendszernek az adott feltételek mellett két, $\mathbf{x}$ és $\mathbf{x}'$ megoldása is létezik. Ekkor a 11.13-12 homogén egyenletrendszer megoldása kell legyen, tehát

$$\mathbf{A}\mathbf{X} = 0.$$

Feltételeink szerint mind $\mathbf{x}$, mind $\mathbf{x}'$ eleget tesz 11.13-10-nek, tehát

$$X_{L_\nu} = 0 \quad (\nu = 1, \ldots, n). \tag{11.13-16}$$

Ez azonban azt jelenti, hogy $\mathbf{X}$ a homogén egyenletrendszernek olyan megoldása, ahol a szabadon választható ismeretlenek értékét zérusnak választottuk. Ebben az esetben azonban $\mathbf{X}$ csak a triviális $\mathbf{X} = \mathbf{0}$ megoldás lehet, tehát

$$\mathbf{x} = \mathbf{x}'.$$

Összefoglalva az eddigieket, megállapítottuk, hogy az n-ed rendben elfajult lineáris egyenletrendszer csak akkor oldható meg, ha a konstans értékek n darab homogén lineáris összefüggésnek tesznek eleget. Ezek az összefüggések homogén lineáris egyenletrendszer esetén mindig teljesülnek.

Inhomogén esetben pedig $N - n$ számú y_k értéket szabadon választhatunk, a fennmaradó n számú y_k értéket pedig a 11.12-9 feltételi egyenletből kell meghatároznunk, ha azt akarjuk, hogy az egyenletrendszer megoldható legyen.

Ha $\mathbf{y}$ *komponenseire a 11.13-9 feltételek teljesülnek, akkor az* $\mathbf{x}$ *megoldás* n *komponensének tetszés szerinti megválasztása után az inhomogén egyenletrendszer egyértelműen megoldható.*

11.14. Alkalmazás

A lineáris egyenletrendszerek a fizika szinte minden területén előfordulnak. A következőkben először egy elvi szempontból fontos — a mátrixok rangjával kapcsolatos — tételt, majd egy áramköri problémát tárgyalunk.

11.14.1. Tétel a mátrixok rangjával kapcsolatban

Legyen $\mathbf{S}$ egy $N \times N$-es mátrix, amelyre

$$\det \mathbf{S} \neq 0. \tag{11.14-1}$$

Képezzük az $\mathbf{S}$ mátrix és az ugyancsak $N \times N$-es $\mathbf{G}$ mátrix segítségével a

$$\mathbf{T} = \mathbf{S}\mathbf{G}\tilde{\mathbf{S}} \tag{11.14-2}$$

mátrixot. Bebizonyítjuk, hogy a $\mathbf{T}$ és $\mathbf{G}$ mátrixok rangja megegyezik, tehát

$$\mathcal{R}(\mathbf{T}) = \mathcal{R}(\mathbf{G}). \tag{11.14-3}$$

Tegyük fel, hogy

$$\mathcal{R}(\mathbf{T})=N-n, \qquad 11.14\text{-}4$$

és vizsgáljuk a

$$\mathbf{Tx}=\mathbf{0} \qquad 11.14\text{-}5$$

egyenletrendszer megoldásait. A 11.12. fejezetben megállapítottuk, hogy a 11.14-5 egyenletrendszernek $\mathcal{R}(\mathbf{T})=N-n$ miatt n lineárisan független megoldásrendszere létezik, tehát

$$\mathbf{Tx}^{(k)}=\mathbf{0} \quad (k=1, 2, \ldots, n). \qquad 11.14\text{-}6$$

Szorozzuk meg 11.14-2-t balról S^{-1}-gyel, jobbról pedig $\mathbf{x}^{(k)}$-val. A

$$\mathbf{G}(\tilde{\mathbf{S}}\mathbf{x}^{(k)})=\mathbf{0} \qquad 11.14\text{-}7$$

egyenlethez jutunk. Bevezetve az

$$\mathbf{y}^{(k)}=\tilde{\mathbf{S}}\mathbf{x}^{(k)} \qquad 11.14\text{-}8$$

jelölést, 11.14-7 a

$$\mathbf{Gy}^{(k)}=\mathbf{0} \quad (k=1, 2, \ldots, n) \qquad 11.14\text{-}9$$

alakot ölti. Minthogy $\det \mathbf{S}\neq 0$, 11.14-8 a 11.14-9 egyenletrendszer lineárisan független megoldásait adja. Ebből viszont következik, hogy

$$\mathcal{R}(\mathbf{G})\leq\mathcal{R}(\mathbf{T}). \qquad 11.14\text{-}10$$

Hasonló módon, ha 11.14-2 helyett az

$$\mathbf{S}^{-1}\mathbf{T}\tilde{\mathbf{S}}^{-1}=\mathbf{G}$$

összefüggésből indulunk ki, azt kapjuk, hogy

$$\mathcal{R}(\mathbf{G})\geq\mathcal{R}(\mathbf{T}). \qquad 11.14\text{-}11$$

11.14-10 és 11.14-11-ből viszont következik, hogy

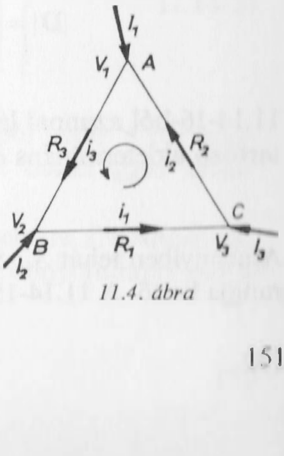
$$\mathcal{R}(\mathbf{G})=\mathcal{R}(\mathbf{T}).$$

Ezzel tételünket bebizonyítottuk.

11.14.2. Egy áramköri probléma

A következőkben egy egyenáramú hálózat potenciálviszonyait határozzuk meg. A probléma tárgyalását egy egyszerű speciális esettel, a *11.4. ábrán* látható áramkör A, B, C pontjaihoz tartozó V_1, V_2, V_3 potenciálok meghatározásával kezdjük.

Legyen az ABC háromszög oldalainak ellenállása R_1, R_2, R_3. Jelöljük a vezetőkörbe az A, B, C pontokon keresztül kívülről befolyó áramokat I_1, I_2, I_3-mal,



11.4. ábra

az egyes ellenállásokon folyó áramokat pedig i_1, i_2, i_3-mal. Az áramok, ill. feszültségek közötti kapcsolatot a Kirchhoff-törvények adják meg. Kirchhoff I. törvénye szerint töltések az áramkör egyetlen pontjában sem halmozódhatnak fel, így az A, B, C pontokban rendre érvényesek a következő egyenletek:

$$\left.\begin{aligned} i_3-i_2&=I_1\\ i_1-i_3&=I_2\\ i_2-i_1&=I_3. \end{aligned}\right| \qquad 11.14\text{-}12$$

A 11.14-12 egyenletrendszer csak akkor teljesülhet, ha

$$I_1+I_2+I_3=0. \qquad 11.14\text{-}13$$

Az egyes ágakra érvényes az Ohm-törvény, így fennállnak az

$$\left.\begin{aligned} i_1&=(V_3-V_2)S_1\\ i_2&=(V_1-V_3)S_2\\ i_3&=(V_2-V_1)S_3 \end{aligned}\right\} \qquad 11.14\text{-}14$$

egyenletek. 11.14-14-ben S_i-vel ($i=1, 2, 3$) az

$$S_i=\frac{1}{R_i}$$

vezetőképességeket jelöltük.

A 11.14-14 egyenleteket 11.14-12-be beírva, a

$$\left.\begin{aligned} -V_1(S_2+S_3)+V_2S_3+V_3S_2&=I_1\\ V_1S_3-V_2(S_3+S_1)+V_3S_1&=I_2\\ V_1S_2+V_2S_1-V_3(S_1+S_2)&=I_3 \end{aligned}\right| \qquad 11.14\text{-}15$$

egyenletrendszerhez jutunk. Az I_k ($k=1, 2, 3$) áramok és S_l ($l=1, 2, 3$) vezetőképességek ismeretében 11.14-15-ből a V_k potenciálok meghatározhatók. A 11.14-15 egyenletrendszer elfajuló, hiszen determinánsa zérus:

$$|\mathbf{D}|=\begin{vmatrix} -(S_2+S_3) & S_3 & S_2\\ S_3 & -(S_3+S_1) & S_1\\ S_2 & S_1 & -(S_1+S_2) \end{vmatrix}. \qquad 11.14\text{-}16$$

11.14-16-ból azonnal látható, hogy a $\mathbf{D}$ mátrix tetszőleges diagonális eleméhez tartozó aldetermináns értéke:

$$\det \mathbf{D}^{(k,k)}=S_1S_2+S_2S_3+S_3S_1.$$

Amennyiben tehát S_1, S_2, S_3 közül legalább kettő nem zérus, akkor a $\mathbf{D}$ mátrix rangja kettő. A 11.14-15 egyenletrendszert ebben az esetben úgy oldhatjuk meg, hogy egy potenciálértéket önkényesen megválasztunk, és a három egyenletből egyet elhagyunk.

Legyen pl. $V_1=0$, és oldjuk meg a

$$\left.\begin{aligned} -V_2(S_1+S_2)+V_3S_1&=I_2 \\ V_2S_1-V_3(S_1+S_2)&=I_3 \end{aligned}\right\} \qquad 11.14\text{-}17$$

egyenletrendszert.

11.14-17 megoldását a

$$\left.\begin{aligned} V_2&=-\frac{I_2(S_1+S_2)+I_3S_1}{S_1S_2+S_2S_3+S_3S_1} \\ &\text{és} \\ V_3&=-\frac{I_2S_1+I_3(S_3+S_1)}{S_1S_2+S_2S_3+S_3S_1} \end{aligned}\right\} \qquad 11.14\text{-}18$$

formulák adják.

Gyakorlati szempontból fontos az $I_2=0$, $S_2=0$ speciális eset. Ekkor 11.14-13 miatt

$$I_1+I_3=0,$$

tehát

$$I_1=I \quad \text{és} \quad I_3=-I.$$

A potenciálértékeket pedig 11.14-18 szerint a

$$\left.\begin{aligned} V_3&=\frac{I(S_1+S_3)}{S_1S_3}=I(R_1+R_3), \\ V_2&=IR_3 \end{aligned}\right\} \qquad 11.14\text{-}19$$

formulák adják.

Ha $S_2\neq 0$, akkor

$$\left.\begin{aligned} V_3&=\frac{I(S_1+S_3)}{S_1S_2+S_2S_3+S_1S_3}=\frac{I}{\dfrac{1}{R_2}+\dfrac{1}{R_1+R_3}}, \\ V_2&=\frac{IS_1}{S_1S_2+S_2S_3+S_3S_1}. \end{aligned}\right\} \qquad 11.14\text{-}20$$

A 11.14-19 formulákból a sorosan, a 11.14-20 formulákból pedig a párhuzamosan kapcsolt ellenállások eredőjére vonatkozó ismert összefüggés olvasható ki.

A következőkben a fenti formalizmust bonyolultabb esetekre általánosítjuk. Természetesen a probléma tárgyalásához a fentinél kissé bonyolultabb jelöléseket kell alkalmaznunk.

Legyen adva egy $N+1$ pontból ($P_1, P_2, \ldots, P_{N+1}$) álló áramkör. Az egyes pontok között elhelyezkedő ellenállásokat jelöljük R_{kl}-lel. Nyilvánvaló, hogy

$$R_{kl}=R_{lk}, \qquad 11.14\text{-}21$$

valamint hogy $l \neq k$. Jelöljük az egyes pontok között a vezetőképességet S_{kl}-lel, tehát

$$S_{kl}=\frac{1}{R_{kl}}.$$

Jelöljük a hálózat egyes pontjainak potenciálját rendre $V_1, V_2, \ldots, V_{N+1}$-gyel, és tegyük fel, hogy a hálózat egyes pontjaiba rendre $I_1^{(0)}, I_2^{(0)}, \ldots, I_{N+1}^{(0)}$ nagyságú külső áramok folynak be.

Kirchhoff első törvénye miatt azonban

$$\sum_{k=1}^{N+1} I_k^{(0)}=0, \qquad 11.14\text{-}22$$

hiszen az áramkörben nem halmozódnak fel töltések.

Jelöljük a hálózat P_k pontjából a P_l pontba folyó áram értékét I_{kl}-lel. Mivel $I_{kl}=-I_{lk}$, a hálózatban folyó áramok egy I antiszimmetrikus mátrixszal jellemezhetők.

A hálózat egyes pontjaira felírva Kirchhoff első törvényét:

$$\sum_{\substack{l=1\\ l\neq k}}^{N+1} I_{kl}+I_k^{(0)}=0 \quad (k=1, \ldots, N+1). \qquad 11.14\text{-}23$$

Kirchhoff második törvénye szerint

$$I_{kl}=S_{kl}(V_k-V_l). \qquad 11.14\text{-}24$$

11.14-24-et 11.14-23-ba beírva, azt kapjuk, hogy

$$\sum_{\substack{l=1\\ k\neq l}}^{N+1} S_{kl}(V_k-V_l)+I_k^{(0)}=0. \qquad 11.14\text{-}25$$

A 11.14-25 egyenletrendszert az

$$\overline{\mathbf{S}}\mathbf{V}=\mathbf{I}^{(0)} \qquad 11.14\text{-}26$$

összefoglaló alakban is felírhatjuk, ha az $\overline{\mathbf{S}}$ mátrixot a következőképpen definiáljuk:

$$\bar{S}_{kl}=\begin{cases} S_{kl} & \text{ha} \quad k\neq l, \\ -\sum\limits_{\substack{m=1\\ m\neq k}}^{N+1} S_{km} & \text{ha} \quad k=l, \end{cases}$$

$$\mathbf{V}=V_1 \dots V_{N+1},$$ 11.14-27

$$I^{(0)}=I_1^{(0)} \dots I_{N+1}^{(0)}.$$

A 11.14-26 egyenletrendszer segítségével az S_{kl} értékek és az $I_k^{(0)}$ külső áramok ismeretében meghatározhatjuk az egyes pontokhoz tartozó potenciálértékeket. 11.14-26 tulajdonképpen az Ohm-törvény általánosítása, amelyben a feszültség és az áramerősség helyén egy egydimenziós, a vezetőképesség helyén pedig egy kétdimenziós mátrix szerepel. A potenciálokat tehát egy lineáris egyenletrendszer megoldásai szolgáltatják. A 11.14-26 egyenletrendszer azonban elfajult, hiszen a 11.14-27 definícióból következik, hogy

$$\det \bar{\mathbf{S}}=0.$$ 11.14-28

Erről meggyőződhetünk pl. úgy, hogy az $\bar{S}$ determináns minden oszlopát összeadjuk. Eredményül csupa zérust kapunk, így 11.14-28 valóban fennáll.

A 11.14-26 egyenletrendszer nem triviális megoldásait az $\bar{\mathbf{S}}$ mátrix legmagasabb rendű el nem tűnő almátrixának segítségével határozhatjuk meg. Tegyük fel, hogy $\mathcal{R}(\bar{\mathbf{S}})=N$, és tételezzük fel, hogy

$$\det \bar{\mathbf{S}}^{(NN)}\neq 0.$$ 11.14-29

11.14-29 szerint azt tesszük fel, hogy az $\bar{\mathbf{S}}$ mátrix jobb alsó sarokeleméhez tartozó aldetermináns nem zérus.

Ekkor a 11.12. fejezet szerint egy ismeretlen értékét szabadon választhatjuk. 11.14-29 teljesülése esetén ez az ismeretlen V_{N+1}. Ez megfelel annak, hogy az $N+1$ pontból álló hálózat esetén abszolút értelemben nem rendelhetünk potenciálokat az egyes pontokhoz. Az egyszerűség kedvéért célszerű a $V_{N+1}=0$ értéket választani, így minden pont potenciálját az $N+1$-edik ponthoz viszonyítjuk.

A fenti feltételek mellett 11.14-26 megoldását 11.12-32 szerint

$$\mathbf{V}=(\bar{\mathbf{S}}^{(NN)})^{-1}\mathbf{I}^{(0)\prime}$$ 11.14-30

szolgáltatja, ahol

$$I^{(0)\prime}=I_1^{(0)},\ I_2^{(0)},\ \dots,\ I_N^{(0)},\ 0.$$ 11.14-31

11.14-30 azonban csak akkor megoldása az eredeti 11.14-26 egyenletrendszernek, ha teljesül a 11.12-8-nak megfelelő

$$\sum_l \bar{S}_{N+1\,l} V_l = I^{(0)}_{N+1} \qquad 11.14\text{-}32$$

feltétel.

11.14-32-nek azonban a 11.14-22 fizikai feltétel miatt automatikusan teljesülnie kell. A 11.14-30 megoldás meghatározásakor ugyanis a 11.14-26 egyenletrendszert a

$$\sum_l \bar{\mathbf{S}}^{(NN)} V_l = I_k^{(0)\prime} \quad (k=1, \ldots, N+1) \qquad 11.14\text{-}33$$

egyenletrendszerrel helyettesítettük. 11.14-33 egyenleteit összeadva, és a 11.14-27 definíciót figyelembe véve, a

$$-\sum_{k=1}^{N+1} \bar{S}_{N+1\,l} V_l = \sum_{k=1}^{N+1} I_k^{(0)\prime} = \sum_{k=1}^{N} I_k^{(0)} \qquad 11.14\text{-}34$$

egyenletet kapjuk.

11.14-32-t és 11.14-34-et összevetve, a

$$-\sum_{k=1}^{N} I_k^{(0)} = I_{N+1}^{(0)} \qquad 11.14\text{-}35$$

feltételhez jutunk, ami azonos 11.14-22-vel.

Bevezetve az

$$\mathbf{R} = (\bar{\mathbf{S}}^{(NN)})^{-1}$$

jelölést és az értelemszerű ellenállásmátrix elnevezést, 11.14-30 a

$$\mathbf{V} = \mathbf{R}\mathbf{I}^{(0)\prime} \qquad 11.14\text{-}36$$

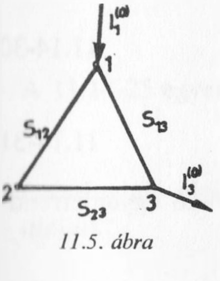
alakot ölti. 11.14-36 pontos analogonja az Ohm-törvénynek.

11.14-36 segítségével $\mathbf{I}^{(0)}$ és $\mathbf{R}$ ismeretében meghatározhatjuk a V_k potenciálértékeket. Nem szükséges azonban, hogy 11.14-36-ban a V_k értékeket tekintsük ismeretlennek. A 11.14-36 egyenlet a $V_{N+1}=0$ és 11.14-35 feltétel teljesülése mellett az I_k és V_k $(k=1, \ldots, N)$ mennyiségek közül bármely N számú mennyiség meghatározására alkalmas, ha a fennmaradó N értéket megadjuk.

Befejezésül megvizsgáljuk, hogy mi a feltétele a

$$\det \bar{\mathbf{S}}^{(NN)} = 0 \qquad 11.14\text{-}37$$

egyenlet teljesülésének.



11.5. ábra

Térjünk vissza a bevezetésben már tárgyalt egyszerű hálózatra. A fenti általános formalizmusból a *11.5. ábrán* látható hálózat esetén adódik, hogy

$$\overline{\mathbf{S}}^{(22)}=\begin{pmatrix} -(S_{12}+S_{13}) & S_{12} & 0 \\ S_{12} & -(S_{12}+S_{23}) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad 11.14\text{-}38$$

$$\det \overline{\mathbf{S}}^{(22)}=S_{12}S_{23}+S_{12}S_{13}+S_{13}S_{23}. \qquad 11.14\text{-}39$$

Mivel $S_{kl} \geq 0$, $\det \overline{\mathbf{S}}^{(22)} \geq 0$. 11.14-39-ből látható, hogy $\det \overline{\mathbf{S}}^{(22)}$ akkor és csak akkor lehet zérus, ha a három különböző vezetőképesség-érték közül kettő zérus.

Fizikailag ez azt jelenti, hogy a három pont közül egy nincs kapcsolatban a másik kettővel. A $\det \overline{\mathbf{S}}^{(22)}=0$ egyenlőség tehát azt jelzi, hogy a potenciáleloszlás a három pont esetén független a bevezetett áramoktól. A probléma tehát fizikai szempontból érdektelen.

Amennyiben 11.14-39 nem zérus, akkor

$$(\overline{\mathbf{S}}^{(22)})^{-1}=\frac{1}{\det \overline{\mathbf{S}}^{(22)}}\begin{pmatrix} -(S_{12}+S_{23}) & -S_{12} & 0 \\ -S_{12} & -(S_{12}+S_{23}) & 0 \\ 0 & 0 & \det \overline{\mathbf{S}}^{(22)} \end{pmatrix}. \qquad 11.14\text{-}40$$

Ennek segítségével

$$\left.\begin{aligned} V_1&=\frac{-(S_{12}+S_{23})I_1^{(0)}-S_{12}I_2^{(0)}}{\det \overline{\mathbf{S}}^{(22)}} \\ &\text{és} \\ V_2&=\frac{-S_{12}I^{(0)}-(S_{12}+S_{13})I_2^{(0)}}{\det \overline{\mathbf{S}}^{(22)}} \end{aligned}\right\}. \qquad 11.14\text{-}41$$

Amennyiben $I_2^{(0)}=0$ és $S_{13}=0$, akkor

$$\frac{1}{\det \overline{\mathbf{S}}^{(22)}}=R_{12}R_{23},$$

és

$$V_1=(R_{12}+R_{23})I_1^{(0)}, \quad V_2=R_{23}I_1^{(0)}, \qquad 11.14\text{-}42$$

valamint mivel

$$V_3=0, \quad I_3^{(0)}=-I_1^{(0)}, \qquad 11.14\text{-}43$$

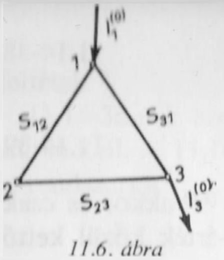
11.4-41 két, sorba kapcsolt ellenállás esetén érvényes formulákat jelent, és a formulából adódik a jól ismert

$$R_{\text{eredő}}=R_{12}+R_{23}$$

összefüggés.

Amennyiben $I_2^{(0)}=0$, de $S_{13}\neq 0$, akkor hasonló módon kaphatjuk az

$$\frac{1}{R_{\text{eredő}}}=\frac{1}{R_{13}}+\frac{1}{R_{12}+R_{23}} \qquad 11.14\text{-}44$$



11.6. ábra

formulát (*11.6. ábra*). Visszakaptuk tehát a bevezető egyszerű példa során nyert eredményeket.

$N \geq 3$ esetén det $\overline{S}^{(NN)}$ ugyancsak nemnegatív tagok összegeként állítható elő. Ezért det $\overline{S}^{(NN)} = 0$ csak akkor lehetséges, ha a determináns minden tagja zérus. Ez az eset bekövetkezhet például akkor, ha egy adott l esetén minden

$$S_{kl} = 0 \quad (k = 1, \ldots, N+1). \qquad 11.14\text{-}45$$

A 11.14-45 feltétel azt jelenti, hogy a P_l pont független a többitől.

Ugyancsak zérussá válik a determináns értéke akkor, ha

$$S_{kl} = 0, \quad \text{ha} \quad \left.\begin{array}{l} k = K_1, \ldots, K_n \\ l = K_{n+1}, K_{n+2}, \ldots, K_{N+1}, \end{array}\right\} \qquad 11.14\text{-}46$$

ahol $K_1 \ldots K_{N+1}$ az $1 \ldots N+1$ számok permutációja. Ez utóbbi esetben a $P_1, \ldots, P_{N+1}$ pontok két csoportba oszthatók úgy, hogy a $P_{K_1}, \ldots, P_{K_n}$ és $P_{K_{n+1}}, \ldots, P_{K_N}$ pontcsoportok függetlenek egymástól. A hálózat tehát két független részre osztható, így fizikailag nincs értelme az együttes tárgyalásnak.

Megjegyezzük még, hogy a megoldás során az $\overline{S}^{(NN)}$ almátrix nem játszik kitüntetett szerepet. A fenti gondolatmenet minden további nélkül elvégezhető tetszőleges $\overline{S}^{(KK)}$ almátrix segítségével.

III. TENZOROK

12. A homogén lineáris vektoroperátor vagy tenzor

A következőkben rendszerezzük és továbbfejlesztjük a 7. fejezetben a lineáris operátorokkal kapcsolatban nyert eredményeket.

A forgatási operátorok tulajdonságait vizsgálva megállapítottuk, hogy ha az $\underline{\underline{\mathbf{O}}}$ operátor egy $\underline{\mathbf{a}}$ és egy $\underline{\mathbf{b}}$ vektort az

$$\underline{\mathbf{a}}^* = \underline{\underline{\mathbf{O}}}\underline{\mathbf{a}}$$

és

$$\underline{\mathbf{b}}^* = \underline{\underline{\mathbf{O}}}\underline{\mathbf{b}}$$

vektorokba visz át, akkor tetszőleges α és β mellett

$$\underline{\underline{\mathbf{O}}}(\alpha\underline{\mathbf{a}} + \beta\underline{\mathbf{b}}) = \alpha\underline{\mathbf{a}}^* + \beta\underline{\mathbf{b}}^*. \qquad \text{12-1}$$

Általában azokat az operátorokat, amelyekre fennáll, hogy

$$\bar{\underline{\mathbf{a}}} = \underline{\underline{\mathbf{A}}}\underline{\mathbf{a}} \quad \text{és} \quad \bar{\underline{\mathbf{b}}} = \underline{\underline{\mathbf{B}}}\underline{\mathbf{b}}$$

esetén

$$\underline{\underline{\mathbf{A}}}(\alpha\underline{\mathbf{a}} + \beta\underline{\mathbf{b}}) = \alpha\bar{\underline{\mathbf{a}}} + \beta\bar{\underline{\mathbf{b}}}, \qquad \text{12-2}$$

homogén lineáris vektoroperátoroknak nevezzük.

A homogén lineáris vektoroperátorokat *tenzoroknak* is nevezzük. A forgatási operátorok tehát tenzorok.

A tenzorokat definiáló 12-2 tulajdonságból az $\alpha = 0$, $\beta = 0$ választással következik, hogy a tenzorok a zérusvektort zérusvektorra képezik le, azaz

$$\underline{\underline{\mathbf{A}}}\underline{\mathbf{n}} = \underline{\mathbf{n}},$$

ahol $\underline{\mathbf{n}}$ a nullvektor.

12.1. A tenzor jellemzése

A következőkben megvizsgáljuk, hogy milyen mennyiségek szükségesek egy tenzor egyértelmű jellemzéséhez.

A tenzort akkor tekintjük ismertnek, ha meg tudjuk adni, hogy a tér tetszőleges vektorát hogyan transzformálja. Ehhez elegendő, ha megadjuk, hogy a tenzor a tér három, nem komplanáris $\underline{a}$, $\underline{b}$, $\underline{c}$ vektorát mely

$$\bar{\underline{a}}=\underline{\underline{A}}\underline{a}, \quad \bar{\underline{b}}=\underline{\underline{A}}\underline{b}, \quad \bar{\underline{c}}=\underline{\underline{A}}\underline{c} \qquad 12.1\text{-}1$$

vektorokba viszi át.

Itt

$$(\underline{a}, \underline{b}, \underline{c}) \neq 0. \qquad 12.1\text{-}2$$

Mivel $\underline{a}$, $\underline{b}$, $\underline{c}$ nem esik egy síkba, a tér tetszőleges $\underline{d}$ vektora előállítható

$$\underline{d}=\alpha\underline{a}+\beta\underline{b}+\gamma\underline{c} \qquad 12.1\text{-}3$$

alakban. 12.1-1 ismeretében a $\underline{d}$ vektor $\bar{\underline{d}}$ transzformáltja a

$$\bar{\underline{d}}=\underline{\underline{A}}\underline{d}=\alpha\bar{\underline{a}}+\beta\bar{\underline{b}}+\gamma\bar{\underline{c}}$$

alakban állítható elő, 12.1-1 ismerete tehát valóban elegendő tetszőleges vektor transzformáltjának meghatározásához.

12.2. Az inverz operátor

Jelöljük az $\underline{\underline{A}}$ operátor inverzét $\underline{\underline{A}}^{-1}$-gyel. Az inverz operátor definíciója értelmében ez azt jelenti, hogy

$$\underline{\underline{A}}^{-1}\bar{\underline{a}}=\underline{a}, \quad \text{ha} \quad \underline{\underline{A}}\underline{a}=\bar{\underline{a}}. \qquad 12.2\text{-}1$$

Abban az esetben, ha az $\bar{\underline{a}}$, $\bar{\underline{b}}$, $\bar{\underline{c}}$ vektorok sem komplanárisak, tehát

$$(\bar{\underline{a}}, \bar{\underline{b}}, \bar{\underline{c}}) \neq 0, \qquad 12.2\text{-}2$$

akkor tetszőleges $\bar{\underline{d}}$ vektor előállítható

$$\bar{\underline{d}}=\alpha\bar{\underline{a}}+\beta\bar{\underline{b}}+\gamma\bar{\underline{c}} \qquad 12.2\text{-}3$$

alakban. Ebből leolvasható, hogy az $\underline{\underline{A}}^{-1}$ operátor $\bar{\underline{d}}$-re való hatásának eredményét egyértelműen meghatározhatjuk, ha ismerjük az $\bar{\underline{a}}$, $\bar{\underline{b}}$, $\bar{\underline{c}}$ vektorok transzformáltjait, vagyis ha ismerjük az

$$\underline{a}=\underline{\underline{A}}^{-1}\bar{\underline{a}}, \quad \underline{b}=\underline{\underline{A}}^{-1}\bar{\underline{b}}, \quad \underline{c}=\underline{\underline{A}}^{-1}\bar{\underline{c}} \qquad 12.2\text{-}4$$

vektorokat.

Amennyiben tehát sem az $\underline{a}$, $\underline{b}$, $\underline{c}$, sem pedig az $\bar{\underline{a}}$, $\bar{\underline{b}}$, $\bar{\underline{c}}$ vektorcsoport nem komplanáris, ezek a vektorok mind az $\underline{\underline{\mathbf{A}}}$, mind az $\underline{\underline{\mathbf{A}}}^{-1}$ operátort egyértelműen definiálják.

Mind az $\underline{\underline{\mathbf{A}}}$, mind az $\underline{\underline{\mathbf{A}}}^{-1}$ operátor homogén lineáris transzformációt definiál, így $\underline{\underline{\mathbf{A}}}^{-1}$ is tenzor.

Egy tenzort nem elfajulónak tekintünk, ha nem komplanáris vektorokat nem komplanáris vektorokra képez le, tehát ha 12.1-2 érvényessége esetén 12.2-2 is fennáll.

A fentiekből látható, hogy a nem elfajuló $\underline{\underline{\mathbf{A}}}$ tenzorok $\underline{\underline{\mathbf{A}}}^{-1}$ inverze vagy reciproka létezik és egyértelműen meghatározható. Továbbá, ha $\underline{\underline{\mathbf{A}}}$ nem elfajuló, akkor $\underline{\underline{\mathbf{A}}}^{-1}$ sem az.

12.3. Műveletek tenzorokkal

12.3.1. Két tenzor szorzata

Legyen

$$\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}^{(k)}=\underline{\bar{\mathbf{a}}}^{(k)} \quad (k=1,2,3) \tag{12.3-1}$$

és

$$\underline{\underline{\mathbf{B}}}\underline{\bar{\mathbf{a}}}^{(k)}=\underline{\bar{\bar{\mathbf{a}}}}^{(k)} \quad (k=1,2,3). \tag{12.3-2}$$

12.3-1-et 12.3-2-be beírva, a

$$\underline{\underline{\mathbf{B}}}(\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}^{(k)})=\underline{\bar{\bar{\mathbf{a}}}}^{(k)} \quad (k=1,2,3) \tag{12.3-3}$$

egyenletet kapjuk. 12.3-3 az $\underline{\mathbf{a}}^{(k)}$ vektort az $\underline{\bar{\bar{\mathbf{a}}}}^{(k)}$ vektorra képezi le. Ezt az átmenetet egyetlen transzformációnak is tekinthetjük. Jelöljük ezt a transzformációt $\underline{\underline{\mathbf{C}}}$-vel. Tehát

$$\underline{\underline{\mathbf{C}}}\underline{\mathbf{a}}^{(k)}=\underline{\bar{\bar{\mathbf{a}}}}^{(k)}. \tag{12.3-4}$$

Amennyiben

$$(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(3)})\neq 0, \quad (\underline{\bar{\mathbf{a}}}^{(1)}, \underline{\bar{\mathbf{a}}}^{(2)}, \underline{\bar{\mathbf{a}}}^{(3)})\neq 0 \tag{12.3-5}$$

és

$$(\underline{\bar{\bar{\mathbf{a}}}}^{(1)}, \underline{\bar{\bar{\mathbf{a}}}}^{(2)}, \underline{\bar{\bar{\mathbf{a}}}}^{(3)})\neq 0,$$

akkor $\underline{\underline{\mathbf{A}}}$ és $\underline{\underline{\mathbf{B}}}$ nem elfajuló tenzorok. Ez esetben $\underline{\underline{\mathbf{C}}}$ is egy homogén lineáris vektortranszformációt jelent. 12.3-5 miatt $\underline{\underline{\mathbf{C}}}$ sem elfajuló tenzor.

Bevezethetjük a

$$\underline{\underline{\mathbf{C}}}=\underline{\underline{\mathbf{B}}}\underline{\underline{\mathbf{A}}} \tag{12.3-6}$$

jelölést, hiszen a $\underline{\underline{\mathbf{C}}}$ tenzor alkalmazása ugyanazt az eredményt adja, mint az $\underline{\underline{\mathbf{A}}}$ és $\underline{\underline{\mathbf{B}}}$ tenzorok egymás utáni alkalmazása; $\underline{\underline{\mathbf{C}}}$-t a $\underline{\underline{\mathbf{B}}}$ és $\underline{\underline{\mathbf{A}}}$ tenzorok szorzatának nevezzük. A szorzási művelet nem kommutatív!

Különleges operátor az $\underline{\mathbf{E}}$ egységoperátor. Az egységoperátor minden vektort önmagára képez le, tehát

$$\underline{\mathbf{E}}\underline{\mathbf{a}}=\underline{\mathbf{a}}. \qquad 12.3\text{-}7$$

Egyszerűen belátható, hogy a tenzorszorzás asszociatív művelet, ezt azonban az olvasóra bízzuk.

12.3.2. Tenzorok lineáris kombinációja

A linearitásból kiindulva, bevezethetjük tenzorok lineáris kombinációját. Legyen

$$\underline{\mathbf{A}}\underline{\mathbf{a}}=\bar{\underline{\mathbf{a}}} \quad \text{és} \quad \underline{\mathbf{B}}\underline{\mathbf{a}}=\bar{\bar{\underline{\mathbf{a}}}},$$

és vezessük be a

$$\underline{\mathbf{C}}=\lambda\underline{\mathbf{A}}+\mu\underline{\mathbf{B}} \qquad 12.3\text{-}8$$

tenzort úgy, hogy tetszőleges $\underline{\mathbf{a}}$-ra

$$\underline{\mathbf{C}}\underline{\mathbf{a}}=\lambda\bar{\underline{\mathbf{a}}}+\mu\bar{\bar{\underline{\mathbf{a}}}}$$

legyen.

Amennyiben $\mu=0$, akkor

$$\underline{\mathbf{C}}=\lambda\underline{\mathbf{A}} \quad \text{és} \quad \underline{\mathbf{C}}\underline{\mathbf{a}}=\lambda\underline{\mathbf{A}}\underline{\mathbf{a}}=\lambda\bar{\underline{\mathbf{a}}}, \qquad 12.3\text{-}9$$

továbbá, ha $\lambda=\mu=1$, akkor

$$\underline{\mathbf{C}}=\underline{\mathbf{A}}+\underline{\mathbf{B}} \quad \text{és} \quad \underline{\mathbf{C}}\underline{\mathbf{a}}=\bar{\underline{\mathbf{a}}}+\bar{\bar{\underline{\mathbf{a}}}}. \qquad 12.3\text{-}10$$

12.3-9 tulajdonképpen a tenzor és skalár szorzatának, 12.3-10 pedig két tenzor összegének definíciója.

12.4. Tenzorok reprezentációja

Megállapítottuk, hogy egy $\underline{\mathbf{A}}$ tenzort egyértelműen jellemezhetünk, ha megadjuk, hogy három nem komplanáris vektort mely vektorokra képez le.

Válasszuk a három nem egy síkba eső vektorként egy $\mathcal{K}$ ortogonális koordináta-rendszer $\underline{\mathbf{e}}^{(k)}$ $(k=1, 2, 3)$ alapvektorait, és adjuk meg az

$$\underline{\mathbf{A}}\underline{\mathbf{e}}^{(k)}=\underline{\mathbf{f}}^{(k)} \quad (k=1, 2, 3) \qquad 12.4\text{-}1$$

vektorokat.

Szorozzuk meg 12.4-1-et $\underline{\mathbf{e}}^{(l)}$-lel, azt kapjuk, hogy

$$(\underline{\mathbf{A}}\underline{\mathbf{e}}^{(k)})\underline{\mathbf{e}}^{(l)}=\underline{\mathbf{e}}^{(l)}\underline{\mathbf{f}}^{(k)}. \qquad 12.4\text{-}2$$

Jelöljük a 12.4-2 jobb oldalán álló skaláris szorzatok értékét

$$\underline{e}^{(l)}\underline{f}^{(k)} = A_{lk}\text{-val.} \qquad 12.4\text{-}3$$

Az A_{lk} értékek egy 3×3-as mátrixba rendezhetők. Ezt az **A** mátrixot az $\underline{A}$ tenzor $\mathcal{K}$-beli reprezentációjának nevezzük, tehát

$$\mathcal{K}(\underline{A}) = \mathbf{A}. \qquad 12.4\text{-}4$$

12.4-4-et azért tekintjük az $\underline{A}$ tenzor reprezentációjának, mert **A** segítségével az

$$\underline{\bar{a}} = \underline{A}\,\underline{a} \qquad 12.4\text{-}5$$

transzformáció leírható.

12.4-5 helyett felírhatjuk, hogy

$$\underline{\bar{a}} = \underline{A} \sum a_k \underline{e}^{(k)}. \qquad 12.4\text{-}6$$

Minthogy azonban $\underline{A}$ homogén lineáris vektortranszformáció, 12.4-2 segítségével 12.4-6 az

$$\underline{\bar{a}} = \sum a_k \underline{f}^{(k)} \qquad 12.4\text{-}7$$

formába írható át. 12.4-7-et balról $\underline{e}^{(l)}$-lel beszorozva, az

$$\bar{a}_l = \sum_k A_{lk} a_k \qquad 12.4\text{-}8$$

formulához jutunk. 12.4-8 mátrixjelöléssel az

$$\bar{\mathbf{a}} = \mathbf{A}\mathbf{a} \qquad 12.4\text{-}9$$

alakot ölti. 12.4-9 megmutatja, hogy a $\mathcal{K}$ reprezentációban hogyan adódnak $\underline{\bar{a}}$ koordinátái **A** és **a** komponensei segítségével. Ezért 12.4-9 a 12.4-5 összefüggés $\mathcal{K}$-beli reprezentációja.

12.4.1. Néhány tenzor mátrixreprezentációja

12.4-9-ből következik, hogy az $\underline{E}$ tenzor $\mathcal{K}(\underline{E})$ reprezentációja:

$$\mathcal{K}(\underline{E}) = \mathbf{E},$$

ahol **E** az egységmátrix.

Ugyancsak egyszerűen megállapíthatjuk annak a tenzornak a reprezentációját, amelyet az

$$\underline{A}\,\underline{x} = \underline{a}(\underline{b}\,\underline{x}) \qquad 12.4\text{-}10$$

összefüggéssel értelmezünk. 12.4-10-ben $\underline{a}$ és $\underline{b}$ rögzített, $\underline{x}$ pedig tetszőleges vektor. Az $\underline{A}$ tenzort tehát az $\underline{a}$ és $\underline{b}$ vektorok jellemzik. 12.4-2 és 12.4-3 szerint

$$A_{lk} = (\underline{e}^{(l)}\underline{a})(\underline{b}\,\underline{e}^{(k)}). \qquad 12.4\text{-}11$$

Az $\underline{e}^{(k)}$ alapvektorokkal megadott koordináta-rendszerben ez azt jelenti, hogy

$$A_{ik} = a_i b_k \qquad 12.4\text{-}12$$

12.4-12 viszont éppen az **a** és **b** egydimenziós mátrixok szorzata. A 12.4-10 által definiált $\underline{\underline{\mathbf{A}}}$ tenzor tetszőleges koordináta-rendszerben vett reprezentációját az **a** és **b** mátrixok diadikus szorzata adja meg. Ezért ezt a tenzort az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok diadikus szorzatának nevezzük, és

$$\underline{\underline{\mathbf{A}}} = \underline{\mathbf{a}} \circ \underline{\mathbf{b}} \qquad 12.4\text{-}13$$

-vel jelöljük.

12.4.2. A transzponált tenzor

Az operátorok és vektorok

$$\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}} = \underline{\mathbf{y}}$$

szorzási művelete azt fejezi ki, hogy az $\underline{\underline{\mathbf{A}}}$ operátort alkalmazzuk $\underline{\mathbf{a}}$ vektorra. Ebben az esetben nincs értelme tehát jobbról, ill. balról való szorzásról beszélni. Amennyiben az eredményül kapott vektort egy $\underline{\mathbf{b}}$ vektorral skalárisan szorozzuk, akkor egy

$$(\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}})\underline{\mathbf{b}} = \underline{\mathbf{b}}(\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}) \qquad 12.4\text{-}14$$

számértéket kapunk. Ez a számérték általában nem egyezik meg az $(\underline{\underline{\mathbf{A}}}\underline{\mathbf{b}})\underline{\mathbf{a}}$ értékkel, tehát

$$(\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}})\underline{\mathbf{b}} \neq (\underline{\underline{\mathbf{A}}}\underline{\mathbf{b}})\underline{\mathbf{a}}.$$

Elképzelhető azonban, hogy adott $\underline{\underline{\mathbf{A}}}$ tenzorhoz tetszőleges $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok esetén létezik olyan tenzor, amelyre teljesül az

$$\underline{\mathbf{a}}(\underline{\underline{\mathbf{A}}}\underline{\mathbf{b}}) = \underline{\mathbf{b}}(\underline{\underline{\mathbf{B}}}\underline{\mathbf{a}}) \qquad 12.4\text{-}15$$

összefüggés. A 12.4-15-nek eleget tevő $\underline{\underline{\mathbf{B}}}$ tenzort az $\underline{\underline{\mathbf{A}}}$ tenzor transzponált operátorának nevezzük, és $\underline{\underline{\mathbf{B}}} = \widetilde{\underline{\underline{\mathbf{A}}}}$-val jelöljük. Egyszerűen belátható, hogy a transzponált operátor mindig létezik. A 12.4-15 egyenletet tetszőleges reprezentációban felírva adódik ugyanis, hogy a

$$\mathcal{K}(\underline{\underline{\mathbf{A}}}) = \mathbf{A} \qquad 12.4\text{-}16$$

mátrix $\widetilde{\mathbf{A}}$ transzponáltjára teljesül az

$$\mathbf{aAb} = \mathbf{bAa} \qquad 12.4\text{-}17$$

összefüggés. A $\mathbf{B} = \widetilde{\mathbf{A}}$ választással tehát tetszőleges reprezentációban meghatározhatjuk a keresett $\underline{\underline{\mathbf{B}}}$ tenzor

$$\mathcal{K}(\underline{\underline{\mathbf{B}}}) = \mathbf{B} = \widetilde{\mathbf{A}}$$

reprezentációját.

A mátrixreprezentáció indokolja a 12.6-15-nek megfelelő $\underline{\underline{\mathbf{B}}}$ tenzor elneveját és a $\underline{\mathbf{B}} = \tilde{\underline{\mathbf{A}}}$ jelölést is. A tenzorok tetszőleges mátrixreprezentációjára fennálló asszociatív törvényt formálisan kiterjeszthetjük az operátorokra vonatkozó 12.4-15 formulára. Definíció szerint megköveteljük tehát az

$$\underline{\mathbf{a}}(\underline{\mathbf{A}}\underline{\mathbf{b}}) = (\underline{\mathbf{a}}\underline{\mathbf{A}})\underline{\mathbf{b}} = \underline{\mathbf{a}}\underline{\mathbf{A}}\underline{\mathbf{b}} \qquad 12.4\text{-}18$$

egyenlőség fennállását. 12.4-18 segítségével 12.4-15 az

$$(\underline{\mathbf{a}}\underline{\mathbf{A}})\underline{\mathbf{b}} = \underline{\mathbf{b}}(\underline{\mathbf{B}}\underline{\mathbf{a}}) \qquad 12.4\text{-}19$$

alakban írható fel. Mivel 12.4-19-nek tetszőleges $\underline{\mathbf{b}}$ vektorra teljesülnie kell, 12.4-19-ből következik, hogy

$$\underline{\mathbf{a}}\underline{\mathbf{A}} = \underline{\mathbf{B}}\underline{\mathbf{a}}. \qquad 12.4\text{-}20$$

12.4-20 az $\underline{\mathbf{a}}$ vektor és $\underline{\mathbf{A}}$ tenzor balról vett szorzatát definiálja.

A balról és jobbról vett szorzás bevezetése nem szükségszerű, és a továbbiakban ezt a megkülönböztetést ritkán használjuk.

Speciális esetben egy $\underline{\mathbf{A}}$ tenzor és transzponáltja azonos lehet, tehát

$$\underline{\mathbf{A}} = \tilde{\underline{\mathbf{A}}}. \qquad 12.4\text{-}21$$

Ezeket az operátorokat szimmetrikus operátoroknak nevezzük. A szimmetrikus operátorokra jellemző, hogy tetszőleges $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ vektorokra

$$\underline{\mathbf{a}}\underline{\mathbf{A}}\underline{\mathbf{b}} = \underline{\mathbf{b}}\underline{\mathbf{A}}\underline{\mathbf{a}}.$$

Azokat az operátorokat, amelyekre

$$\underline{\mathbf{A}} = -\tilde{\underline{\mathbf{A}}}, \qquad 12.4\text{-}22$$

antiszimmetrikus operátoroknak nevezzük.

12.5. Tenzorműveletek koordinátareprezentációja

A 12.4-6 reprezentáció segítségével összetett tenzorműveleteket is reprezentálhatunk.

Legyen

$$\bar{\underline{\mathbf{a}}} = \underline{\mathbf{A}}\underline{\mathbf{a}} \quad \text{és} \quad \bar{\bar{\underline{\mathbf{a}}}} = \underline{\mathbf{B}}\bar{\underline{\mathbf{a}}} = (\underline{\mathbf{B}}\underline{\mathbf{A}})\underline{\mathbf{a}}, \qquad 12.5\text{-}1$$

tehát

$$\bar{\bar{\underline{\mathbf{a}}}} = \underline{\mathbf{C}}\underline{\mathbf{a}}, \quad \text{ahol} \quad \underline{\mathbf{C}} = (\underline{\mathbf{B}}\underline{\mathbf{A}}). \qquad 12.5\text{-}2$$

A $\mathscr{X}$ reprezentációban 12.5-1-nek az

$$\bar{\mathbf{a}} = \mathbf{A}\mathbf{a} \quad \text{és} \quad \bar{\bar{\mathbf{a}}} = \mathbf{B}\bar{\mathbf{a}} = (\mathbf{B}\mathbf{A})\mathbf{a},$$

12.5-2-nek pedig az

$$\bar{\bar{\mathbf{a}}} = \mathbf{C}\mathbf{a} \quad (\text{ahol} \quad \mathbf{C} = \mathbf{B}\mathbf{A})$$

összefüggések felelnek meg. Ezekben a formulákban a szorzások mátrixszorzást jelentenek, tehát $C_{kl}=\sum_m B_{km}A_{ml}$.

Hasonló módon legyen

$$\underline{\bar{\mathbf{a}}}=\underline{\mathbf{A}}\underline{\mathbf{a}} \quad \text{és} \quad \underline{\bar{\bar{\mathbf{a}}}}=\underline{\mathbf{B}}\underline{\mathbf{a}},$$

valamint

$$\underline{\mathbf{C}}\underline{\mathbf{a}}=\lambda\underline{\bar{\mathbf{a}}}+\mu\underline{\bar{\bar{\mathbf{a}}}},$$

ahol

$$\underline{\mathbf{C}}=\lambda\underline{\mathbf{A}}+\mu\underline{\mathbf{B}}.$$

Ezen összefüggések $\mathcal{K}$-beli megfelelői:

$$\bar{\mathbf{a}}=\mathbf{A}\mathbf{a} \quad \text{és} \quad \bar{\bar{\mathbf{a}}}=\mathbf{B}\mathbf{a},$$

valamint ebből következően

$$\mathbf{C}\mathbf{a}=\lambda\bar{\mathbf{a}}+\mu\bar{\bar{\mathbf{a}}},$$

ahol

$$\mathbf{C}=\lambda\mathbf{A}+\mu\mathbf{B}.$$

A tenzoriális összefüggéseknek ugyanis tetszőleges reprezentációban teljesülniük kell.

Összefoglalva a fentieket megállapíthatjuk, hogy a tenzorokat mátrixokkal, a tenzorműveleteket pedig mátrixműveletekkel reprezentálhatjuk.

12.6. Összefüggés a tenzorok reprezentációi között

Egy $\underline{\mathbf{A}}$ tenzor fizikai mennyiséget jellemez, így nem függ attól, hogy milyen koordináta-rendszerben reprezentáljuk.

Szükséges azonban annak megállapítása, hogy az $\underline{\mathbf{A}}$ tenzor különböző koordinátareprezentációi között milyen összefüggés van.

Legyen $\mathcal{K}$ és $\mathcal{K}'$ két ortogonális koordináta-rendszer, amelynek alapvektorait a $\underline{\mathbf{P}}$ forgatás viszi át egymásba, tehát

$$\underline{\mathbf{e}}^{(k)'}=\underline{\mathbf{P}}\underline{\mathbf{e}}^{(k)}. \qquad 12.6\text{-}1$$

A $\underline{\mathbf{P}}$ operátor $\mathcal{K}$-beli reprezentációjának elemeit 12.4-3 szerint (ha $\underline{\mathbf{f}}^{(k)}=\underline{\mathbf{e}}^{(k)'}$-t írunk), az

$$\underline{\mathbf{e}}^{(l)}\underline{\mathbf{e}}^{(k)'}=P_{lk} \qquad 12.6\text{-}2$$

formula adja.

Ha a transzformációt a P_{lk} mátrixelemekkel adjuk meg, akkor 12.6-2 felhasználásával az $\underline{\mathbf{e}}^{(k)'}$ vektorok kifejezhetők a P_{lk} mátrixelemek segítségével. $\underline{\mathbf{e}}^{(k)'}$ ugyanis előállítható az $\underline{\mathbf{e}}^{(k)}$ alapvektorok lineáris kombinációjaként, tehát

$$\underline{\mathbf{e}}^{(k)'}=\sum_l P_{lk}\underline{\mathbf{e}}^{(l)}. \qquad 12.6\text{-}3$$

12.6-3-ban a P_{lk} mátrixelemek helyett felhasználhatjuk a

$$P_{lk} = \tilde{P}_{kl} = O_{kl}$$

transzponált mátrixelemeket is. Ezekkel

$$\underline{e}^{(k)'} = \sum_{l} O_{kl} \underline{e}^{(l)}. \qquad 12.6\text{-}4$$

A forgatási operátorok tulajdonságaiból következik, hogy a $\tilde{\mathbf{P}} = \mathbf{O}$ mátrix az

$$\underline{O} = \underline{P}^{-1}$$

forgatást reprezentálja.

Határozzuk meg először az $\underline{a}$ vektor

$$\mathcal{K}(\underline{a}) = \mathbf{a} \quad \text{és} \quad \mathcal{K}'(\underline{a}) = \mathbf{a}'$$

reprezentációi között fennálló összefüggést.

Legyen

$$\underline{a} = \sum_{k=1}^{3} a_k \underline{e}^{(k)} = \sum_{m=1}^{3} a'_m \underline{e}^{(m)'}. \qquad 12.6\text{-}5$$

Szorozzuk be 12.6-5-öt $\underline{e}^{(l)'}$-vel. Az

$$\underline{a}\underline{e}^{(l)'} = \sum_{k=1}^{3} a_k \underline{e}^{(k)} \underline{e}^{(l)'} = a'_l \qquad 12.6\text{-}6$$

összefüggéshez jutunk.

Felhasználva 12.6-2-t és 12.6-4-et, adódik, hogy

$$a'_l = \sum_{k=1} O_{lk} a_k, \qquad 12.6\text{-}7$$

ill.

$$\mathbf{a}' = \mathbf{O}\mathbf{a}.$$

Foglalkozzunk most az $\underline{A}$ tenzor

$$\mathcal{K}(\underline{A}) = \mathbf{A} \quad \text{és} \quad \mathcal{K}'(\underline{A}) = \mathbf{A}'$$

reprezentációi között fennálló összefüggés meghatározásával.

12.4-2 szerint

$$A'_{lk} = \underline{e}^{(l)'} \underline{A} \underline{e}^{(k)'}. \qquad 12.6\text{-}8$$

12.6-8-ba beírva a 12.6-4-et, az

$$A'_{lk} = \sum_{m,\, n=1}^{3} O_{lm} \underline{e}^{(m)} \underline{A} \underline{e}^{(n)} O_{kn}$$

formulához jutunk.

Minthogy

$$\underline{e}^{(m)} \underline{A} \underline{e}^{(n)} = A_{mn},$$

a két reprezentáció között az

$$A'_{lk} = \sum_{m,\, n=1}^{3} O_{lm} A_{mn} O_{kn} \tag{12.6-9}$$

összefüggés áll fenn.

12.6-9-et mátrixalakban az

$$\mathbf{A}' = \mathbf{O}\mathbf{A}\tilde{\mathbf{O}} \tag{12.6-10}$$

formula reprezentálja.

A fenti transzformációs formulák az eddigiekben talált összefüggésekkel összhangban vannak. Ennek illusztrációjaként foglalkozzunk néhány speciális esettel.

Képezzük 12.6-10 determinánsát. Azt kapjuk, hogy

$$\det \mathbf{A} = \det \mathbf{A}', \tag{12.6-11}$$

mert

$$\det \mathbf{O} = \det \tilde{\mathbf{O}} = \pm 1.$$

A transzformáció során tehát az operátort reprezentáló mátrix determinánsa nem változik.

Megmutatjuk még, hogy az $\mathbf{AB} = \mathbf{C}$ szorzat a $\mathcal{K}'$ koordináta-rendszerben az $\mathbf{A}'\mathbf{B}' = \mathbf{C}'$ alakot ölti. Ugyanis

$$\mathbf{C}' = \mathbf{OAB}\tilde{\mathbf{O}} = \mathbf{OA}\tilde{\mathbf{O}}\mathbf{OB}\tilde{\mathbf{O}} = (\mathbf{OA}\tilde{\mathbf{O}})(\mathbf{OB}\tilde{\mathbf{O}}),$$

vagyis

$$\mathbf{C}' = \mathbf{A}'\mathbf{B}'.$$

Itt felhasználtuk, hogy $\tilde{\mathbf{O}}\mathbf{O} = \mathbf{E}$. Látható az is, hogy

$$\mathbf{E}' = \mathbf{OE}\tilde{\mathbf{O}},$$

vagyis az $\underline{\mathbf{E}}$ operátor tetszőleges reprezentációban az

$$\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

alakot ölti. Ebből következik, hogy ha $\mathcal{K}(\underline{\mathbf{A}}) = \mathbf{A}$, akkor $\mathcal{K}(\underline{\mathbf{A}}^{-1}) = \mathbf{A}^{-1}$.

A tenzor definícióját 12.6-9 segítségével is meg tudjuk adni. Ha egy $\underline{\mathbf{A}}$ operátor különböző, $\mathcal{K}$ és $\mathcal{K}'$ koordináta-rendszerbeli $\mathcal{K}(\underline{\mathbf{A}}) = \mathbf{A}$ és $\mathcal{K}'(\underline{\mathbf{A}}) = \mathbf{A}'$ reprezentációi között fennáll az

$$\mathbf{A}' = \mathbf{OA}\tilde{\mathbf{O}}$$

összefüggés, akkor $\underline{\mathbf{A}}$ tenzor. A reprezentációk ismeretében is eldönthetjük tehát, hogy egy operátor tenzor vagy sem. Ennek megfelelően egy mennyiséget akkor nevezünk tenzornak, ha az szabály szerint transzformálódik.

Mi a tenzort inkább a fejezet kezdetén kifejtett fizikai meggondolások alapján definiáljuk. Ezekből a megfontolásokból szükségszerűen jutottunk el a 12.6-12 transzformációs szabályhoz.

A 12.6-12 szabály szerint transzformálódó mennyiségeket kovariáns mennyiségeknek is nevezik.

Ez esetben pontosabbnak tartjuk, ha úgy fogalmazunk, hogy **a** és **A** az $\underline{a}$ és $\underline{A}$ mennyiségek kovariáns reprezentációi.

Tetszőleges fizikai mennyiséget akkor nevezünk kovariánsnak, ha az adott mennyiségnek létezik kovariáns reprezentációja. A kovariáns reprezentáció létezése egyáltalán nem triviális. Ezekkel a kérdésekkel a következőkben még részletesebben foglalkozunk.

12.7. Alkalmazások

12.7.1. A tehetetlenségi tenzor

A következőkben a tenzorfogalom alkalmazásaként merev testek bizonyos dinamikai jellemzőit tárgyaljuk. Vizsgáljunk egy merev testet, amely $m^{(i)}$ tömegű ($i = 1, 2, \ldots$) diszkrét tömegpontokból áll. Az egyes pontok az $\underline{\mathbf{X}}^{(i)}$ helyvektorokkal jellemezhetők. A tetszőlegesen választott origójú $\underline{\mathbf{X}}^{(i)}$ koordinátákról azonban térjünk át az ún. súlyponti koordinátákra. Egy test súlypontját definíció szerint az

$$\underline{\mathbf{X}}^{(S)} = \frac{\sum\limits_i m^{(i)} \underline{\mathbf{X}}^{(i)}}{\sum\limits_i m^{(i)}} \qquad 12.7\text{-}1$$

összefüggés határozza meg, ahol $\underline{\mathbf{X}}^{(i)}$ végigfut a pontrendszer minden pontján. Célszerű, ha az $\underline{\mathbf{X}}^{(i)}$ koordináták helyett az

$$\underline{\mathbf{x}}^{(i)} = \underline{\mathbf{X}}^{(i)} - \underline{\mathbf{X}}^{(s)} \qquad 12.7\text{-}2$$

koordinátákat használjuk, azaz, ha az origót a test súlypontjába helyezzük át. Ezekre a koordinátákra 12.7-1 alapján fennáll, hogy

$$\sum_i \underline{\mathbf{x}}^{(i)} = \underline{\mathbf{0}}. \qquad 12.7\text{-}3$$

Határozzuk meg az $m^{(i)}$ tömegpontokból álló test mozgási energiáját. Tegyük fel először, hogy a test súlypontja rögzített, és a test a rögzített súlypont körül $\underline{\omega}$ szögsebességgel forog. Ez esetben az i-edik pont sebessége 5.6-22 szerint

$$\underline{\mathbf{v}}^{(i)}=\underline{\omega}\times\underline{\mathbf{x}}^{(i)}. \qquad 12.7\text{-}4$$

Az i-edik pont mozgási energiája

$$\frac{1}{2}m^{(i)}\underline{\mathbf{v}}^{(i)2}=\frac{1}{2}m^{(i)}(\underline{\omega}\times\underline{\mathbf{x}}^{(i)})(\underline{\omega}\times\underline{\mathbf{x}}^{(i)}). \qquad 12.7\text{-}5$$

A 4.8. fejezet vektoriális négyesszorzatra vonatkozó 4.8-8 összefüggése szerint ez átírható az

$$\frac{1}{2}m^{(i)}\underline{\mathbf{v}}^{(i)2}=\frac{1}{2}m^{(i)}\left[\underline{\mathbf{x}}^{(i)2}\underline{\omega}^{2}-(\underline{\mathbf{x}}^{(i)}\underline{\omega})^{2}\right] \qquad 12.7\text{-}6$$

alakra. 12.7-6-ot az $\underline{\omega}$ vektor kiemelése céljából alakítsuk át a diadikus szorzat definíciója és az egységtenzor segítségével az

$$\frac{1}{2}m^{(i)}\underline{\mathbf{v}}^{(i)2}=\frac{1}{2}m^{(i)}\underline{\omega}\left[\underline{\mathbf{E}}\underline{\mathbf{x}}^{(i)2}-\underline{\mathbf{x}}^{(i)}\circ\underline{\mathbf{x}}^{(i)}\right]\underline{\omega} \qquad 12.7\text{-}7$$

formára.

A teljes mozgási energia ennek alapján az

$$E_k=\frac{1}{2}\sum_i m^{(i)}\underline{\mathbf{v}}^{(i)2}=\frac{1}{2}\underline{\omega}\underline{\mathbf{T}}\underline{\omega} \qquad 12.7\text{-}8$$

alakban írható fel, ahol

$$\underline{\mathbf{T}}=\sum_i m^{(i)}\left[\underline{\mathbf{E}}\underline{\mathbf{x}}^{(i)2}-\underline{\mathbf{x}}^{(i)}\circ\underline{\mathbf{x}}^{(i)}\right]. \qquad 12.7\text{-}9$$

$\underline{\mathbf{T}}$-t a test tehetetlenségi tenzorának nevezzük.

A jobb áttekinthetőség kedvéért fejtsük ki a $\underline{\mathbf{T}}$ tenzort egy $\mathcal{K}$ koordinátarendszerben, ahol $\mathbf{T}=\mathcal{K}(\underline{\mathbf{T}})$. Azt kapjuk, hogy

$$\left.\begin{aligned}
T_{11}&=\sum_i m^{(i)}(x_2^{(i)2}+x_3^{(i)2}),\\
T_{22}&=\sum_i m^{(i)}(x_3^{(i)2}+x_1^{(i)2}),\\
T_{33}&=\sum_i m^{(i)}(x_1^{(i)2}+x_2^{(i)2}),
\end{aligned}\right\} \qquad 12.7\text{-}10$$

valamint

$$\left.\begin{aligned} T_{21}=T_{12}&=-\sum_i m^{(i)}x_1^{(i)}x_2^{(i)},\\ T_{32}=T_{23}&=-\sum_i m^{(i)}x_3^{(i)}x_2^{(i)},\\ T_{31}=T_{13}&=-\sum_i m^{(i)}x_3^{(i)}x_1^{(i)}.\end{aligned}\right\} \qquad 12.7\text{-}11$$

A 12.7-10 formulából leolvasható, hogy a T_{11}, T_{22}, T_{33} értékek a merev test tehetetlenségi nyomatékai a $\mathcal{K}$ koordináta-rendszer tengelyeire vonatkozóan. A 12.7-11 értékeket *deviációs nyomatékoknak* nevezzük, fizikai jelentésük magyarázatára most nem térünk ki.

A 12.7-8 formula emlékeztet a rögzített tengely körül forgó merev test forgási energiájára vonatkozó

$$E_{\mathrm{kin}}=\frac{1}{2}\,\Theta\omega^2 \qquad 12.7\text{-}12$$

formulára, ezért a $\underline{\mathbf{T}}$ tehetetlenségi tenzort a tehetetlenségi nyomaték általánosításának tekintjük.

Merev test kinetikus energiája

Ha a merev test súlypontját nem rögzítjük, és a súlypont $\underline{\mathbf{v}}^{(0)}$ sebességgel mozog, akkor az i-edik pont sebessége

$$\underline{\mathbf{V}}^{(i)}=\underline{\mathbf{v}}^{(0)}+(\underline{\boldsymbol{\omega}}\times\underline{\mathbf{x}}^{(i)})=\underline{\mathbf{v}}^{(0)}+\underline{\mathbf{v}}^{(i)}. \qquad 12.7\text{-}13$$

A kinetikus energiát ekkor

$$E_{\mathrm{kin}}=\frac{1}{2}\sum_i m^{(i)}\underline{\mathbf{V}}^{(i)2}=\frac{1}{2}\,\underline{\mathbf{v}}^{(0)2}\sum_i m^{(i)}+$$

$$+\frac{1}{2}\,\underline{\boldsymbol{\omega}}\underline{\mathbf{T}}\underline{\boldsymbol{\omega}}+\underline{\mathbf{v}}^{(0)}\sum_i m^{(i)}(\underline{\boldsymbol{\omega}}\times\underline{\mathbf{x}}^{(i)})$$

alakban írhatjuk fel. Mivel $\underline{\mathbf{x}}^{(i)}$ súlyponti koordináta, 12.7-3 értelmében

$$\sum m^{(i)}(\underline{\boldsymbol{\omega}}\times\underline{\mathbf{x}}^{(i)})=\underline{\boldsymbol{\omega}}\times\sum_i m^{(i)}\underline{\mathbf{x}}^{(i)}=0.$$

A kinetikus energia tehát

$$E_{\mathrm{kin}}=\frac{1}{2}\,m(\underline{\mathbf{v}}^{(0)2})+\frac{1}{2}\,\underline{\boldsymbol{\omega}}\underline{\mathbf{T}}\underline{\boldsymbol{\omega}} \qquad 12.7\text{-}14$$

alakban írható fel, ahol $m=\sum m^{(i)}$ a test össztömege. 12.7-14 szerint egy merev test kinetikus energiája a súlypont körüli forgásból származó forgási energiából és a súlypont haladása miatt fellépő transzlációs mozgási energiából tevődik össze.

12.7.2. A merev test impulzusmomentuma

A forgó test mozgásmennyiségét az

$$\underline{\mathbf{I}} = \sum_i m^{(i)}\underline{\mathbf{v}}^{(i)} = m\underline{\mathbf{v}}^{(0)} + \sum_i m^{(i)}(\underline{\boldsymbol{\omega}} \times \underline{\mathbf{x}}^{(i)}) \qquad 12.7\text{-}15$$

formula adja meg. A súlypont kezdőpontú koordináta-rendszerben 12.7-15 második tagja eltűnik, tehát a test teljes mozgásmennyisége a transzlációs mozgásból származik.

A forgó testnek a súlypontjára vonatkozó *impulzusmomentuma* az

$$\underline{\mathbf{N}} = \sum_i \underline{\mathbf{N}}^{(i)} = \sum_i m^{(i)}(\underline{\mathbf{x}}^{(i)} \times \underline{\mathbf{v}}^{(i)}) \qquad 12.7\text{-}16$$

összefüggéssel határozható meg. 12.7-4 felhasználásával ez az

$$\underline{\mathbf{N}} = \sum_i m^{(i)}[\underline{\mathbf{x}}^{(i)} \times (\underline{\boldsymbol{\omega}} \times \underline{\mathbf{x}}^{(i)})] \qquad 12.7\text{-}17$$

alakra hozható. A vektoriális hármasszorzatot a kifejtési tétel segítségével átírva:

$$\underline{\mathbf{N}} = \sum_i m^{(i)}[\underline{\boldsymbol{\omega}}\underline{\mathbf{x}}^{(i)^2} - \underline{\mathbf{x}}^{(i)}(\underline{\boldsymbol{\omega}}\underline{\mathbf{x}}^{(i)})]. \qquad 12.7\text{-}18$$

Minthogy

$$\underline{\mathbf{x}}^{(i)}(\underline{\boldsymbol{\omega}}\underline{\mathbf{x}}^{(i)}) = \underline{\boldsymbol{\omega}}(\underline{\mathbf{x}}^{(i)} \circ \underline{\mathbf{x}}^{(i)}),$$

azt kapjuk, hogy

$$\underline{\mathbf{N}} = \underline{\boldsymbol{\omega}}\underline{\underline{\mathbf{T}}} = \underline{\underline{\mathbf{T}}}\underline{\boldsymbol{\omega}}. \qquad 12.7\text{-}19$$

12.7-19 és 12.7-14 összehasonlításából azt kapjuk, hogy

$$\frac{1}{2}\underline{\boldsymbol{\omega}}\underline{\mathbf{N}} = E_{\text{kin}}. \qquad 12.7\text{-}20$$

A merev test impulzusmomentumát tehát a szögsebesség-vektor és tehetetlenségi tenzor szorzata adja. 12.7-19-ből leolvasható, hogy az impulzusmomentum és a szögsebesség-vektor (tehát a pillanatnyi forgástengely) iránya általában nem azonos. Ezzel függ össze a pörgettyűmozgásnál fellépő precessziós jelenség, amelyre a későbbiek során még visszatérünk.

13. A sajátérték-probléma

Felmerül a kérdés, hogy vajon tetszőleges $\underline{\mathbf{A}}$ tenzor esetén léteznek-e olyan $\mathbf{a}$ vektorok, amelyeket az adott tenzor az eredeti vektorral azonos állású vektorokra képez le.

Az előző fejezetben tárgyalt

$$\underline{\mathbf{N}} = \underline{\mathbf{T}}\boldsymbol{\omega}$$

összefüggéssel kapcsolatban ez annak a vizsgálatát jelenti, hogy vajon léteznek-e olyan $\boldsymbol{\omega}^{(s)}$ vektorok, amelyekre a megfelelő $\underline{\mathbf{N}}$ vektorral fennáll az, hogy $\boldsymbol{\omega}$ és $\underline{\mathbf{N}}$ iránya megegyezzen, azaz

$$\underline{\mathbf{N}} = \underline{\mathbf{T}}\boldsymbol{\omega}^{(s)} = \lambda^{(s)}\boldsymbol{\omega}^{(s)}$$

legyen. Tapasztalataink alapján tudjuk, hogy vannak ilyen vektorok. Általában tetszőleges $\underline{\mathbf{A}}$ tenzor esetén kereshetjük azokat az $\underline{\mathbf{x}}^{(s)} \neq \underline{\mathbf{0}}$ vektorokat, amelyekre

$$\underline{\mathbf{A}}\underline{\mathbf{x}}^{(s)} = \lambda^{(s)}\underline{\mathbf{x}}^{(s)}. \qquad 13\text{-}1$$

A 13-1 formulát kielégítő $\underline{\mathbf{x}}^{(s)}$ vektorokat az $\underline{\mathbf{A}}$ tenzor sajátvektorainak, a megfelelő $\lambda^{(s)}$ értékeket pedig sajátértékeknek nevezzük.

13.1. A szekuláris egyenlet

A következőkben egy adott $\underline{\mathbf{A}}$ operátor sajátértékeinek és sajátvektorainak meghatározásával foglalkozunk. Adott $\underline{\mathbf{A}}$ esetén keressük tehát azokat az $\underline{\mathbf{s}}'$ vektorokat, amelyekre

$$\underline{\mathbf{A}}\underline{\mathbf{s}}' = \lambda\underline{\mathbf{s}}'. \qquad 13.1\text{-}1$$

Ez az egyenlet az

$$(\underline{\mathbf{A}} - \lambda\underline{\mathbf{E}})\underline{\mathbf{s}}' = 0 \qquad 13.1\text{-}2$$

formára is átrendezhető, ami egy adott reprezentációban kifejezve, lineáris egyenletrendszer.

Az egyenletrendszernek csak akkor van a triviálistól különböző megoldása, ha

$$\det(\mathbf{A} - \lambda\mathbf{E}) = 0. \qquad 13.1\text{-}3$$

13.1-3 koordinátás alakban az

$$\begin{vmatrix} A_{11}-\lambda & A_{12} & A_{13} \\ A_{21} & A_{22}-\lambda & A_{23} \\ A_{31} & A_{32} & A_{33}-\lambda \end{vmatrix} = 0 \qquad 13.1\text{-}4$$

alakot ölt. A 13.1-4 típusú egyenletet szekuláris egyenletnek nevezzük. Az elnevezés onnan ered, hogy a csillagászatban ilyen egyenleteket használnak a bolygók pályájának lassú, évszázadok alatti (szekuláris) változásainak kiszámítására.

A 13.1-4 determinánst kifejtve, a

$$\lambda^3-(A_{11}+A_{22}+A_{33})\lambda^2+(A_{11}A_{22}-A_{12}A_{21}+A_{22}A_{33}-A_{23}A_{32}+ \\ +A_{11}A_{33}-A_{31}A_{13})\lambda-\det \mathbf{A}=0 \qquad 13.1\text{-}5$$

harmadfokú egyenletet kapjuk. A 13.1-5 egyenlet együtthatóit az adjungált mátrix elemeivel is kifejezhetjük, ugyanis

$$A_{kk}A_{ll}-A_{kl}A_{lk}=A^{(mm)}, \qquad 13.1\text{-}6$$

ahol k, l, m az 1, 2, 3 számok ciklikus permutációja. Ezzel 13.1-5 a

$$\lambda^3-\lambda^2(A_{11}+A_{22}+A_{33})+\lambda(A^{(11)}+A^{(22)}+A^{(33)})-\det \mathbf{A}=0 \qquad 13.1\text{-}7$$

rövidebb formában is felírható, ahol $A^{(kk)}=\det \mathbf{A}^{(kk)}$ $(k=1, 2, 3)$.

Egy mátrix diagonálelemeinek összegét a mátrix spurjának nevezzük és sp **A**-val jelöljük. Ennek felhasználásával a 13.1-7 egyenlet a

$$\lambda^3-\mathrm{sp}\,\mathbf{A}\lambda^2+\mathrm{sp}\,\bar{\mathbf{A}}\lambda-\det \mathbf{A}=0 \qquad 13.1\text{-}8$$

formában is felírható. 13.1-8-ban $\bar{\mathbf{A}}$ az adjungált mátrixot jelenti.

A harmadfokú egyenlet együtthatói és gyökei közötti ismert összefüggések felhasználásával a

$$\begin{aligned} \mathrm{sp}\,\mathbf{A}&=\lambda_1+\lambda_2+\lambda_3, \\ \mathrm{sp}\bar{\mathbf{A}}&=\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3, \\ \det \mathbf{A}&=\lambda_1\lambda_2\lambda_3 \end{aligned} \qquad 13.1\text{-}9$$

formulákhoz juthatunk.

Mivel a sajátértékeket és sajátvektorokat reprezentációtól függetlenül definiáltuk, a 13.1-9 összefüggések is reprezentációtól függetlenül érvényesek. A szekuláris egyenlet együtthatói tehát nem függnek attól a koordináta-rendszertől, amelyben az egyenletet felírtuk.

A 13.1-9 mennyiségeket, mivel reprezentációtól függetlenek, *invariánsoknak* nevezzük.

A tenzorinvariánsokat szokás A_{I}, A_{II}, A_{III}-mal is jelölni. Ezzel a jelöléssel a szekuláris egyenlet a

$$\lambda^3-A_{\mathrm{I}}\lambda^2+A_{\mathrm{II}}\lambda-A_{\mathrm{III}}=0 \qquad 13.1\text{-}10$$

alakban írható fel. 13.1-5-ből leolvasható, hogy

$$A_{\mathrm{I}}=\mathrm{sp}\,\mathbf{A},$$

$$A_{II}=\frac{1}{2}\sum_{i,j,k,l,m}\varepsilon_{ijk}\varepsilon_{lmk}A_{il}A_{jm}$$ 13.1-11

$$A_{III}=\det \mathbf{A}.$$

13.2. Tenzorok hatványai és a hatvány sajátértékei

Legyen

$$\underline{\bar{\mathbf{a}}}=\underline{\mathbf{A}}\underline{\mathbf{a}}$$ 13.2-1

egy homogén lineáris vektortranszformáció. A művelet megismétlésének jelölésére vezessük be az

$$\underline{\bar{\bar{\mathbf{a}}}}=\underline{\mathbf{A}}\underline{\bar{\mathbf{a}}}=\underline{\mathbf{A}}^2\underline{\mathbf{a}}$$ 13.2-2

szimbólumot.

Ha $\underline{\mathbf{s}}$ az $\underline{\mathbf{A}}$ tenzor sajátvektora, akkor

$$\underline{\mathbf{A}}\underline{\mathbf{s}}=\lambda\underline{\mathbf{s}},$$ 13.2-3

ahol λ a megfelelő sajátérték, és

$$\underline{\mathbf{A}}^2\underline{\mathbf{s}}=\underline{\mathbf{A}}\lambda\underline{\mathbf{s}}=\lambda^2\underline{\mathbf{s}}.$$ 13.2-4

Tehát az $\underline{\mathbf{A}}^2$ tenzor sajátértékei $\underline{\mathbf{A}}$ sajátértékeinek négyzetével egyenlők. Folytatva ezt az eljárást, azt kapjuk, hogy az $\underline{\mathbf{A}}^n$ tenzor sajátértékei a λ^n számok. (Itt $\underline{\mathbf{A}}^n$-nel az operáció n-szer egymás után történő alkalmazását jelöltük.) Mint tudjuk, ha $\det \mathbf{A}\neq 0$, akkor létezik az $\underline{\mathbf{A}}^{-1}$ operátor. Szorozzuk 13.2-3 mindkét oldalát $\underline{\mathbf{A}}^{-1}$-gyel. Az

$$\underline{\mathbf{s}}=\lambda\underline{\mathbf{A}}^{-1}\underline{\mathbf{s}}$$

eredmény adódik, ahonnan

$$\underline{\mathbf{A}}^{-1}\underline{\mathbf{s}}=\frac{1}{\lambda}\underline{\mathbf{s}}\quad(\lambda\neq 0).$$ 13.2-5

13.2-5 azt fejezi ki, hogy a reciprok tenzor sajátvektorai megegyeznek az eredeti tenzor sajátvektoraival, sajátértékei azonban az eredeti sajátértékek reciprokai.

Azonnal látható az is, hogy ha az $f(\underline{\mathbf{A}})$ mátrixfüggvényt az

$$f(\underline{\mathbf{A}})=a_n\underline{\mathbf{A}}^n+a_{n-1}\underline{\mathbf{A}}^{n-1}+\ldots+a_0\underline{\mathbf{E}}$$ 13.2-6

mátrixszal definiáljuk, akkor az $f(\underline{\mathbf{A}})$ mátrix sajátértékeit a

$$\Lambda=f(\lambda)=a_n\lambda+a_{n-1}\lambda^{n-1}+\ldots+a_0$$ 13.2-7

függvény adja.

13.3. A sajátértékek és sajátvektorok meghatározása speciális esetekben

Adott $\underline{\underline{A}}$ tenzor sajátértékeit a szekuláris egyenlet segítségével határozhatjuk meg. Az algebra alaptétele szerint (lásd függelék) minden n-ed fokú egyenlet a komplex számsíkon n és csak n gyökkel rendelkezik. Ha egy egyenletnek a z komplex szám gyöke, akkor a z^* komplex konjugált is gyöke az egyenletnek. Következésképpen a harmadfokú valós együtthatós egyenletek legalább egy valós gyökkel rendelkeznek.

Legyen λ a szekuláris egyenlet egy valós gyöke. Ekkor a λ sajátértékhez tartozó $\underline{s}$ sajátvektor a 13.1-2 egyenletből határozható meg. Tehát az

$$\begin{aligned}(A_{11}-\lambda)s_1+A_{12}s_2+A_{13}s_3&=0\\ A_{21}s_1+(A_{22}-\lambda)s_2+A_{23}s_3&=0\\ A_{31}s_1+A_{32}s_2+(A_{33}-\lambda)s_3&=0\end{aligned} \qquad 13.3\text{-}1$$

egyenletrendszer adja az s_1, s_2, s_3 komponenseket. A 13.3-1 egyenletrendszer elfajuló, tehát megoldása nem egyértelmű. Ha az $\underline{\underline{A}}-\lambda\underline{\underline{E}}$ mátrix rangja kettő, akkor az $\underline{s}$ vektor iránya meghatározott, hossza azonban nem. Magasabb elfajulás esetén $\underline{s}$ iránya sem határozott. Amennyiben a szekuláris egyenlet λ' és λ'' megoldásai is valósak, a hozzájuk tartozó $\underline{s}'$ és $\underline{s}''$ sajátvektorok hasonló módon határozhatók meg.

Mielőtt részletesen vizsgálnánk a sajátvektorok egzisztenciáját, néhány speciális problémával foglalkozunk.

13.3.1. A tehetetlenségi tenzor sajátértékei és sajátvektorai

Megállapítható, és a későbbiek során be is bizonyítjuk, hogy a szimmetrikus operátorok három valós sajátértékkel rendelkeznek.

A tehetetlenségi tenzor szimmetrikus tenzor, így sajátértékei valósak. Mivel

$$\frac{1}{2}\underline{\omega}\underline{\underline{T}}\underline{\omega}=E_k \qquad 13.3\text{-}2$$

a rendszer forgási energiája, az

$$\underline{\omega}\underline{\underline{T}}\underline{\omega}>0 \qquad 13.3\text{-}3$$

összefüggés tetszőleges $\underline{\omega}$ esetén fennáll. A 13.3-3-at tetszőleges vektorra kielégítő tenzorokat *pozitív definit tenzoroknak* nevezzük.

Ha $\underline{\omega}$ sajátvektor, akkor

$$\underline{\omega}\underline{\underline{T}}\underline{\omega}=\lambda\underline{\omega}^2>0, \qquad 13.3\text{-}4$$

azaz

$$\lambda \geq 0.$$

A tehetetlenségi tenzor sajátértékei tehát mindig pozitívak.

Egyszerűen belátható, hogy ha $\lambda^{(1)} \neq \lambda^{(2)} \neq \lambda^{(3)}$, akkor az $\underline{\boldsymbol{\omega}}^{(k)}$ $(k=1, 2, 3)$ sajátvektorok ortogonálisak. Legyenek $\underline{\boldsymbol{\omega}}^{(1)}$, ill. $\underline{\boldsymbol{\omega}}^{(2)}$ a $\underline{\mathbf{T}}$ tenzor $\lambda^{(1)}$, $\lambda^{(2)}$ sajátértékeihez tartozó sajátvektorok. Ekkor

$$\underline{\mathbf{T}}\underline{\boldsymbol{\omega}}^{(1)} = \lambda^{(1)}\underline{\boldsymbol{\omega}}^{(1)},$$
$$\underline{\mathbf{T}}\underline{\boldsymbol{\omega}}^{(2)} = \lambda^{(2)}\underline{\boldsymbol{\omega}}^{(2)}.$$

Szorozzuk be az első egyenletet $\underline{\boldsymbol{\omega}}^{(2)}$-vel, a másodikat $\underline{\boldsymbol{\omega}}^{(1)}$-gyel, és vonjuk ki a második egyenletet az elsőből. Mivel $\underline{\mathbf{T}}$ szimmetrikus,

$$0 = \underline{\boldsymbol{\omega}}^{(2)}\underline{\mathbf{T}}\underline{\boldsymbol{\omega}}^{(1)} - \underline{\boldsymbol{\omega}}^{(1)}\underline{\mathbf{T}}\underline{\boldsymbol{\omega}}^{(2)} = (\lambda^{(2)} - \lambda^{(1)})\underline{\boldsymbol{\omega}}^{(1)}\underline{\boldsymbol{\omega}}^{(2)}.$$

Ez azonban csak akkor teljesülhet, ha $\underline{\boldsymbol{\omega}}^{(1)} \perp \underline{\boldsymbol{\omega}}^{(2)}$.

Foglalkozzunk most azzal az esettel, amikor az operátor sajátértékei nem mind különböznek.

Legyen pl.

$$\lambda^{(1)} = \lambda^{(2)} = \lambda.$$

Ekkor a megfelelő sajátvektorokra

$$\underline{\mathbf{T}}\underline{\boldsymbol{\omega}}^{(1)} = \lambda\underline{\boldsymbol{\omega}}^{(1)}, \quad \text{és} \quad \underline{\mathbf{T}}\underline{\boldsymbol{\omega}}^{(2)} = \lambda\underline{\boldsymbol{\omega}}^{(2)}.$$

Ebben az esetben azonban az

$$\underline{\boldsymbol{\omega}} = \alpha\underline{\boldsymbol{\omega}}^{(1)} + \beta\underline{\boldsymbol{\omega}}^{(2)}$$

vektorra is teljesül a

$$\underline{\mathbf{T}}\underline{\boldsymbol{\omega}} = \lambda\underline{\boldsymbol{\omega}}$$

egyenlet, vagyis az $\underline{\boldsymbol{\omega}}^{(1)}$ és $\underline{\boldsymbol{\omega}}^{(2)}$ által meghatározott sík minden vektora $\underline{\mathbf{T}}$ sajátvektora.

Természetesen ezek közül a vektorok közül is kiválaszthatunk ortogonálisakat.

A

$$\lambda^{(1)} = \lambda^{(2)} = \lambda^{(3)} = \lambda$$

eset akkor áll fenn, ha a 13.1-5 szekuláris egyenletnek egyetlen háromszoros gyöke van, tehát

$$(T - \lambda)^3 = 0$$

alakra hozható. Ez azonban csak akkor lehetséges, ha

$$T_{11} = T_{22} = T_{33} = T$$

és

$$T_{12} = T_{23} = T_{31} = 0,$$

vagyis a $\underline{\mathbf{T}}$ tenzor mátrixa tetszőleges koordináta-rendszerben

$$\mathbf{T}=T\mathbf{E}$$

alakú diagonálmátrixszal reprezentálható. Ezek az operátorok azonban egyetlen vektor irányát sem változtatják meg, vagyis a tér minden vektora sajátvektor.

Most is kiválaszthatunk azonban — bár nem egyértelműen — három egymásra merőleges sajátvektort. Általánosan kimondhatjuk tehát, hogy a tehetetlenségi tenzorhoz mindig található három ortogonális sajátvektor.

13.3.2. A forgatási operátor sajátértékei

A forgatási operátorok ortogonális mátrixokkal írhatók le. Az ortogonális mátrixok esetén

$$\mathbf{O}\tilde{\mathbf{O}}=\mathbf{E},$$

tehát $\tilde{\mathbf{O}}^{-1}=\mathbf{O}$, és mivel $\det\mathbf{O}=+1$, ezért

$$\mathbf{O}^{-1}=\bar{\mathbf{O}}=\tilde{\mathbf{O}},$$

vagyis az ortogonális mátrix transzponáltja és adjungáltja azonos. Mivel

$$\mathrm{sp}\,\mathbf{O}=\mathrm{sp}\,\tilde{\mathbf{O}},$$

a

$$\mathrm{sp}\,\mathbf{O}=\mathrm{sp}\,\bar{\mathbf{O}}=a$$

egyenlőség is fennáll. A szekuláris egyenlet tehát a

$$\lambda^3-a\lambda^2+a\lambda-1=0 \qquad 13.3\text{-}5$$

formát ölti. Behelyettesítéssel meggyőződhetünk, hogy a 13.3-5 egyenlet egyik megoldása

$$\lambda_1=1. \qquad 13.3\text{-}6$$

A 13.3-5 egyenlet elosztható a $\lambda-1$ tényezővel. Ennek eredményeként a

$$\lambda^2-(a-1)\lambda+1=0 \qquad 13.3\text{-}7$$

egyenlet adódik. A 13.3-7 egyenletet egyszerűen megoldhatjuk, ha megvizsgáljuk az $\underline{\mathbf{O}}$ tenzor fizikai jelentését. Az ortogonális operátorral — mint már láttuk — merev testek elforgatását írhatjuk le. Ha egy P pont $\underline{\mathbf{r}}$ koordinátavektorát a forgatás egy P^* pont $\underline{\mathbf{r}}^*$ helyvektorába viszi át, akkor

$$\underline{\mathbf{r}}^*=\underline{\mathbf{O}}\underline{\mathbf{r}},$$

ahol a koordináta-rendszert úgy választjuk, hogy a forgatás alatt az origóban fekvő pont nem mozog, tehát

$$\mathbf{r}_0 = \mathbf{O}\mathbf{r}_0.$$

A $\lambda_1 = 1$ sajátértékhez tartozó sajátvektort $\boldsymbol{\omega}_1$-gyel jelölve,

$$\mathbf{O}\boldsymbol{\omega}_1 = \boldsymbol{\omega}_1,$$

tehát ha

$$\mathbf{r} = \alpha\boldsymbol{\omega}_1, \tag{13.3-8}$$

akkor

$$\underline{\mathbf{r}}^* = \underline{\mathbf{r}},$$

vagyis az origón átmenő és $\underline{\boldsymbol{\omega}}_1$ irányú egyenes pontjai az elforgatás során helyükön maradnak.

Ebből látható, hogy egy merev test olyan elmozdulását, amelynek során legalább egy P_0 pont nem változtatja helyét, egy a P_0 ponton átmenő tengely körüli forgással helyettesíthetjük. A tengely egyenletét 13.3-8 adja meg.

Ha az $\underline{\mathbf{O}}$ operátort olyan $\mathcal{K}$ rendszerben reprezentáljuk, amelynek x_1 tengelye megegyezik a forgástengellyel, akkor

$$\mathbf{O} = \mathcal{K}(\underline{\mathbf{O}}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix},$$

ahol φ a tengely körüli elfordulás szöge. A fenti kifejezésből látjuk, hogy

$$a = \operatorname{sp} \mathbf{O} = 1 + 2\cos\varphi, \tag{13.3-9}$$

tehát

$$-1 \leqq a \leqq +3.$$

sp $\mathbf{O}$ számértéke viszont nem függ a reprezentációtól, tehát 13.3-9 általánosan érvényes. 13.3-9-et felhasználva, a 13.3-7 egyenlet a

$$\lambda^2 - 2\cos\varphi\lambda + 1 = 0$$

alakot ölti, ahonnan

$$\lambda_{2,3} = \cos\varphi \pm i\sin\varphi = e^{\pm i\varphi}. \tag{13.3-10}$$

Látjuk tehát, hogy $\underline{\mathbf{O}}$ sajátértékei 1, $e^{i\varphi}$, $e^{-i\varphi}$.

Ennek megfelelően, ha $\varphi \neq 0, k\pi$, akkor csak egy valós főirány létezik. Ez megfelel annak a szemléletes ténynek, hogy a tengelyen fekvő pontok kivételével egy merev test tetszés szerinti pontjának helyzetvektora a forgatás során változtatja irányát.

13.4. Komplex sajátértékek és sajátvektorok

A szekuláris egyenletnek a komplex számsíkon három megoldása van. Tetszőleges komplex sajátértéket a 13.3-1 egyenletrendszerbe behelyettesítve, meghatározhatjuk az s_1, s_2, s_3 számokat. Valós sajátérték esetén ezek a tenzor sajátvektorának komponensei. Komplex sajátérték esetén az s_1, s_2, s_3 számok is komplex értéket vesznek fel.

Formálisan a három komplex számértéket egy komplex komponensű vektor három komponensének tekinthetjük. Kérdéses azonban, hogy egy komplex reprezentációval bíró vektor leírhat-e fizikai mennyiséget.

Legyen

$$\underline{\mathbf{a}} = \alpha_1 + i\beta_1, \quad \alpha_2 + i\beta_2, \quad \alpha_3 + i\beta_3,$$

ahol α_k és β_k valós számok, amelyek egy $\underline{\boldsymbol{\alpha}}$ és egy $\underline{\boldsymbol{\beta}}$ vektor $\mathcal{K}$-beli reprezentációjának tekinthetők. Tehát az α_k és β_k számok az

$$\underline{\mathbf{a}} = \underline{\boldsymbol{\alpha}} + i\underline{\boldsymbol{\beta}}$$

komplex vektor reprezentációjának tekinthetők. Ebben a formulában $\underline{\boldsymbol{\alpha}}$ és $\underline{\boldsymbol{\beta}}$ valós komponensű vektorok. A szokásos számítási szabályok felhasználásával valóban fennáll a

$$\mathcal{K}(\underline{\boldsymbol{\alpha}} + i\underline{\boldsymbol{\beta}}) = \mathcal{K}(\underline{\boldsymbol{\alpha}}) + i\mathcal{K}(\underline{\boldsymbol{\beta}}) = \boldsymbol{\alpha} + i\boldsymbol{\beta},$$

tehát

$$\mathcal{K}(\underline{\mathbf{a}}) = \mathbf{a}.$$

Komplex vektorral írhatunk le tehát olyan fizikai mennyiségeket, amelyeket a valós mennyiségek körében két vektorral jellemezhetnénk.

Példaként megemlítjük, hogy az

$$\underline{\mathbf{a}} = \underline{\mathbf{a}}_0 e^{i\omega t}$$

vektor az

$$\underline{\boldsymbol{\alpha}} = \underline{\mathbf{a}}_0 \cos \omega t \quad \text{és} \quad \underline{\boldsymbol{\beta}} = \underline{\mathbf{a}}_0 \sin \omega t$$

vektorokat adja meg egyszerre. Az $\underline{\boldsymbol{\alpha}}$ és $\underline{\boldsymbol{\beta}}$ vektorok egy forgó vektor komponenseit írják le, hiszen

$$\underline{\boldsymbol{\alpha}}^2 + \underline{\boldsymbol{\beta}}^2 = a_0^2.$$

Anélkül, hogy a komplex vektorokkal és tenzorokkal végzett műveleteket részletesen tárgyalnánk, mégegyszer kiemeljük, hogy a komplex vektorokkal és tenzorokkal olyan fizikai mennyiségeket írhatunk le, amelyek két valós vektorral, ill. tenzorral jellemezhetők. A valós komponensekre kidolgozott algebrai szabályok minden további nélkül alkalmazhatók a komplex esetre is.

A fizikai alkalmazásokban az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ komplex komponensű vektorok skaláris szorzatát célszerű az $\underline{\mathbf{a}}\underline{\mathbf{b}}^*$ összefüggéssel definiálni. Tehát a skaláris szorzat második tényezőjében a $\underline{\mathbf{b}}$ vektor komponenseinek komplex konjugáltjával számolunk. Ez a definíció a valós komponensű vektorok skaláris szorzatán nem változtat. A komplex vektorok skaláris szorzatára azonban ekkor a kommutatív törvény nem teljesül.

A fenti definíciót a sokrétű fizikai alkalmazások teszik indokolttá.

Összefoglalásként megállapíthatjuk, hogy minden $\underline{\mathbf{A}}$ operátor három sajátértékkel rendelkezik. Amennyiben $\underline{\mathbf{A}}$ valós operátor (tehát reprezentációi valós számokkal adhatók meg), akkor a λ komplex sajátértékek mellett azok λ^* komplex konjugáltjai is sajátértékek.

13.5. Hermite-operátorok

Tárgyalásmódunk komplex komponensű tenzorokra is kiterjeszthető. Példaként a kvantummechanikában nagyon fontos ún. *hermitikus operátorokkal* foglalkozunk.

Ezen operátorok tetszőleges reprezentációjára fennáll, hogy

$$\mathbf{A} = \tilde{\mathbf{A}}^*, \qquad 13.5\text{-}1$$

azaz az operátor megegyezik transzponáltjának konjugáltjával.

Legyenek a komplex komponensű $\underline{\mathbf{A}}$ operátor sajátértékei a $\lambda^{(k)}$ $(k=1, 2, 3)$ számok, sajátvektorai az $\underline{\mathbf{s}}^{(k)}$ vektorok. Természetesen ezek most általában komplex számok, ill. komplex komponensű vektorok.

A következőkben belátjuk, hogy a hermitikus operátorok sajátértékei mindig valósak.

Legyen

$$\underline{\mathbf{A}}\underline{\mathbf{s}}^{(k)} = \lambda^{(k)}\underline{\mathbf{s}}^{(k)}, \qquad 13.5\text{-}2$$

és képezzük ezen egyenlet

$$\underline{\mathbf{A}}^*\underline{\mathbf{s}}^{(k)*} = \lambda^{(k)*}\underline{\mathbf{s}}^{(k)*} \qquad 13.5\text{-}3$$

konjugáltját. Mivel $\underline{\mathbf{A}}$ hermitikus, 13.5-3 átírható az

$$\underline{\mathbf{A}}^*\underline{\mathbf{s}}^{(k)*} = \underline{\mathbf{s}}^{(k)*}\tilde{\underline{\mathbf{A}}} = \underline{\mathbf{s}}^{(k)*}\underline{\mathbf{A}} = \lambda^{(k)*}\underline{\mathbf{s}}^{(k)*} \qquad 13.5\text{-}4$$

formára. Szorozzuk be 13.5-2-t $\underline{\mathbf{s}}^{(k)*}$-gal, 13.5-4-et $\underline{\mathbf{s}}^{(k)}$-val. Azt kapjuk, hogy

$$\underline{\mathbf{s}}^{(k)*}\underline{\mathbf{A}}\underline{\mathbf{s}}^{(k)} = \lambda^{(k)}\underline{\mathbf{s}}^{(k)}\underline{\mathbf{s}}^{(k)*},$$

$$\underline{\mathbf{s}}^{(k)*}\underline{\mathbf{A}}\underline{\mathbf{s}}^{(k)} = \lambda^{(k)*}\underline{\mathbf{s}}^{(k)}\underline{\mathbf{s}}^{(k)*}.$$

A második egyenletet kivonva az elsőből, a

$$(\lambda^{(k)} - \lambda^{(k)*})\underline{\mathbf{s}}^{(k)}\underline{\mathbf{s}}^{(k)*} = 0$$

egyenlethez jutunk. Mivel

$$\underline{\mathbf{s}}^{(k)}\underline{\mathbf{s}}^{(k)*} > 0,$$

kell, hogy

$$\lambda^{(k)} = \lambda^{(k)*}$$

legyen, vagyis a hermitikus mátrixok sajátértékei tényleg valósak.

A szimmetrikus operátorok a hermitikus operátorok különleges esetét képezik. A $\underline{\mathbf{T}}$ tehetetlenségi tenzor tehát hermitikus operátor. Ezzel viszont bebizonyítottuk, hogy a tehetetlenségi tenzor három valós sajátértékkel rendelkezik.

14. Tenzorok előállítása diádok segítségével

Megállapítottuk, hogy egy $\mathbf{P}$ homogén lineáris vektoroperáció meghatározásához elegendő három, nem komplanáris vektor leképezését megadni. Legyen pl.

$$\mathbf{P}\mathbf{a}^{(k)}=\mathbf{b}^{(k)} \quad (k=1, 2, 3), \tag{14-1}$$

ahol

$$(\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)})\neq 0. \tag{14-2}$$

Jelöljük $\boldsymbol{\alpha}^{(k)}$-val az $\mathbf{a}^{(k)}$ vektorok reciprok vektorhármasát ($k=1, 2, 3$). Tehát

$$\mathbf{a}^{(k)}\boldsymbol{\alpha}^{(l)}=\delta_{kl}. \tag{14-3}$$

Az $\boldsymbol{\alpha}^{(k)}$ vektorok segítségével a $\mathbf{P}$ operátort a

$$\mathbf{P}=\sum_{l=1}^{3}\mathbf{b}^{(l)}\circ\boldsymbol{\alpha}^{(l)} \tag{14-4}$$

alakban állíthatjuk elő, hiszen

$$\left[\sum_{l=1}^{3}\mathbf{b}^{(l)}\circ\boldsymbol{\alpha}^{(l)}\right]\mathbf{a}^{(k)}=\sum_{l=1}^{3}\mathbf{b}^{(l)}(\boldsymbol{\alpha}^{(l)}\mathbf{a}^{(k)})=\mathbf{b}^{(k)}, \tag{14-5}$$

a 14.1 definíciónak megfelelően.

14.1. Elfajuló operátorok

Elfajuló operátorok esetén az inverz operátor nem létezik, tehát a 12.2. fejezet értelmében a $\mathbf{b}^{(k)}$ vektorok nem lineárisan függetlenek.

Tegyük fel, hogy a $\mathbf{b}^{(k)}$ vektorok komplanárisak, azaz legyen pl.

$$\mathbf{b}^{(k)}=\varphi_k\bar{\mathbf{b}}^{(1)}+\eta_k\bar{\mathbf{b}}^{(2)}, \quad (k=1, 2, 3), \tag{14.1-1}$$

ahol

$$\bar{\mathbf{b}}^{(1)}\times\bar{\mathbf{b}}^{(2)}\neq\mathbf{0} \quad \text{és} \quad \varphi_k^2+\eta_k^2>0.$$

14.1-1 segítségével a $\mathbf{P}$ operátor a

$$\mathbf{P}=\sum_{l}(\varphi_l\bar{\mathbf{b}}^{(1)}+\eta_l\bar{\mathbf{b}}^{(2)})\circ\boldsymbol{\alpha}^{(l)} \tag{14.1-2}$$

alakban állítható elő, ahol $\boldsymbol{\alpha}^{(l)}$ ($l=1, 2, 3$) az $\mathbf{a}^{(k)}$ vektorok reciprok hármasát jelenti. Vezessük be az

$$\bar{\boldsymbol{\alpha}}^{(1)}=\varphi_1\boldsymbol{\alpha}^{(1)}+\varphi_2\boldsymbol{\alpha}^{(2)}+\varphi_3\boldsymbol{\alpha}^{(3)},$$

$$\bar{\boldsymbol{\alpha}}^{(2)}=\eta_1\boldsymbol{\alpha}^{(1)}+\eta_2\boldsymbol{\alpha}^{(2)}+\eta_3\boldsymbol{\alpha}^{(3)}$$

jelöléseket. Ezzel 14.1-2 átírható a

$$\mathbf{P}=\bar{\mathbf{b}}^{(1)}\circ\bar{\boldsymbol{\alpha}}^{(1)}+\bar{\mathbf{b}}^{(2)}\circ\bar{\boldsymbol{\alpha}}^{(2)} \tag{14.1-3}$$

formába. Látható tehát, hogy azok az operátorok, amelyek minden vektort egy síkba képeznek le, két diád összegeként állíthatók elő.

Amennyiben egy $\mathbf{P}$ operátor minden vektort egy $\mathbf{b}$ vektor irányába képez le, azaz

$$\mathbf{P}\mathbf{a}^{(k)}=\varphi_k\mathbf{b} \quad (k=1, 2, 3),$$

akkor $\mathbf{P}$ előállítható a

$$\mathbf{P}=\sum_{k=1}^{3}\varphi_k(\mathbf{b}\circ\boldsymbol{\alpha}^{(k)}) \tag{14.1-4}$$

alakban.

Bevezetve az

$$\bar{\boldsymbol{\alpha}}=\sum_{l}\varphi_l\boldsymbol{\alpha}^{(l)}$$

jelölést, 14.1-4 a

$$\mathbf{P}=\mathbf{b}\circ\bar{\boldsymbol{\alpha}} \tag{14.1-5}$$

alakot ölti. Ebben az esetben tehát a $\mathbf{P}$ operátor egyetlen diád segítségével is meghatározható.

A legerősebb degenerációt az jelenti, hogy egy operátor minden vektort a zérusvektorra képez le. Az ilyen operátort zérusoperátornak nevezzük.

14.2. Sajátértékek és sajátvektorok

Legyen adva a $\mathbf{P}$ operátor a

$$\mathbf{P}\mathbf{a}^{(k)}=\mathbf{b}^{(k)} \quad (k=1, 2, 3) \tag{14.2-1}$$

összefüggések segítségével. Vizsgáljuk meg, hogy milyen kapcsolat van az $\mathbf{a}^{(k)}$, $\mathbf{b}^{(k)}$ vektorok és a $\mathbf{P}$ operátor sajátvektorai között.

Legyen

$$\mathbf{s}=\sum_{k}s_k\mathbf{a}^{(k)} \tag{14.2-2}$$

$\mathbf{P}$ egy sajátvektora. Ekkor

$$\mathbf{P}\mathbf{s}=\sum_{k=1}^{3}s_k\mathbf{b}^{(k)}, \tag{14.2-3}$$

másrészt, mivel $\mathbf{s}$ sajátvektor,

$$\mathbf{P}\mathbf{s}=\lambda\mathbf{s}. \tag{14.2-4}$$

14.2-3-ból kivonva 14.2-4-et, adódik, hogy

$$\sum_{k=1}^{3} s_k(\underline{\mathbf{b}}^{(k)} - \lambda\underline{\mathbf{a}}^{(k)}) = 0. \qquad 14.2\text{-}5$$

$\sum s_k^2 > 0$ esetén ez csak akkor teljesülhet, ha a

$$\underline{\mathbf{b}}^{(k)} - \lambda\underline{\mathbf{a}}^{(k)} \qquad 14.2\text{-}6$$

vektorok nem függetlenek, tehát ha

$$(\underline{\mathbf{b}}^{(1)} - \lambda\underline{\mathbf{a}}^{(1)})[(\underline{\mathbf{b}}^{(2)} - \lambda\underline{\mathbf{a}}^{(2)}) \times (\underline{\mathbf{b}}^{(3)} - \lambda\underline{\mathbf{a}}^{(3)})] = 0. \qquad 14.2\text{-}7$$

14.2-7-et kifejtve, λ-ra a

$$c_0\lambda^3 - c_1\lambda^2 + c_2\lambda - c_3 = 0 \qquad 14.2\text{-}8$$

alakú harmadfokú egyenlet adódik, ahol

$$c_0 = (\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(3)}) \neq 0,$$

$$c_1 = \frac{1}{2} \sum_{k,l,m} \varepsilon_{klm}(\underline{\mathbf{a}}^{(k)}, \underline{\mathbf{a}}^{(l)}, \underline{\mathbf{b}}^{(m)}), \qquad 14.2\text{-}9$$

$$c_2 = \frac{1}{2} \sum_{k,l,m} (\underline{\mathbf{a}}^{(k)}, \underline{\mathbf{b}}^{(l)}, \underline{\mathbf{b}}^{(m)}),$$

$$c_3 = (\underline{\mathbf{b}}^{(1)}, \underline{\mathbf{b}}^{(2)}, \underline{\mathbf{b}}^{(3)}).$$

14.2-8 tulajdonképpen a szekuláris egyenlet (13.1 pont), csak az egyenlet együtthatóit most az $\underline{\mathbf{a}}^{(k)}$, $\underline{\mathbf{b}}^{(k)}$ vektorok segítségével határoztuk meg.

A szekuláris egyenletnek van legalább egy valós megoldása, így minden $\underline{\mathbf{P}}$ operátor legalább egy valós sajátvektorral rendelkezik.

Amennyiben a szekuláris egyenlet megoldásai komplex értékek, akkor ezekhez komplex sajátvektorok tartoznak. Egy háromdimenziós operátorhoz legfeljebb három független sajátvektor tartozhat. Kétdimenziós esetben pedig a független sajátvektorok száma legfeljebb kettő.

Léteznek azonban olyan operátorok, amelyek az operátor dimenziójánál kevesebb független sajátvektorral rendelkeznek. Most ezt az esetet illusztráljuk két példával.

Legyen

$$\underline{\mathbf{P}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A szekuláris egyenlet

$$(1-\lambda)^2 = 0.$$

Ebből a $\lambda_1 = 1$ kétszeres sajátérték adódik.

Megállapítható, hogy ehhez az $(s_1, 0)$ komponensű sajátvektor tartozik. A $\underline{\mathbf{P}}$ operátornak tehát csak egyetlen sajátvektora létezik.

Példaként foglalkozzunk még a

$$\underline{\mathbf{P}} = A\underline{\mathbf{E}} + \mathbf{a} \circ \mathbf{b} \qquad 14.2\text{-}10$$

háromdimenziós operátor sajátvektorainak vizsgálatával. A 14.2-10 operátor egy tetszőleges $\underline{\mathbf{x}}$ vektort a

$$\underline{\mathbf{P}}\underline{\mathbf{x}} = A\underline{\mathbf{x}} + \underline{\mathbf{a}}(\underline{\mathbf{b}}\underline{\mathbf{x}})$$

vektorra képez le. A

$$\underline{\mathbf{P}}\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}}$$

egyenlőség az $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ vektorra akkor teljesülhet, ha

$$\underline{\mathbf{b}}\underline{\mathbf{x}} = 0, \quad \text{és ekkor} \quad \lambda = A,$$

vagy ha

$$\underline{\mathbf{x}} = \alpha\underline{\mathbf{a}}, \quad \text{és ekkor} \quad \lambda = A + \underline{\mathbf{a}}\underline{\mathbf{b}}.$$

Ennek megfelelően $\underline{\mathbf{P}}$ sajátvektorai az $\underline{\mathbf{a}}$ irányába mutató vektorok, ill. a $\underline{\mathbf{b}}$-re merőleges vektorok. Válasszunk most a $\underline{\mathbf{b}}$-re merőleges síkban két, nem párhuzamos $\underline{\mathbf{s}}_1$ és $\underline{\mathbf{s}}_2$ vektort. Ha $\underline{\mathbf{a}}$ nem ebben a síkban fekszik, akkor $\underline{\mathbf{s}}_1$, $\underline{\mathbf{s}}_2$ és $\underline{\mathbf{s}}_3 = d\underline{\mathbf{a}}$ a $\underline{\mathbf{P}}$ operátor három lineárisan független sajátvektora.

Abban az esetben azonban, ha teljesül az $\underline{\mathbf{a}}\underline{\mathbf{b}} = 0$ feltétel, tehát $\underline{\mathbf{a}}$ is a $\underline{\mathbf{b}}$-re merőleges síkban fekszik, akkor $\underline{\mathbf{P}}$ összes sajátvektorai egy síkban vannak, így $\underline{\mathbf{P}}$ nek legfeljebb két lineárisan független sajátvektora létezhet.

A 14.2-10 operátorra vonatkozó $\det(\mathbf{P} - \lambda\mathbf{E}) = 0$ szekuláris egyenlet tetszőleges reprezentációban a

$$\det\left[(A - \lambda)\mathbf{E} + \mathbf{a} \circ \mathbf{b}\right] = 0$$

alakot ölti. A 13.1-3 egyenletben az $\mathbf{A} \to \mathbf{a} \circ \mathbf{b}$ és $\lambda \to \lambda - A$ helyettesítést elvégezve, 13.1-8 alapján adódik, hogy

$$(A - \lambda)^3 - \operatorname{sp}(\mathbf{a} \circ \mathbf{b})(A - \lambda)^2 + \operatorname{sp}(\overline{\mathbf{a} \circ \mathbf{b}})(A - \lambda) - \det(\mathbf{a} \circ \mathbf{b}) = 0.$$

Itt $\overline{\mathbf{a} \circ \mathbf{b}}$ az $\mathbf{a} \circ \mathbf{b}$ mátrix adjungáltját jelenti. Mivel $\det(\mathbf{a} \circ \mathbf{b}) = 0$ és $\overline{\mathbf{a} \circ \mathbf{b}} = \mathbf{0}$, a sajátértékegyenlet az

$$(A - \lambda)^3 - (A - \lambda)^2(\mathbf{ab}) = 0$$

alakra egyszerűsödik.

Tehát

$$\lambda = A$$

kétszeres megoldása a szekuláris egyenletnek, míg a harmadik megoldást a

$$\lambda = A + \underline{\mathbf{a}}\underline{\mathbf{b}}$$

összefüggés adja. $\underline{\mathbf{a}}\underline{\mathbf{b}} = 0$ esetén mindhárom sajátérték egybeesik, és az operátor harmadik sajátvektora is az első két sajátvektor síkjába kerül.

14.3. Független sajátvektorokkal rendelkező operátorok előállítása

Megállapítottuk, hogy három lineárisan független vektor és a hozzájuk tartozó reciprok vektorok segítségével tetszőleges tenzor diádok összegeként állítható elő (14. fejezet).

Amennyiben egy operátor három független sajátvektorral rendelkezik, akkor a diádokba fejtést a sajátvektorok segítségével is elvégezhetjük.

Legyen

$$\underline{\mathbf{P}}\underline{\mathbf{s}}^{(k)} = \lambda^{(k)}\underline{\mathbf{s}}^{(k)} \quad (k=1, 2, 3),$$

ahol

$$\det(\underline{\mathbf{s}}^{(1)}, \underline{\mathbf{s}}^{(2)}, \underline{\mathbf{s}}^{(3)}) \neq 0.$$

Ekkor a $\underline{\mathbf{P}}$ operátor 14.5 értelmében a

$$\underline{\mathbf{P}} = \sum \lambda^{(k)}(\underline{\mathbf{s}}^{(k)} \circ \underline{\mathbf{S}}^{(k)}) \qquad 14.3\text{-}1$$

alakban állítható elő, ahol $\underline{\mathbf{S}}^{(k)}$ az $\underline{\mathbf{s}}^{(k)}$ $(k=1, 2, 3)$ vektorhármas reciprok vektorait jelenti.

Megjegyezzük még, hogy a 14.3-1 előállítás akkor is érvényes, ha $\underline{\mathbf{P}}$ háromnál több sajátvektorral rendelkezik.

15. Néhány különleges operátor

15.1. A szimmetrikus operátor sajátvektorainak vizsgálata

Tegyük fel, hogy a $\underline{\mathbf{P}}$ operátor három, páronként ortogonális $\underline{\mathbf{s}}^{(k)}$ sajátvektorral rendelkezik, és legyenek $\underline{\mathbf{s}}^{(k)}$-k egységvektorok. Ekkor

$$\underline{\mathbf{s}}^{(i)}\underline{\mathbf{s}}^{(j)} = \delta_{ij} \quad (i, j=1, 2, 3). \qquad 15.1\text{-}1$$

Ez esetben az $\underline{\mathbf{s}}^{(k)}$ vektorok reciprok vektorhármasa egybeesik az eredeti vektorokkal, tehát

$$\underline{\mathbf{s}}^{(k)} = \underline{\mathbf{S}}^{(k)}.$$

A $\underline{\mathbf{P}}$ operátor tehát a

$$\underline{\mathbf{P}} = \sum_{k=1}^{3} \lambda^{k}(\underline{\mathbf{s}}^{(k)} \circ \underline{\mathbf{s}}^{(k)})$$

alakban állítható elő. Ennek következtében tetszőleges vektorok esetében fennáll, hogy

$$\underline{\mathbf{x}}\underline{\mathbf{P}}\underline{\mathbf{y}} = \underline{\mathbf{y}}\underline{\mathbf{P}}\underline{\mathbf{x}},$$

azaz $\underline{\mathbf{P}} = \underline{\widetilde{\mathbf{P}}}$, tehát $\underline{\mathbf{P}}$ szimmetrikus operátor.

Amennyiben tehát egy operátornak van három, páronként ortogonális sajátvektora, akkor az operátor szimmetrikus.

A fenti tétel megfordítható! Legyen ugyanis $\underline{\mathbf{T}}=\tilde{\underline{\mathbf{T}}}$, akkor

$$\underline{\mathbf{T}}\underline{\mathbf{s}}=\underline{\mathbf{s}}\tilde{\underline{\mathbf{T}}}=\underline{\mathbf{s}}\underline{\mathbf{T}}=\lambda_s\underline{\mathbf{s}},$$

tehát $\underline{\mathbf{T}}$ jobb és bal oldali sajátvektorai megegyeznek. A

$$\underline{\mathbf{T}}=\sum \lambda^{(k)}(\underline{\mathbf{s}}^{(k)}\circ\underline{\mathbf{S}}^{(k)})$$

előállításból azonban következik, hogy ha $\underline{\mathbf{s}}^{(k)}$ jobb oldali sajátvektor-rendszer, akkor $\underline{\mathbf{S}}^{(k)}$ bal oldali és megfordítva. Ily módon azonban a $\underline{\mathbf{T}}$ operátor esetén az $\underline{\mathbf{s}}^{(k)}$ vektorok megegyeznek saját reciprok vektoraikkal. Következésképpen az $\underline{\mathbf{s}}^{(k)}$ vektorok páronként ortogonálisak. A tételnek ezt a részét a $\underline{\mathbf{T}}$ tehetetlenségi tenzor tárgyalásakor már bizonyítottuk.

15.2. Az antiszimmetrikus operátor

Ha egy $\underline{\mathbf{A}}$ operátort egy adott koordináta-rendszerben antiszimmetrikus mátrix reprezentál, akkor $\underline{\mathbf{A}}$ tetszőleges reprezentációja antiszimmetrikus mátrix.

Az antiszimmetrikus mátrix esetén

$$\operatorname{spur}\mathbf{A}=0,$$

$$\operatorname{spur}\tilde{\tilde{\mathbf{A}}}=A_{12}^2+A_{23}^2+A_{31}^2=a^2>0,$$

és

$$\det\mathbf{A}=0.$$

Így a szekuláris egyenlet

$$\lambda^3+a^2\lambda=0$$

alakú, tehát

$$\lambda_3=0,\quad \text{és}\quad \lambda_{1,2}=\pm ia. \qquad 15.2\text{-}1$$

Az antiszimmetrikus operátor rangja tehát kettő. Az ia és $-ia$ sajátértékekhez az $\underline{\mathbf{a}}^{(1)}$ és $\underline{\mathbf{a}}^{(1)*}$ komplex konjugált sajátvektorok tartoznak. Ezek segítségével az antiszimmetrikus operátor az

$$\underline{\mathbf{A}}=ia[(\underline{\mathbf{a}}^{(1)}\circ\underline{\boldsymbol{\alpha}}^{(1)})-(\underline{\mathbf{a}}^{(1)*}\circ\underline{\boldsymbol{\alpha}}^{(1)*})] \qquad 15.2\text{-}2$$

alakban állítható elő, ahol $\underline{\boldsymbol{\alpha}}^{(1)}$ és $\underline{\boldsymbol{\alpha}}^{(1)*}$ az $\underline{\mathbf{a}}^{(1)}$ és $\underline{\mathbf{a}}^{(1)*}$ vektorok reciprok vektorai.

15.2-2-ből leolvasható az is, hogy amennyiben egy operátor sajátértékei 15.2-1-nek megfelelőek, akkor az operátor antiszimmetrikus.

15.3. A vektoriális szorzat tenzorreprezentációja

A vektoriális szorzás egy vektor—vektor hozzárendelés. Ezt a hozzárendelést lineáris operációként is kezelhetjük. Legyen ugyanis

$$\underline{\mathbf{a}}\times\underline{\mathbf{x}}=\underline{\mathbf{A}}\underline{\mathbf{x}}. \qquad 15.3\text{-}1$$

A vektoriális szorzat definíciója alapján megállapítható, hogy 15.3-1-ben az $\underline{\mathbf{A}}$ operátor elemei az

$$A_{kl}=\sum_m \varepsilon_{ikm}a_m \qquad 15.3\text{-}2$$

formulával határozhatók meg.

A vektoriális szorzat tulajdonságaiból 15.3-2 felhasználása nélkül is megállapíthatjuk az $\underline{\mathbf{A}}$ operátor tulajdonságait. A sajátértékei és sajátvektorai eleget tesznek az

$$\underline{\mathbf{a}}\times\underline{\mathbf{a}}^{(s)}=\lambda_s\underline{\mathbf{a}}^{(s)} \quad (s=1,2,3) \qquad 15.3\text{-}3$$

egyenletnek. 15.3-3 valós megoldását az

$$\underline{\mathbf{a}}^{(1)}=\underline{\mathbf{a}} \quad \text{és} \quad \lambda_s=0$$

értékek szolgáltatják.

Komplex megoldásokat azon $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok segítségével képezhetünk, amelyekre

$$\underline{\mathbf{b}}\times\underline{\mathbf{c}}=\underline{\mathbf{a}}, \quad \underline{\mathbf{b}}^2=\underline{\mathbf{c}}^2=a \quad \text{és} \quad \underline{\mathbf{b}}\underline{\mathbf{c}}=0. \qquad 15.3\text{-}4$$

15.3-4 szerint tehát $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ $\sqrt{a}$ abszolút értékű vektorok és $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$, $\underline{\mathbf{a}}$ páronként ortogonálisak. A $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok segítségével képzett $\underline{\mathbf{b}}+i\underline{\mathbf{c}}$ és $\underline{\mathbf{b}}-i\underline{\mathbf{c}}$ vektorokra fennáll a 15.3-3 összefüggés, hiszen a 15.3-4 összefüggések felhasználásával

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}+i\underline{\mathbf{c}})=(\underline{\mathbf{b}}\times\underline{\mathbf{c}})\times(\underline{\mathbf{b}}+i\underline{\mathbf{c}})=-ia(\underline{\mathbf{b}}+i\underline{\mathbf{c}}) \qquad 15.3\text{-}5$$

és

$$\underline{\mathbf{a}}\times(\underline{\mathbf{b}}-i\underline{\mathbf{c}})=(\underline{\mathbf{b}}\times\underline{\mathbf{c}})\times(\underline{\mathbf{b}}-i\underline{\mathbf{c}})=ia(\underline{\mathbf{b}}-i\underline{\mathbf{c}}). \qquad 15.3\text{-}6$$

A $\underline{\mathbf{b}}+i\underline{\mathbf{c}}$ sajátvektor tehát a $-ia$, a $\underline{\mathbf{b}}-i\underline{\mathbf{c}}$ sajátvektor pedig az ia sajátértékhez tartozik.

Megállapítható tehát, hogy a vektoriális szorzást reprezentáló operátor sajátértékei és sajátvektorai azonosak az általános antiszimmetrikus operátor sajátértékeivel és sajátvektoraival.

Következésképpen egy $\underline{\mathbf{a}}$ vektorral való vektoriális szorzás mindig helyettesíthető egy antiszimmetrikus operátor alkalmazásával, és tetszőleges antiszimmetrikus operátorhoz is mindig megtalálható az a vektor, amellyel vektoriálisan szorozva, ugyanarra az eredményre jutunk, mint az antiszimmetrikus operátor alkalmazásával.

A 15.3-4 feltételek szerint a $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorok abszolút értéke adott, valamint $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ merőlegesek egymásra.

Ez azonban azt jelenti, hogy $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ nem határozható meg egyértelműen, hiszen $\underline{\mathbf{b}}$-t és $\underline{\mathbf{c}}$-t $\underline{\mathbf{a}}$ körül elforgatva, a kapott $\underline{\mathbf{b}}'$ és $\underline{\mathbf{c}}'$-ből képzett $\underline{\mathbf{b}}'+i\underline{\mathbf{c}}'$ és $\underline{\mathbf{b}}'-i\underline{\mathbf{c}}'$ vektorok sajátvektorok maradnak.

A többértelműség azonban csak látszólagos. Ha a $\underline{\mathbf{b}}$ és $\underline{\mathbf{c}}$ vektorokat φ szöggel elforgatjuk $\underline{\mathbf{a}}$ körül, akkor a

$$\underline{\mathbf{b}} \to \underline{\mathbf{b}}' = \underline{\mathbf{b}} \cos\varphi - \underline{\mathbf{c}} \sin\varphi,$$

$$\underline{\mathbf{c}} \to \underline{\mathbf{c}}' = \underline{\mathbf{b}} \sin\varphi + \underline{\mathbf{c}} \cos\varphi$$

vektorokhoz jutunk. A kapott sajátvektorok

$$\underline{\mathbf{b}}' + i\underline{\mathbf{c}}' = (\underline{\mathbf{b}} + i\underline{\mathbf{c}})e^{i\varphi}$$

és

$$\underline{\mathbf{b}}' - i\underline{\mathbf{c}}' = (\underline{\mathbf{b}} - i\underline{\mathbf{c}})e^{-i\varphi}.$$

Az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok elforgatása miatt a sajátvektorok egy egységnyi abszolút értékű komplex számmal szorzódnak. A sajátvektorok azonban, mint tudjuk, csak egy szorzótényező erejéig meghatározottak. A komplex sajátvektorok „iránya" tehát nem változik a komponensek elforgatása során.

Megállapítható tehát, hogy az antiszimmetrikus operátorok pontosan három sajátvektorral rendelkeznek. Ügyelnünk kell azonban arra, hogy a komplex sajátvektorok komplex számmal történő szorzása megengedett.

16. Geometriai alkalmazások

16.1. A másodrendű görbék és felületek általános egyenlete

Másodrendű felületeknek nevezzük azokat a felületeket, amelyeknek a helyvektorai kvadratikus egyenletet elégítenek ki.

Tehát

$$\underline{\mathbf{x}}\underline{\mathbf{A}}\underline{\mathbf{x}} + 2\boldsymbol{\beta}\underline{\mathbf{x}} + \gamma = 0, \qquad 16.1\text{-}1$$

ahol $\underline{\mathbf{x}}$ a felület futóvektora, $\underline{\mathbf{A}}$ pedig egy szimmetrikus tenzor.

A 16.1-1 alakú egyenlet nemcsak háromdimenziós térben érvényes, hanem az egyenlet és a következő meggondolások minden további nélkül érvényben maradnak kétdimenziós esetben is. A 16.1-1 típusú egyenlet kétdimenziós esetben másodrendű görbét ír le. A következő gondolatmenet algebrailag egyszerűen általánosítható n dimenzióra is, geometriai tartalmat azonban csak a két- és háromdimenziós esetnek tulajdoníthatunk. Átvitt értelemben a relativitáselméletben négydimenziós esetek is értelmezhetők (lásd 20. fejezet).

16.1.1. A centrális egyenletek

A 16.1-1 egyenletet homogén kvadratikus formává alakíthatjuk át, ha létezik az $\underline{\mathbf{A}}^{-1}$ reciprok tenzor.

Legyen ugyanis

$$\underline{\mathbf{x}}=\underline{\mathbf{y}}-\underline{\boldsymbol{\alpha}}, \tag{16.1-2}$$

ahol $\underline{\boldsymbol{\alpha}}$ egy állandó vektor. 16.1-1-ben **x** helyett 16.1-2 felhasználásával az

$$(\underline{\mathbf{y}}-\underline{\boldsymbol{\alpha}})\underline{\mathbf{A}}(\underline{\mathbf{y}}-\underline{\boldsymbol{\alpha}})+2\underline{\boldsymbol{\beta}}(\underline{\mathbf{y}}-\underline{\boldsymbol{\alpha}})+\gamma=0 \tag{16.1-3}$$

formulát használhatjuk.

Felhasználva, hogy $\underline{\mathbf{A}}$ szimmetrikus tenzor, az

$$\underline{\mathbf{y}}\underline{\mathbf{A}}\underline{\mathbf{y}}+2(\underline{\boldsymbol{\beta}}-\underline{\boldsymbol{\alpha}}\underline{\mathbf{A}})\underline{\mathbf{y}}+\underline{\boldsymbol{\alpha}}\underline{\mathbf{A}}\underline{\boldsymbol{\alpha}}-2\underline{\boldsymbol{\beta}}\underline{\boldsymbol{\alpha}}+\gamma=0 \tag{16.1-4}$$

formulához jutunk.

Az utóbbi kifejezésben a lineáris tag eltűnik, ha

$$\underline{\boldsymbol{\beta}}=\underline{\boldsymbol{\alpha}}\underline{\mathbf{A}};$$

vagyis ha $\det \underline{\mathbf{A}}\neq 0$ és $\underline{\mathbf{A}}^{-1}$ létezik, akkor az $\underline{\boldsymbol{\alpha}}$ vektort

$$\underline{\boldsymbol{\alpha}}=\underline{\mathbf{A}}^{-1}\underline{\boldsymbol{\beta}}\text{-nak} \tag{16.1-5}$$

választva, a másodrendű felület egyenletét az

$$\underline{\mathbf{y}}\underline{\mathbf{A}}\underline{\mathbf{y}}=p \tag{16.1-6}$$

formában írhatjuk fel, ahol

$$p=\underline{\boldsymbol{\beta}}\underline{\mathbf{A}}^{-1}\underline{\boldsymbol{\beta}}-\gamma.$$

16.1-6-ot a másodrendű felület centrális egyenletének, az

$$\underline{\boldsymbol{\alpha}}=\underline{\mathbf{A}}^{-1}\underline{\boldsymbol{\beta}}$$

vektort pedig a felület centrumának nevezzük.

16.1.2. A kanonikus egyenlet

$\underline{\mathbf{A}}$ szimmetrikus tenzor és $\det \mathbf{A}\neq 0$, így $\underline{\mathbf{A}}$ három valós sajátértékkel és három, páronként egymásra merőleges sajátvektorral rendelkezik. Amennyiben $\underline{\mathbf{A}}$-t sajátvektorainak koordináta-rendszerében reprezentáljuk, akkor

$$\mathcal{K}_s(\underline{\mathbf{A}})=\mathbf{A}$$

diagonálmátrix, tehát

$$A_{kl}=\delta_{kl}\lambda_k. \tag{16.1-7}$$

Ebben a reprezentációban a másodrendű felületek 16.1-1 egyenlete a

$$\sum_{k=1}^{n} \lambda_k x_k^2 = p \qquad 16.1\text{-}8$$

alakot ölti. Bevezetve a

$$\left|\frac{p}{\lambda_k}\right| = a_k^2$$

jelölést, 16.1-8 a

$$\pm\frac{x_1^2}{a_1^2} \pm \frac{x_2^2}{a_2^2} \pm \frac{x_3^2}{a_3^2} = 1 \qquad 16.1\text{-}9$$

formába írható át, ahol az előjelek $\frac{p}{\lambda_k}$ előjelével egyeznek meg. A 16.1-9 alakú egyenletet a másodrendű felületek kanonikus egyenletének nevezzük.

12.7-20 szerint egy merev test kinetikus energiájára fennáll az

$$E_{\text{kin}} = \tfrac{1}{2}\underline{\omega}\underline{\underline{T}}\underline{\omega}$$

összefüggés. Az állandó kinetikus energiájú pontok tehát egy másodrendű felületen helyezkednek el. Mivel $\underline{\underline{T}}$ sajátértékei pozitív számok, ez a felület mindig ellipszoid. Az adott $\underline{\underline{T}}$ tehetetlenségi tenzor által definiált

$$\underline{\mathbf{x}}\underline{\underline{\mathbf{T}}}\underline{\mathbf{x}} = 1$$

egyenletű ellipszoidot tehetetlenségi ellipszoidnak nevezzük. A tehetetlenségi ellipszoid ismeretében egy merev test tetszőleges tengelyre vonatkozó tehetetlenségi nyomatéka meghatározható.

16.1.3. A másodrendű görbék részletes leírása

A 16.1-1 formulának megfelelő egyenlet kétdimenziós esetben is transzformálható a sajátvektorok koordináta-rendszerébe. A 16.1.2. fejezetben alkalmazott gondolatmenettel a másodrendű görbék kanonikus alakjához juthatunk.

A kanonikus egyenlet ebben az esetben a

$$\pm\frac{x_1^2}{a_1^2}\pm\frac{x_2^2}{a_2^2}=1 \qquad 16.1\text{-}10$$

alakot ölti.

A különböző előjelkombinációkat megvizsgálva, látható, hogy amennyiben $\frac{p}{\lambda_1}$ és $\frac{p}{\lambda_2}$ pozitív, akkor a másodrendű görbe ellipszis, ha $\frac{p}{\lambda_1}$ és $\frac{p}{\lambda_2}$ különböző előjelű, akkor hiperbola.

Amennyiben mind $\frac{p}{\lambda_1}$, mind $\frac{p}{\lambda_2}$ negatív, akkor 16.1-10 *üres alakzat*[1] egyenlete.

Az $\underline{\mathbf{A}}^{-1}$ létezése azt jelenti, hogy az $\underline{\mathbf{A}}$ tenzor λ_1 és λ_2 sajátértékei közül egyik sem zérus.

Foglalkoznunk kell még azzal az esettel, amikor $\lambda_1\neq 0$, de $\lambda_2=0$. Ekkor az előző módszer nem alkalmazható. Az $\underline{\mathbf{A}}$ tenzorhoz azonban most is találhatunk olyan $\mathcal{K}$ koordináta-rendszert, amelyben

$$\mathbf{A}=\mathcal{K}(\underline{\mathbf{A}})=\begin{bmatrix}\lambda_1 & 0\\ 0 & 0\end{bmatrix} \qquad 16.1\text{-}11$$

diagonális formát ölt.

Írjuk ki most 16.1-11 felhasználásával a 16.1-1 egyenlet koordinátás alakját:

$$\lambda_1 x_1^2+2\beta_1 x_1+2\beta_2 x_2+\gamma=0. \qquad 16.1\text{-}12$$

Alakítsuk az x_1-et tartalmazó tagokat teljes négyzetté:

$$\left(x_1+\frac{\beta_1}{\lambda_1}\right)^2=-\frac{2\beta_2}{\lambda_1}x_2+\frac{\beta_1^2}{\lambda_1^2}-\frac{\gamma}{\lambda_1},$$

illetve

$$\left(x_1+\frac{\beta_1}{\lambda_1}\right)^2=-\frac{2\beta_2}{\lambda_1}\left[x_2-\left(\frac{\beta_1^2}{2\beta_2\lambda_1}-\frac{\gamma}{2\beta_2}\right)\right]. \qquad 16.1\text{-}13$$

16.1-13-at összehasonlítva az

$$(x_1-u_1)^2=2p(x_2-u_2)$$

[1] Üres alakzatról akkor beszélünk, ha nem létezik olyan pont, amelynek koordinátái kielégítik az adott egyenletet.

egyenlettel, látható, hogy ez esetben olyan parabolát kaptunk, amelynek csúcsa az

$$u_1=\frac{\beta_1}{\lambda_1},\quad u_2=\frac{\beta_1^2}{2\beta_2\lambda_1}-\frac{\gamma}{2\beta_2}$$

pontban helyezkedik el.

Ha λ_1 is zérus, akkor 16.1-12

$$2\underline{\boldsymbol{\beta}}\underline{\mathbf{x}}+\gamma=0 \qquad 16.1\text{-}14$$

alakra egyszerűsödik, ami egyenes egyenlete.

Amennyiben még β_1, β_2 és γ értékek valamelyike is zérus, további elfajuló görbékhez, ill. üres alakzatokhoz jutunk.

Az alábbiakban táblázatosan összefoglaljuk a másodrendű görbetípusokat:

λ_1	λ_2	β_1	β_2	Kanonikus alak $p>0$	Kanonikus alak $p=0$
>0	>0	tetszőleges		$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}=1$ (ellipszis)	$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}=0$ (az origó)
>0	<0	tetszőleges		$\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}=1$ (hiperbola)	$\left(\frac{x_1}{a_1}+\frac{x_2}{a_2}\right)\left(\frac{x_1}{a_1}-\frac{x_2}{a_2}\right)=0$ metsző egyenespár
<0	<0	tetszőleges		üres alakzat	$\frac{x_1}{a_1^2}+\frac{x_2}{a_2^2}=0$ (az origó)

λ_1	λ_2	β_1	β_2	kanonikus egyenlet
$\neq 0$	0	$\geqq 0$	>0	$(x_1-u_1)^2=2p(x_2-u_2)$ (parabola)
$\neq 0$	0	$\geqq 0$	0	$(x_1-u_1)^2=$ konstans (egyenespár)
0	0	$\beta_1^2+\beta_2^2>0$		$2\beta_1 x_1+2\beta_2 x_2=$ konstans (egyenes)
0	0	0	0	semmitmondó egyenlet

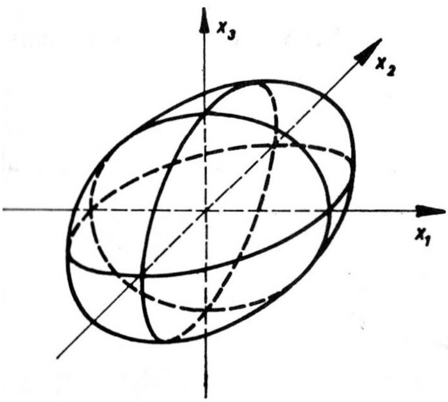
16.1.4. A másodrendű felületek részletes leírása

Megállapítottuk, hogy ha az $\underline{\mathbf{A}}$ tenzor sajátértékei nem egyenlők zérussal, akkor a másodrendű felületek egyenlete

$$\pm\frac{x_1^2}{a_1^2}\pm\frac{x_2^2}{a_2^2}\pm\frac{x_3^2}{a_3^2}=1$$

alakra hozható. Itt az előjelek $\frac{p}{\lambda_k}$ előjelével egyeznek meg, ahol $p=\boldsymbol{\beta}\underline{\mathbf{A}}^{-1}\boldsymbol{\beta}-\gamma$.

Az előjelektől függően a következő lényegesen különböző felületeket kapjuk:



16.1. ábra

1. Ha mindhárom előjel pozitív, az

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\frac{x_3^2}{a_3^2}=1 \qquad 16.1\text{-}15$$

felülethez, az ún. *ellipszoidhoz* jutunk *(16.1. ábra)*. Ha $a_1=a_2=a_3$, akkor 16.1-15 gömb egyenlete.

2. Ha két előjel pozitív, egy negatív, akkor például az

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}=1 \qquad 16.1\text{-}16$$

egyenletű felületet kapjuk. Az ilyen típusú felületeket *egyköpenyű hiperboloidok-nak* nevezzük *(16.2. ábra)*.

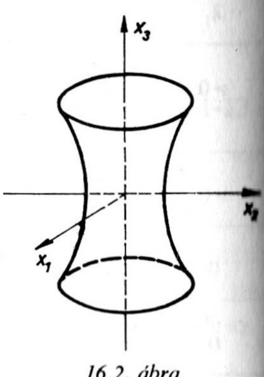
3. Amennyiben két előjel negatív és egy pozitív, az

$$\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}=1 \qquad 16.1\text{-}17$$

ún. *kétköpenyű hiperboloid* egyenletéhez jutunk *(16.3. ábra)*. Ha mindhárom előjel negatív, a kapott egyenlet üres alakzat egyenlete.

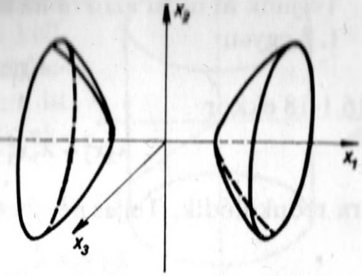
Ez az eljárás nem alkalmazható, ha az $\underline{\mathbf{A}}$ tenzor sajátértékei között zérus is található. Ekkor ugyanis az $\underline{\mathbf{A}}^{-1}$ tenzor nem létezik.

Az $\underline{\mathbf{A}}$ tenzor természetesen most is felírható diagonálmátrix formájában, ha megfelelő $\mathcal{K}$ koordinátarendszert választunk. A diszkusszió során fellépő különböző esetek taglalásának megkönnyítésére írjuk



16.2. ábra

fel a 16.1-1 egyenletet a $\mathscr{X}$ rendszerben, és fejtsük ki koordinátás alakban:



16.3. ábra

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + 2\beta_1 x_1 + 2\beta_2 x_2 + 2\beta_3 x_1 + \gamma = 0. \qquad 16.1\text{-}18$$

Alakítsuk 16.1-18-at minden koordinátában tiszta másodfokú egyenletté. A

$$\lambda_1\left(x_1+\frac{\beta_1}{\lambda_1}\right)^2+\lambda_2\left(x_2+\frac{\beta_2}{\lambda_2}\right)^2+\left(x_3+\frac{\beta_3}{\lambda_3}\right)=c \qquad 16.1\text{-}19$$

egyenletet kapjuk, ahol c állandó. (Csak a λ_i-t és β_i-t, valamint γ-t tartalmazza.) Tulajdonképpen ugyanazt az átalakítást végeztük el koordinátákkal, amit az A^{-1} létezése esetén teljes általánosságban megtettünk. 16.1-19 azonban mutatja, hogy amennyiben valamely λ_i érték 0, a teljes négyzetté való kiegészítés értelmetlen. Mielőtt ezekre az esetekre térnénk át, foglalkozzunk még azzal az esettel, amikor $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$, ellenben $c=0$.

Ekkor

$$\lambda_1\left(x_1+\frac{\beta_1}{\lambda_1}\right)^2+\lambda_2\left(x_2+\frac{\beta_2}{\lambda_2}\right)^2+\lambda_3\left(x_3+\frac{\beta_3}{\lambda_3}\right)^2=0. \qquad 16.1\text{-}20$$

Ha sgn $\lambda_1 =$ sgn $\lambda_2 =$ sgn λ_3, akkor 16.1-20-at csak az

$$\alpha_1=-\frac{\beta_1}{\lambda_1}, \quad \alpha_2=-\frac{\beta_2}{\lambda_2}, \quad \alpha_3=-\frac{\beta_3}{\lambda_3} \qquad 16.1\text{-}21$$

pont elégíti ki.[1]

Amennyiben sgn $\lambda_1 \neq$ sgn $\lambda_2 =$ sgn λ_3 és $c=0$, akkor 16.1-20 a

$$\operatorname{sgn}\lambda_1[|\lambda_1|(x_1-\alpha_1)^2-|\lambda_2|(x_2-\alpha_2)^2-|\lambda_3|(x_3-\alpha_3)^2]=0$$

alakot ölti.

Ez az alakzat egy ún. másodrendű kúp egyenlete. Ennek bizonyítására fontoljuk meg a következőket: ha a koordináta-rendszer középpontját az α_1, α_2, α_3 pontba helyezzük, akkor

$$\operatorname{sgn}\lambda_1[|\lambda_1|x_1'^2-|\lambda_2|x_2'^2-|\lambda_3|x_3'^2]=0, \qquad 16.1\text{-}22$$

ahol x_1', x_2', x_3' az új rendszerben vett koordináták. Ennek a felületnek az $x_1=0$ síkban csak egyetlen pontja van, az origó. Azonnal látható az is, hogy amennyiben egy $P(z_1, z_2, z_3)$ pont a felületen van, akkor az OP egyenes minden pontja hozzátartozik a felülethez, hiszen $P(z_1, z_2, z_3)$-mal együtt minden $P(\mu z_1, \mu z_2, \mu z_3)$ pont is kielégíti 16.1-22-t. Ez pedig éppen a kúpra jellemző tulajdonság.

[1] Az sgn jelölés a signum (előjel) szó rövidítése, és sgn λ a λ szám előjelét jelenti.

Térjünk át most azokra az esetekre, amikor a sajátértékek valamelyike zérus.

1. Legyen

$$\lambda_1=0,\quad \lambda_2>0,\quad \lambda_3>0.$$

16.1-18 ekkor

$$\lambda_2 x_2^2+\lambda_3 x_3^2+2\beta_1 x_1+2\beta_2 x_2+2\beta_3 x_3+\gamma=0 \qquad 16.1\text{-}23$$

-ra redukálódik. Teljes négyzetté alakítva, és a 16.1-21 jelöléseket felhasználva,

$$\lambda_2(x_2-\alpha_2)^2+\lambda_3(x_3-\alpha_3)^2+2\beta_1(x_1-c)=0 \qquad 16.1\text{-}24$$

alakú egyenlethez jutunk, ahol

$$c=\frac{\beta_2^2}{2\lambda_2^2\beta_1}+\frac{\beta_3^2}{2\lambda_3^2\beta_1}-\frac{\gamma}{2\beta_1}.$$

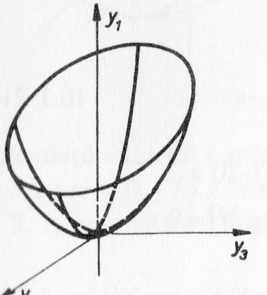
A koordináta-rendszer kezdőpontját a $P(c, \alpha_1, \alpha_2)$ pontba tolva (feltéve, hogy $\beta_1\neq 0$), a

$$\lambda_2 y_2^2+\lambda_3 y_3^2=-2\beta_1 y_1$$

egyenlethez jutunk, ahol y_i $(i=1, 2, 3)$ az új rendszerben vett koordináta.

Bevezetve az $a_2=\dfrac{1}{\sqrt{\lambda_2}}$, $a_3=\dfrac{1}{\sqrt{\lambda_3}}$ és $-\beta_1=p$ jelöléseket, az

$$\frac{y_2^2}{a_2^2}+\frac{y_3^2}{a_3^2}=2py_1 \qquad 16.1\text{-}25$$



16.4. ábra

egyenletet kapjuk.

A 16.1-25 alakra hozható felületeket *elliptikus paraboloidoknak* nevezzük *(16.4. ábra)*.

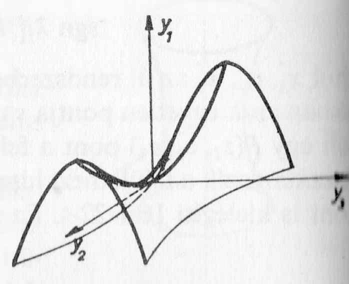
2. Ha $\lambda_1=0$, $\lambda_2<0$, $\lambda_3>0$, akkor a 16.1-18 egyenletet az előzőkhöz hasonló átalakításokkal az

$$\frac{y_2^2}{a_2^2}-\frac{y_3^2}{a_3^2}=2py_1 \qquad 16.1\text{-}26$$

alakra hozhatjuk. Ezeket a felületeket *hiperbolikus paraboloidoknak* nevezzük *(16.5. ábra)*.

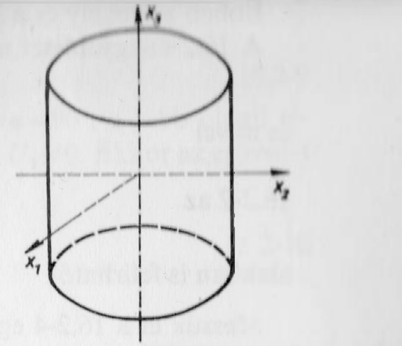
3. Ha $\lambda_1=0_1$, $\lambda_2>0$, $\lambda_3>0$ és $\beta=0$, akkor tulajdonképpen az 1. eset egy speciális fajtájával van dolgunk. 16.1-24-et felhasználva a

$$\lambda_2(x_2-\alpha_2)^2+\lambda_3(x_3-\alpha_3)^2=c \qquad 16.1\text{-}27$$



16.5. ábra

egyenletet kapjuk. Ez az (x_2, x_3) síkban egy ellipszis egyenlete (ha $c>0$). Mivel x_1 nem szerepel az egyenletben, 16.1-27-nek tetszőleges x_1-re teljesülnie kell. Ez azt jelenti, hogy a felület az (x_2, x_3) síkra merőleges *elliptikus henger (16.6. ábra)*. Amennyiben itt $c=0$, akkor egyenes egyenletét kapjuk; ha $c<0$, akkor üres alakzathoz jutunk.



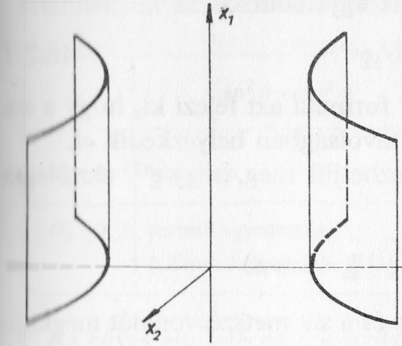
16.6. ábra

4. Ha $\lambda_1=0$, $\beta_1=0$, $\lambda_2<0$, $\lambda_3>0$, akkor az 1. eset alapján

$$\lambda_3(x_3-\alpha_3)^2-\lambda_2(x_2-\alpha_2)^2=c. \qquad 16.1\text{-}28$$

$c\neq 0$ esetén ez az (x_2, x_3) síkban fekvő hiperbola egyenlete, tehát a felület egy *hiperbola alapú henger (16.7. ábra)*, $c=0$ esetén *16.1-28* egy metsző síkpárt határoz meg.

5. $\lambda_1=\lambda_2=0$, $\lambda_3\neq 0$. Mivel két sajátérték megegyezik, a sajátirányok koordináta-rendszere nem egyértelműen meghatározott (a λ_3 értékhez tartozó irányra merőleges sík minden vektora sajátvektor). Megválaszthatjuk tehát koordináta-rendszerünket úgy, hogy pl. $\beta_1=0$ legyen. Így 16.1-18 a következő lesz:



16.7. ábra

$$\lambda_3 x_3^2+2\beta_2 x_2+2\beta_3 x_3+c=0. \qquad 16.1\text{-}29$$

Teljes négyzetté alakítva:

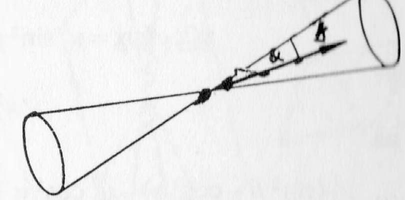
$$\lambda_3(x_3-\alpha_3)^2=-2\beta_2(x_2-c'), \qquad 16.1\text{-}30$$

ahol $c'=\dfrac{c}{2\beta_3}+\dfrac{\beta_3}{2\lambda_3^2}$. Az egyenlet parabola alapú hengert ír le.

16.2. Kúp metszése síkkal

A kúp síkmetszeteit kúpszeleteknek nevezzük. A kúp — mint láttuk — másodrendű felület. Vegyünk fel egy kúpot, amelynek tengelye a $\underline{\mathbf{k}}$ egységvektor irányában fekszik, csúcsa az origóban van és fél nyílásszöge α. Az $\underline{\mathbf{x}}$ helyvektorú pont a kúp felületén fekszik, ha

$$\underline{\mathbf{k}}\underline{\mathbf{x}}=\pm x\cos\alpha, \qquad 16.2\text{-}1$$



16.8. ábra

ahol $x=|\underline{\mathbf{x}}|$.

A két előjelre azért van szükség, mert rögtön a kettős kúp egyenletét írtuk fel *(16.8. ábra)*.

Ebben a tengely és a helyvektor hajlásszöge tompaszög is lehet.

A 16.2-1. egyenletet négyzetre emelve azt kapjuk, hogy

$$(\underline{\mathbf{k}}\underline{\mathbf{x}})^2 - x^2 \cos^2 \alpha = 0, \qquad 16.2\text{-}2$$

és mivel

$$(\underline{\mathbf{k}}\underline{\mathbf{x}})^2 = \underline{\mathbf{x}}(\underline{\mathbf{k}} \circ \underline{\mathbf{k}})\underline{\mathbf{x}}, \qquad 16.2\text{-}3$$

16.2-2 az

$$\underline{\mathbf{x}}(\underline{\mathbf{k}} \circ \underline{\mathbf{k}} - \underline{\mathbf{E}} \cos^2 \alpha)\underline{\mathbf{x}} = 0 \qquad 16.2\text{-}4$$

alakban is felírható.

Messük el a 16.2-4 egyenletű kúpot az

$$\underline{\mathbf{x}} = \underline{\mathbf{a}} + p\underline{\mathbf{K}} + g\underline{\mathbf{L}} \qquad 16.2\text{-}5$$

egyenletű síkkal, ahol

$$\underline{\mathbf{M}} = \underline{\mathbf{K}} \times \underline{\mathbf{L}} \neq \underline{\mathbf{0}}. \qquad 16.2\text{-}6$$

Egyszerűbben jutunk eredményhez, ha $\underline{\mathbf{K}}$, $\underline{\mathbf{L}}$, $\underline{\mathbf{M}}$-et három, páronként egymásra merőleges és az adott sorrendben jobbsodrású koordináta-rendszert alkotó egységvektoroknak választjuk.

Ebben a koordináta-rendszerben a 16.2-5 sík egyenlete az

$$\underline{\mathbf{x}} = u_1\underline{\mathbf{e}}^{(1)} + u_2\underline{\mathbf{e}}^{(2)} + U_3\underline{\mathbf{e}}^{(3)} \qquad 16.2\text{-}7$$

alakban írható fel, ahol U_3 állandó. A 16.2-7 formula azt fejezi ki, hogy a sík merőleges az $\underline{\mathbf{e}}^{(3)}$ tengelyre és az origótól U_3 távolságban helyezkedik el.

Az általánosság megszorítása nélkül feltételezhetjük még, hogy $\underline{\mathbf{e}}^{(2)}$ merőleges a kúp $\underline{\mathbf{k}}$ tengelyére. Ez esetben

$$\underline{\mathbf{e}}^{(1)}\underline{\mathbf{k}} = \sin\beta; \quad \underline{\mathbf{e}}^{(2)}\underline{\mathbf{k}} = 0; \quad \underline{\mathbf{e}}^{(3)}\underline{\mathbf{k}} = \cos\beta,$$

ahol β a $\underline{\mathbf{k}}$ és $\underline{\mathbf{e}}^{(3)}$ vektorok közötti szög. A kúp és a sík metszésvonalát megkapjuk, ha a 16.2-7 egyenletet a kúp

$$\underline{\mathbf{x}}(\underline{\mathbf{k}} \circ \underline{\mathbf{k}} - \underline{\mathbf{E}} \cos^2 \alpha)\underline{\mathbf{x}} = 0$$

egyenletébe behelyettesítjük.

Mivel

$$\underline{\mathbf{x}}(\underline{\mathbf{k}} \circ \underline{\mathbf{k}})\underline{\mathbf{x}} = u_1^2 \sin^2\beta + U_3^2 \cos^2\beta + 2u_1U_3 \cos\beta \sin\beta$$

és

$$x^2 = u_1^2 + u_2^2 + U_3^2,$$

az

$$u_1^2(\sin^2\beta - \cos^2\alpha) - u_2^2\cos^2\alpha + U_3^2(\cos^2\beta - \cos^2\alpha) + 2u_1U_3\cos\beta\sin\beta = 0 \qquad 16.2\text{-}8$$

egyenlethez jutunk. Feltételezve, hogy $\sin^2\beta - \cos^2\alpha \neq 0$, ezzel a tényezővel az egyenletet végigoszthatjuk és az

$$\left(u_1+U_3\frac{\cos\beta\sin\beta}{\sin^2\beta-\cos^2\alpha}\right)^2-u_2^2\frac{\cos^2\alpha}{\sin^2\beta-\cos^2\alpha}=U_3^2\frac{\cos^2\alpha\sin^2\alpha}{(\sin^2\beta-\cos^2\alpha)^2} \qquad 16.2\text{-}9$$

alakra rendezhetjük. Mivel a kúp fél nyílásszöge $0<\alpha<90°$, a jobb oldali tényező csak $U_3=0$ esetén lehet zérus. Tegyük fel, hogy $U_3\neq 0$. Ekkor az egyenlet az

$$\frac{(u_1-z)^2}{a^2}+\frac{u_2^2}{b^2}=1 \qquad 16.2\text{-}10$$

alakra hozható, ahol

$$z=U_3\frac{\cos\beta\sin\beta}{\sin^2\beta-\cos^2\alpha};\quad a^2=U_3^2\frac{\cos^2\alpha\sin^2\alpha}{(\sin^2\beta-\cos^2\alpha)^2};\quad b^2=-U_3^2\frac{\sin^2\alpha}{\sin^2\beta-\cos^2\alpha}. \qquad 16.2\text{-}11$$

A formulákból leolvasható, hogy a metszetgörbe a $\sin^2\beta>\cos^2\alpha$ esetben hiperbola, a $\sin^2\beta<\cos^2\alpha$ esetben pedig ellipszis. Ha $\sin\beta=0$, akkor a metszetgörbe kör.

A fenticket, valamint az $U_3=0$, illetve $\sin^2\beta=\cos^2\alpha$ esetek vizsgálatának eredményeit az alábbi táblázat foglalja össze:

	$\sin^2\beta>\cos^2\alpha$	$\sin^2\beta<\cos^2\alpha$	$\sin^2\beta=\cos^2\alpha$
$U_3\neq 0$	hiperbola	ellipszis	parabola
$U_3=0$	metsző egyenespár (elfajult hiperbola)	egyetlen egyenes (elfajult ellipszis)	az origó (elfajult parabola)

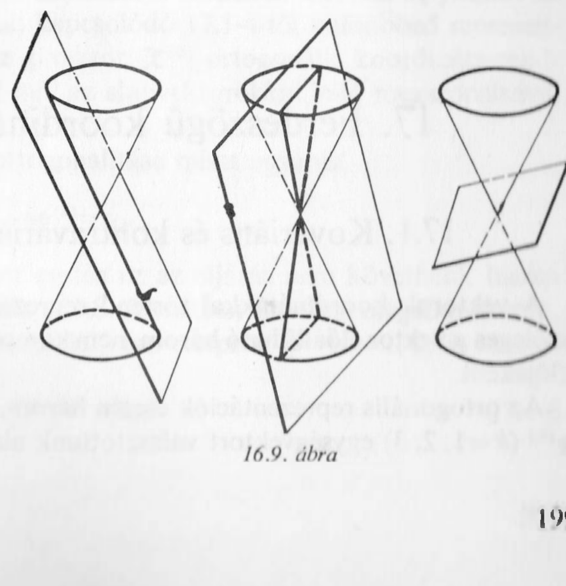
Az egyes elfajuló esetek nyilvánvalóan akkor jönnek létre, ha:

a) a metsző sík éppen érinti a kúpot. (Az ekkor létrejövő egyenes éppen a kúp alkotója);

b) a metsző sík a kúp tengelyét is tartalmazza, az egyenespár két, egymást az origóban metsző alkotó;

c) a sík a kúp szögtartományán kívül halad, akkor a sík és a kúp közös része az origó.

Az egyes eseteket a *16.9. ábra* mutatja.



16.9. ábra

Messük el a 16.2-7-nek megfelelő

$$\underline{x}=u_1\underline{e}^{(1)}+u_2\underline{e}^{(2)}+U\underline{e}^{(3)} \qquad 16.3\text{-}1$$

egyenletű síkkal az

$$\underline{x}\underline{\underline{A}}\underline{x}+2\underline{q}\underline{x}+c=0 \qquad 16.3\text{-}2$$

egyenletű másodrendű felületet.

16.3-1-et 16.3-2-be behelyettesítve, azt kapjuk, hogy

$$\sum_{k,\,l=1}^{2} u_k u_l \underline{e}^{(k)}\underline{\underline{A}}\underline{e}^{(l)}+2\sum_{k=1}^{2} u_k U\underline{e}^{(3)}\underline{\underline{A}}\underline{e}^{(k)}+U^2\underline{e}^{(3)}\underline{\underline{A}}\underline{e}^{(3)}+ \qquad 16.3\text{-}3$$

$$+2\sum_{k=1}^{2} u_k\underline{e}^{(k)}\underline{q}+U\underline{e}^{(3)}\underline{q}+c=0.$$

Vezessük be az

$$(\underline{\underline{A}}^{(2)})_{kl}=\underline{e}^{(k)}\underline{\underline{A}}\underline{e}^{(l)} \quad (k,l=1,2) \qquad 16.3\text{-}4$$

kétszer kettős mátrixot, valamint a

$$(q^{(2)})_k=\underline{e}^{(k)}\underline{q}+U\underline{e}^{(3)}\underline{\underline{A}}\underline{e}^{(k)} \quad (k=1,2) \qquad 16.3\text{-}5$$

kétdimenziós vektort. Ezzel 16.3-4 az

$$\underline{u}\underline{\underline{A}}^{(2)}\underline{u}+2\underline{q}^{(2)}\underline{u}+c^{(2)}=0 \qquad 16.3\text{-}6$$

alakban írható fel, ahol

$$c^{(2)}=U\underline{e}^{(3)}\underline{q}+c+U^2\underline{e}^{(3)}\underline{\underline{A}}\underline{e}^{(3)}.$$

16.3-6 megegyezik a másodrendű görbék általános egyenletével. Megállapítható tehát, hogy a másodrendű felületek síkmetszetei másodrendű görbék.

17. Ferdeszögű koordináta-rendszerek

17.1. Kovariáns és kontravariáns reprezentációk

A vektorok koordinátákkal történő reprezentációja azon alapul, hogy tetszőleges $\underline{a}$ vektor előállítható három, nem komplanáris vektor lineáris kombinációjaként.

Az ortogonális reprezentációk esetén három, páronként egymásra merőleges $\underline{e}^{(k)}$ $(k=1, 2, 3)$ egységvektort választottunk alapvektorként, s minden vektort azok lineáris kombinációjával állítottunk elő. Az $\underline{e}^{(k)}$ alapvektorok egy $\mathcal{K}^{(0)}$ ortogonális koordináta-rendszert határoznak meg, s ha

$$\underline{a} = \sum a_k \underline{e}^{(k)}, \qquad 17.1\text{-}1$$

akkor

$$\mathcal{K}^{(0)}(\underline{a}) = \mathbf{a}^{(0)} = a_1, a_2, a_3.$$

Az $\underline{a}$ vektor azonban tetszőleges három, nem komplanáris $\underline{f}^{(k)}$ $(k=1, 2, 3)$ vektor lineáris kombinációjaként is előállítható.

Az $\underline{f}^{(k)}$ $(k=1, 2, 3)$ vektorokra tehát fennáll az

$$\underline{f}^{(1)}(\underline{f}^{(2)} \times \underline{f}^{(3)}) = V \neq 0 \qquad 17.1\text{-}2$$

összefüggés. 17.1-2-ben V az $\underline{f}^{(k)}$ $(k=1, 2, 3)$ vektorok által kifeszített paralelepipedon térfogata.

Az $\underline{f}^{(k)}$ $(k=1, 2, 3)$ vektorok egyenesei egy ferdeszögű koordináta-rendszert feszítenek ki. Az $\underline{f}^{(k)}$ vektorok segítségével tetszőleges vektor az

$$\underline{a} = \sum_{k=1}^{3} \bar{a}_k \underline{f}^{(k)} \qquad 17.1\text{-}3$$

alakban írható fel. 17.1-3-ban az $\bar{a}_k$ tényezők a koordináta-rendszer választásától függően változnak. Általában azt mondhatjuk, hogy az $\underline{a}$ vektort a $\mathcal{K}$ koordináta-rendszerben az $\bar{a}_k$ $(k=1, 2, 3)$ komponensekkel jellemezhetjük. A szokásos jelölésekkel

$$\mathcal{K}(\underline{a}) = \bar{\mathbf{a}} = \bar{a}_1, \bar{a}_2, \bar{a}_3. \qquad 17.1\text{-}4$$

A derékszögű esethez hasonlóan a koordináta-rendszert jobbsodrásúnak tekintjük, ha

$$V > 0. \qquad 17.1\text{-}5$$

A komponenseket azért jelöljük felülhúzott betűkkel, mert az a_k jelölést az $\underline{a}$ vektornak a fentiekhez szorosan kapcsolódó 17.1-4-től különböző reprezentációja számára tartjuk fenn. Az $\underline{a}$ vektor $\mathcal{K}^{(0)}$ ortogonális koordináta-rendszerben vett komponensei 17.1-1-nek az alapvektorokkal való megszorzásával fejezhetők ki.

Az $\underline{e}^{(k)}$ $(k=1, 2, 3)$ vektorok ortogonalitása miatt ugyanis

$$\underline{a}\underline{e}^{(l)} = a_l.$$

Ferdeszögű koordináta-rendszer esetén ez az eljárás nem követhető, hiszen az $\underline{f}^{(k)}$ vektorok nem ortogonálisak. A 4.8. fejezetben azonban megállapítottuk, hogy létezik a 17.1-5 feltételnek megfelelő ún. reciprok $\bar{\underline{f}}^{(k)}$ $(k=1, 2, 3)$ vektorhármas, amelyre teljesül az

$$\underline{f}^{(k)} \bar{\underline{f}}^{(l)} = \delta_{kl} \qquad 17.1\text{-}6$$

összefüggés.

17.1-6-ot felhasználva, az $\mathbf{a}$ vektor $\mathcal{K}$-ban vett $\bar{a}_k$ komponenseit a 17.1-3 összefüggésből az $\bar{\mathbf{f}}^{(l)}$ reciprok vektorokkal való szorzással fejezhetjük ki. Tehát

$$\mathbf{a}\bar{\mathbf{f}}^{(l)} = \sum_k \bar{a}_k \mathbf{f}^{(k)} \bar{\mathbf{f}}^{(l)} \quad (l=1, 2, 3), \tag{17.1-7}$$

ahonnan

$$\mathbf{a}\bar{\mathbf{f}}^{(l)} = \bar{a}_l. \tag{17.1-8}$$

A 17.1-8 összefüggés mutatja, hogy a ferdeszögű koordináta-rendszerek esetén az alapvektorok reciprok vektorhármasa is jelentős szerepet játszik.

A reciprok vektorokról tudjuk, hogy nem eshetnek egy síkba, hiszen

$$\bar{\mathbf{f}}^{(1)}(\bar{\mathbf{f}}^{(2)} \times \bar{\mathbf{f}}^{(3)}) = \frac{1}{V}, \tag{17.1-9}$$

így az $\mathbf{a}$ vektor az $\bar{\mathbf{f}}^{(l)}$ $(l=1, 2, 3)$ vektorok lineáris kombinációjaként is előállítható.

Legyen

$$\mathbf{a} = \sum_l a_l \bar{\mathbf{f}}^{(l)}. \tag{17.1-10}$$

Az a_k komponensek a 17.1-10 összefüggésnek az $\mathbf{f}^{(k)}$ vektorokkal történő szorzásával fejezhetők ki. 17.1-6 miatt

$$\mathbf{a}\mathbf{f}^{(l)} = a_l. \tag{17.1-11}$$

A $\mathcal{K}$ ferdeszögű koordináta-rendszerben az $\mathbf{a}$ vektort tehát mind az $\bar{a}_1, \bar{a}_2, \bar{a}_3$, mind az a_1, a_2, a_3 számhármassal jellemezhetjük. A 17.1-8 előállítást az $\mathbf{a}$ vektor $\mathcal{K}$-beli *kontravariáns*, a 17.1-11-et pedig *kovariáns* reprezentációjának nevezzük. A kétféle reprezentáció jelölésére a továbbiakban az

$$\bar{\mathbf{a}} = \bar{a}_1, \bar{a}_2, \bar{a}_3 \tag{17.1-12}$$

és

$$\mathbf{a} = a_1, a_2, a_3 \tag{17.1-13}$$

szimbólumokat használjuk.

17.2. A kovariáns és kontravariáns reprezentációk geometriai jelentése

Megállapítottuk, hogy az

$$\mathbf{a} = \sum a_k \bar{\mathbf{f}}^{(k)} \tag{17.2-1}$$

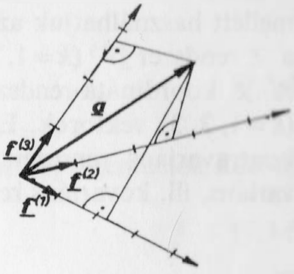
kifejezést $\mathbf{f}^{(l)}$-lel szorozva, az

$$\mathbf{a}\mathbf{f}^{(l)} = a_l \tag{17.2-2}$$

összefüggést kapjuk.

A 17.2-2 kifejezés bal oldala az $\underline{a}$ vektornak az $\underline{f}^{(l)}$ irányba eső vetületét adja megfelelő egységekben. Az $\underline{f}^{(l)}$ ($l=1, 2, 3$) vektorok nem egységvektorok, így 17.2-2 csak akkor egyezik meg az $\underline{a}$ vektor vetületével, ha távolságegységként az $|\underline{f}^{(l)}|=f^{(l)}$ mennyiséget választjuk. Amennyiben ezt nem tesszük meg, akkor a vetület értéke

$$\alpha_l=\frac{a_l}{f^{(l)}} \quad (l=1, 2, 3).$$



17.1. ábra

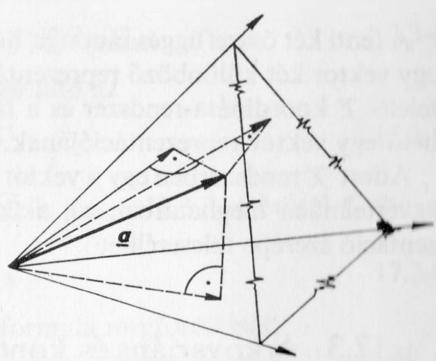
A kovariáns komponensek tehát az $\underline{a}$ vektornak a $\mathcal{X}$ rendszer tengelyeire vett vetületeit adják meg (*17.1. ábra*).

Ugyancsak megállapítottuk, hogy az

$$\underline{a}=\sum \bar{a}_k \underline{f}^{(k)} \qquad 17.2\text{-}3$$

formulát $\underline{\bar{f}}^{(l)}$-lel szorozva, az

$$(\underline{a}\underline{\bar{f}}^{(l)})=\bar{a}_l \qquad 17.2\text{-}4$$



17.2. ábra

összefüggéshez jutunk. Az $\bar{a}_l$ komponens tehát megfelelő mérték választása mellett az $\underline{a}$ vektornak az $\underline{\bar{f}}^{(l)}$ irányba eső vetülete. Mivel az $\underline{\bar{f}}^{(l)}$ vektor merőleges az $\underline{f}^{(m)}$, $\underline{f}^{(k)}$ ($\varepsilon_{klm}\neq 0$) vektorok síkjára, az a kontravariáns reprezentáció komponensei az $\underline{a}$ vektor végpontjának koordinátasíkoktól mért távolságait jelentik (*17.2. ábra*).

Adott vektor kovariáns reprezentációja tehát a tengelyekre vett vetületeket, kontravariáns reprezentációja pedig a vektor végpontjának a koordináta-rendszer alapsíkjaitól mért távolságait jelenti (megfelelő mértékek bevezetése esetén).

A fentiek alapján világossá válik, hogy derékszögű koordináta-rendszer esetén a kétféle reprezentáció nem választható szét, hiszen tetszőleges vektornak a koordinátatengelyekre vett vetületei megegyeznek a vektor végpontjának a megfelelő koordinátasíkoktól mért távolságaival.

Ez az egybeesés tulajdonképpen a derékszögű koordináta-rendszer alapvektorai ortogonalitásának következménye, hiszen

$$\underline{e}^{(k)}\underline{e}^{(l)}=\delta_{kl} \qquad 17.2\text{-}5$$

miatt az $\underline{e}^{(k)}$ ($k=1, 2, 3$) vektorhármas azonos saját reciprok vektorhármasával.

Megjegyzés:

Az $\underline{a}$, ill. $\bar{a}$ szimbólumok egy vektor kétféle reprezentációját jelölik. A két reprezentációt kovariánsnak, ill. kontravariánsnak nevezzük. Maga az $\underline{a}$ vektor azonban sem kovariáns, sem pedig kontravariáns, hanem egy olyan fizikai mennyiség, amelyet konzekvensen jellemezhetünk két különböző módszerrel is.

Némi félreértésre adhat okot időnként az, hogy a $\mathcal{K}$ koordináta-rendszer mellett használhatjuk azt a $\bar{\mathcal{K}}$ koordináta-rendszert is, amelynek alapvektorai a $\mathcal{K}$ rendszer $\underline{f}^{(k)}$ $(k=1, 2, 3)$ alapvektorainak $\bar{\underline{f}}^{(l)}$ $(l=1, 2, 3)$ reciprok vektorai. A $\bar{\mathcal{K}}$ koordináta-rendszer alapvektorainak reciprok vektorai az eredeti $\underline{f}^{(k)}$ $(k=1, 2, 3)$ vektorok. Ez azt jelenti, hogy az $\underline{a}$ vektor $\mathcal{K}$-beli kovariáns, ill. kontravariáns reprezentációja számértékében megegyezik a $\bar{\mathcal{K}}$-beli kontravariáns, ill. kovariáns reprezentációval. Tehát

$$\mathcal{K}(\underline{a})_{\text{kovariáns}}=\bar{\mathcal{K}}(\underline{a})_{\text{kontravariáns}},$$

ill.

$$\mathcal{K}(\underline{a})_{\text{kontravariáns}}=\bar{\mathcal{K}}(\underline{a})_{\text{kovariáns}}.$$

A fenti két összefüggés mutatja, hogy még ugyanazon számhármas is leírhatja egy vektor két különböző reprezentációját. Egy adott számhármas csak a megfelelő $\mathcal{K}$ koordináta-rendszer és a reprezentáció típusának ismeretében tekinthető egy vektor reprezentációjának.

Adott $\mathcal{K}$ rendszerben egy $\underline{a}$ vektor kovariáns és kontravariáns reprezentációja egyértelműen meghatározható, a $\bar{\mathcal{K}}$ rendszerre áttérve azonban a két reprezentáció szerepe felcserélhető.

17.3. A kovariáns és kontravariáns komponensek közötti összefüggés

Állítsuk elő az $\underline{a}$ vektort mind kovariáns, mind kontravariáns komponensek segítségével. Legyen

$$\underline{a}=\sum_{k=1}^{3} a_k\bar{\underline{f}}^{(k)}=\sum_{k=1} \bar{a}_k\underline{f}^{(k)}. \qquad 17.3\text{-}1$$

Szorozzuk be a 17.3-1 egyenlőséget $\underline{f}^{(l)}$-lel; az

$$a_l=\sum_{k=1}^{3} \bar{a}_k\underline{f}^{(k)}\underline{f}^{(l)} \qquad 17.3\text{-}2$$

összefüggéshez jutunk. 17.3-2 segítségével a kontravariáns reprezentációról áttérhetünk a kovariáns reprezentációra.

A kétféle reprezentáció közötti kapcsolatot az alapvektorok

$$\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}=f^{(k)}f^{(l)}\cos\vartheta_{kl}\,,\quad (k,\, l=1,\, 2,\, 3) \qquad 17.3\text{-}3$$

skalárszorzatai határozzák meg, ahol ϑ_{kl} a k-adik és l-edik alapvektor által bezárt szög. Mivel a szorzatok két indextől függenek, felfoghatók egy kétdimenziós mátrix elemeiként.

Legyen $\mathbf{G}$ egy mátrix, amelynek elemeit az

$$\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}=G_{kl}\quad (k,\, l=1,\, 2,\, 3) \qquad 17.3\text{-}4$$

formula határozza meg.

A $\mathbf{G}$ mátrix segítségével a kovariáns és kontravariáns reprezentációk közötti 17.3-2 kapcsolat az

$$\mathbf{a}=\bar{\mathbf{a}}\mathbf{G} \qquad 17.3\text{-}5$$

formában írható fel. A 17.3-4 formulából és a skaláris szorzás kommutativitásából azonnal következik, hogy $\mathbf{G}$ szimmetrikus mátrix, hiszen

$$\widetilde{G}_{kl}=G_{lk}=\underline{\mathbf{f}}^{(l)}\underline{\mathbf{f}}^{(k)}=\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}=G_{kl}. \qquad 17.3\text{-}6$$

A 17.3-5 összefüggés ennek megfelelően átírható az

$$\mathbf{a}=\mathbf{G}\bar{\mathbf{a}} \qquad 17.3\text{-}7$$

alakba.

Amennyiben a $\mathbf{G}$ mátrix $\mathbf{G}^{-1}$ inverze létezik, akkor 17.3-7-et balról $\mathbf{G}^{-1}$-gyel szorozva, következik, hogy az

$$\bar{\mathbf{a}}=\mathbf{G}^{-1}\mathbf{a} \qquad 17.3\text{-}8$$

egyenlőség adja a 17.3-7 transzformációs formula megfordítását.

A kontravariáns reprezentációt a 17.3-1 formula alapján is kifejezhetjük a kovariáns reprezentáció elemeivel. Szorozzuk be a 17.3-1 formulát $\bar{\underline{\mathbf{f}}}^{(l)}$-lel. Ekkor az

$$\bar{a}_l=\sum_k a_k\bar{\underline{\mathbf{f}}}^{(k)}\bar{\underline{\mathbf{f}}}^{(l)} \qquad 17.3\text{-}9$$

összefüggéshez jutunk. A transzformációt az $\bar{\underline{\mathbf{f}}}^{(k)}$ reciprok vektorok skaláris szorzatai határozzák meg.

Vezessük be a $\bar{\mathbf{G}}$ mátrixot, amelynek elemei:

$$\bar{G}_{kl}=\bar{\underline{\mathbf{f}}}^{(k)}\bar{\underline{\mathbf{f}}}^{(l)}\quad (k,\, l=1,\, 2,\, 3). \qquad 17.3\text{-}10$$

Mivel $\bar{\mathbf{G}}$ szimmetrikus mátrix, a 17.3-9 összefüggést $\bar{\mathbf{G}}$ segítségével az

$$\bar{\mathbf{a}}=\bar{\mathbf{G}}\mathbf{a} \qquad 17.3\text{-}11$$

alakban is felírhatjuk. Összehasonlítva a 17.3-8 és 17.3-11 formulákat, adódik hogy

$$\mathbf{G}^{-1}=\bar{\mathbf{G}}, \qquad 17.3\text{-}12$$

azaz a **G** transzformációs mátrix reciproka létezik, és megegyezik a 17.3-10 által definiált **G** mátrix elemeivel. Tehát

$$G^{+}_{kl} = \bar{G}_{kl} = \bar{\underline{\mathbf{f}}}^{(k)}\bar{\underline{\mathbf{f}}}^{(l)}, \qquad 17.3\text{-}13$$

ahol G^{+}_{kl}-lel az inverz mátrix elemeit jelöltük.

17.4. Vektorok összeadása ferdeszögű reprezentációkban

A ferdeszögű reprezentációkban megadott vektorok összegezésére a következő szabályok adódnak.

Legyen

$$\underline{\mathbf{a}} = \sum_{k=1}^{3} \bar{a}_k \underline{\mathbf{f}}^{(k)}, \quad \underline{\mathbf{b}} = \sum_{k=1}^{3} \bar{b}_k \underline{\mathbf{f}}^{(k)} \qquad 17.4\text{-}1$$

és

$$\underline{\mathbf{c}} = \underline{\mathbf{a}} + \underline{\mathbf{b}}. \qquad 17.4\text{-}2$$

A vektorösszegezés szabályait koordináta-rendszertől függetlenül vezettük be, ezért minden további nélkül alkalmazhatjuk az adott esetre is. 17.4-1-et beírva 17.4-2-be:

$$\underline{\mathbf{c}} = \sum_{k=1}^{3} \bar{a}_k \underline{\mathbf{f}}^{(k)} + \sum_{k=1}^{3} \bar{b}_k \underline{\mathbf{f}}^{(k)} = \sum_{k=1}^{3} (\bar{a}_k + \bar{b}_k) \underline{\mathbf{f}}^{(k)}. \qquad 17.4\text{-}3$$

17.4-3-at a

$$\bar{c}_k = \bar{a}_k + \bar{b}_k \quad (k = 1, 2, 3) \qquad 17.4\text{-}4$$

jelölés bevezetésével

$$\underline{\mathbf{c}} = \sum \bar{c}_k \underline{\mathbf{f}}^{(k)}$$

alakban is felírhatjuk. Tömörebben fogalmazva,

$$\bar{\underline{\mathbf{c}}} = \bar{\underline{\mathbf{a}}} + \bar{\underline{\mathbf{b}}}. \qquad 17.4\text{-}5$$

A fentiekhez hasonló módon az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok, valamint a $\underline{\mathbf{c}} = \underline{\mathbf{a}} + \underline{\mathbf{b}}$ összegvektor kovariáns reprezentációjára a

$$\mathbf{c} = \mathbf{a} + \mathbf{b}$$

összefüggés adódik.

Teljesen analóg módon látható be, hogy amennyiben

$$\alpha\underline{\mathbf{a}} + \beta\underline{\mathbf{b}} = \underline{\mathbf{c}}, \qquad 17.4\text{-}6$$

akkor tetszőleges reprezentációban

$$\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{c}$$

és

$$\alpha\bar{\mathbf{a}} + \beta\bar{\mathbf{b}} = \bar{\mathbf{c}}. \qquad 17.4\text{-}7$$

17.4-6 és 17.4-7 azt jelenti, hogy amennyiben a $\underline{\mathbf{c}}$ vektor az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok meghatározott módon vett lineáris kombinációja, akkor ez tetszőleges reprezentáció (koordinátarendszer-választás) esetén is fennáll.

A fenti szabályok egyszerűen kiterjeszthetők vektorok tetszőleges csoportosításaira is. Ennek részletezésével most nem foglalkozunk. Általánosan igaz az, hogy ferdeszögű koordináták esetén is ugyanazok az összegezési szabályok érvényesek, mint derékszögű esetben.

17.5. A skaláris szorzat ferdeszögű reprezentációja

Határozzuk meg az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok skaláris szorzatának értékét az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok kovariáns és kontravariáns reprezentációjának segítségével.

Legyen

$$\underline{\mathbf{a}} = \sum \bar{a}_k \underline{\mathbf{f}}^{(k)} = \sum a_l \underline{\bar{\mathbf{f}}}^{(l)} \qquad 17.5\text{-}1$$

és

$$\underline{\mathbf{b}} = \sum \bar{b}_k \underline{\mathbf{f}}^{(k)} = \sum b_l \underline{\bar{\mathbf{f}}}^{(l)}. \qquad 17.5\text{-}2$$

Foglalkozzunk először a kontravariáns reprezentációval. Ekkor

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \sum \bar{a}_k \bar{b}_l \underline{\mathbf{f}}^{(k)} \underline{\mathbf{f}}^{(l)}. \qquad 17.5\text{-}3$$

17.5-3-ból leolvasható, hogy a skaláris szorzat értékét az alapvektorok skaláris szorzatainak ismeretében határozhatjuk meg. A 17.5-3 formula a 17.3-4 által definiált **G** mátrix elemeinek felhasználásával az

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \sum \bar{a}_k G_{kl} \bar{b}_l, \qquad 17.5\text{-}4$$

illetve mátrixjelölésben az

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \bar{\mathbf{a}}\mathbf{G}\bar{\mathbf{b}} \qquad 17.5\text{-}5$$

alakot ölti.

Hasonló módon a kovariáns előállításból kiindulva kapjuk, hogy

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \sum a_k G^{+}_{kl} b_l, \qquad 17.5\text{-}6$$

ill.

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \mathbf{a}\mathbf{G}^{-1}\mathbf{b}. \qquad 17.5\text{-}7$$

Megállapítható tehát, hogy a skaláris szorzat a kovariáns és kontravariáns reprezentációkban a **G** mátrix segítségével írható fel.

Ha az $\underline{\mathbf{f}}^{(k)}$ vektorok páronként ortogonálisak, tehát derékszögű koordinátarendszerrel van dolgunk, akkor

$$G_{kl}=\delta_{kl}, \tag{17.5-8}$$

azaz

$$\mathbf{G}=\mathbf{E},$$

és így

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\bar{\mathbf{a}}\bar{\mathbf{E}}\bar{\mathbf{b}}=\mathbf{a}\mathbf{E}\mathbf{b}=\mathbf{ab}, \tag{17.5-9}$$

vagyis a skaláris szorzatra vonatkozó ismert összefüggéshez jutunk.

A 17.5-7 formulában $\mathbf{G}^{-1}\mathbf{b}$ 17.3-11 értelmében éppen a $\underline{\mathbf{b}}$ vektor $\bar{\mathbf{b}}$ kontravariáns reprezentációját adja, tehát a skaláris szorzat az

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\mathbf{a}\bar{\mathbf{b}},$$

illetve

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\sum a_k\bar{b}_k \tag{17.5-10}$$

formában is kifejezhető.

Hasonló módon az

$$\bar{\mathbf{a}}=\mathbf{G}^{-1}\mathbf{a}=\mathbf{a}\mathbf{G}^{-1}$$

összefüggés miatt a skaláris szorzat az

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\bar{\mathbf{a}}\mathbf{b},$$

illetve

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\sum \bar{a}_k b_k \tag{17.5-11}$$

alakban is felírható.

17.5-10 és 17.5-11 erősen emlékeztet a skaláris szorzat derékszögű komponensekben történő meghatározására, a különbség csak annyi, hogy az egyik vektor esetén a kovariáns, a másik esetén pedig a kontravariáns reprezentációt kell felhasználni.

17.6. A vektoriális szorzat ferdeszögű reprezentációja

Legyen

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=\underline{\mathbf{c}}. \tag{17.6-1}$$

Az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok kontravariáns reprezentációinak felhasználásával

$$\underline{\mathbf{a}}=\sum_k \bar{a}_k\underline{\mathbf{f}}^{(k)}, \tag{17.6-2}$$

$$\underline{\mathbf{b}}=\sum_l \bar{b}_l\underline{\mathbf{f}}^{(l)}. \tag{17.6-3}$$

17.6-2 és 17.6-3 felhasználásával a 17.6-1 vektoriális szorzat a

$$\underline{\mathbf{c}}=\sum_{k,\,l} \bar{a}_k\bar{b}_l(\underline{\mathbf{f}}^{(k)}\times\underline{\mathbf{f}}^{(l)}) \tag{17.6-4}$$

alakot ölti. Minthogy

$$(\underline{\mathbf{f}}^{(k)}\times\underline{\mathbf{f}}^{(l)})\underline{\mathbf{f}}^{(m)}=\varepsilon_{klm}V, \qquad 17.6\text{-}5$$

17.6-4-ből egyszerűen kifejezhetjük $\underline{\mathbf{c}}$ kovariáns komponenseit. Mivel

$$c_m=\underline{\mathbf{c}}\,\underline{\mathbf{f}}^{(m)},$$

17.6-4-ből 17.6-5 felhasználásával adódik, hogy

$$c_m=\sum_{k,\,l}\varepsilon_{klm}\bar{a}_k\bar{b}_lV. \qquad 17.6\text{-}6$$

17.6-6 a

$$\underline{\mathbf{c}}=(\bar{\underline{\mathbf{a}}}\times\bar{\underline{\mathbf{b}}})V \qquad 17.6\text{-}7$$

alakban írható fel összefoglaló módon. 17.6-7-ben $\bar{\underline{\mathbf{a}}}\times\bar{\underline{\mathbf{b}}}$ jelentése a szokásos, tehát

$$\left.\begin{aligned}
(\bar{\underline{\mathbf{a}}}\times\bar{\underline{\mathbf{b}}})_1&=\bar{a}_2\bar{b}_3-\bar{a}_3\bar{b}_2,\\
(\bar{\underline{\mathbf{a}}}\times\bar{\underline{\mathbf{b}}})_2&=\bar{a}_3\bar{b}_1-\bar{b}_3\bar{a}_1,\\
(\bar{\underline{\mathbf{a}}}\times\bar{\underline{\mathbf{b}}})_3&=\bar{a}_1\bar{b}_2-\bar{a}_2\bar{b}_1.
\end{aligned}\right\} \qquad 17.6\text{-}8$$

Ha a fenti meggondolást az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok

$$\underline{\mathbf{a}}=\sum_k a_k\bar{\underline{\mathbf{f}}}^{(k)} \quad \text{és} \quad \underline{\mathbf{b}}=\sum_l b_l\bar{\underline{\mathbf{f}}}^{(l)}$$

előállításaival, tehát a kovariáns reprezentációk segítségével végezzük el, akkor a

$$\bar{c}_m=\frac{1}{V}\sum_{k,l,}\varepsilon_{klm}a_kb_l \qquad 17.6\text{-}9$$

eredményre jutunk. 17.6-9 indexek nélkül a

$$\bar{\mathbf{c}}=(\mathbf{a}\times\mathbf{b})\,\frac{1}{V} \qquad 17.6\text{-}10$$

alakban fogalmazható meg.

A vektoriális szorzat ferdeszögű reprezentációja tehát a V és $\frac{1}{V}$ tényezőktől eltekintve a derékszögű komponenseknél megszokott módon képezhető, a kovariáns komponensekből kiindulva azonban a vektoriális szorzat kontravariáns komponenseihez, a kontravariáns komponensekből kiindulva pedig a vektoriális szorzat kovariáns komponenseihez jutunk.

Egységvektorokkal megadott ortogonális koordináták esetén mind 17.6-7, mind 17.6-10 a vektoriális szorzat szokásos reprezentációjába megy át, hiszen ebben az esetben

$$\mathbf{a}=\bar{\mathbf{a}}, \quad \mathbf{b}=\bar{\mathbf{b}} \quad \text{és} \quad V=1.$$

A 17.6-7 és 17.6-10 formulákban kényelmetlen, hogy a

$$V = \underline{\mathbf{f}}^{(1)}(\underline{\mathbf{f}}^{(2)} \times \underline{\mathbf{f}}^{(3)})$$

érték explicit módon szerepel. A következőkben megmutatjuk, hogy a V értéket kifejezhetjük a **G** mátrix segítségével.

17.3-7 szerint

$$\mathbf{a} = \mathbf{G}\bar{\mathbf{a}}, \quad \mathbf{b} = \mathbf{G}\bar{\mathbf{b}} \quad \text{és} \quad \mathbf{c} = \mathbf{G}\bar{\mathbf{c}}. \tag{17.6-11}$$

A 17.6-11 összefüggések felhasználásával 17.6-10-ben az **a** és **b** kovariáns reprezentációk a kontravariáns komponensekkel fejezhetők ki. Tehát

$$\bar{\mathbf{c}} = \frac{1}{V}(\mathbf{G}\bar{\mathbf{a}} \times \mathbf{G}\bar{\mathbf{b}}). \tag{17.6-12}$$

17.6-12-ről **G**-vel való szorzással térhetünk át $\underline{\mathbf{c}}$ kovariáns reprezentációjára, vagyis

$$\mathbf{c} = \mathbf{G}\bar{\mathbf{c}} = \frac{1}{V}\mathbf{G}(\mathbf{G}\bar{\mathbf{a}} \times \mathbf{G}\bar{\mathbf{b}}). \tag{17.6-13}$$

17.6-13-at komponensekben kiírva, a

$$c_m = \frac{1}{V} \sum_{K,L,M,k,l} \varepsilon_{KLM} G_{mM} G_{kK} G_{lL} \bar{a}_k \bar{b}_l \tag{17.6-14}$$

formulához jutunk. (17.6-14 felírásakor felhasználtuk, hogy **G** szimmetrikus mátrix.)

Amennyiben először a K, L, M szerinti összegezést végezzük el, akkor a determinánsokra vonatkozó

$$G = \det \mathbf{G} = \varepsilon_{klm} \sum_{K,L,M} \varepsilon_{KLM} G_{kK} G_{lL} G_{mM} \tag{17.6-15}$$

azonosság felhasználásával 17.6-14 a

$$c_m = \frac{1}{V} \det \mathbf{G} \sum_{kl} \varepsilon_{klm} \bar{a}_k \bar{b}_l \tag{17.6-16}$$

alakot ölti. 17.6-16 és 17.6-6 összehasonlításából adódik, hogy

$$\det \mathbf{G} = V^2, \tag{17.6-17}$$

azaz

$$V = \pm\sqrt{G}. \tag{17.6-18}$$

17.6-18-ban a négyzetgyök előjelét aszerint kell választani, hogy V pozitív vagy negatív. Tehát ha kizárólag jobbsodrású koordináta-rendszereket engedünk meg, akkor

$$V=\sqrt{G}.$$

A 17.7-18 összefüggés felhasználásával a vektoriális szorzat kovariáns és kontravariáns előállítása a

$$\underline{\mathbf{c}}=(\bar{\underline{\mathbf{a}}}\times\bar{\underline{\mathbf{b}}})\sqrt{G} \qquad 17.6\text{-}19$$

és

$$\bar{\underline{\mathbf{c}}}=(\underline{\mathbf{a}}\times\underline{\mathbf{b}})\,\frac{1}{\sqrt{G}} \qquad 17.6\text{-}20$$

alakban fogalmazható meg.

17.7. Tenzorok kovariáns és kontravariáns reprezentációja

A tenzorok ferdeszögű reprezentációjának előállításához a tenzor valamely általános tulajdonságából kell kiindulnunk. A kétdimenziós (két indexű) tenzorok esetén egyszerűen célhoz érünk, ha felhasználjuk azt, hogy egy tenzor egy homogén lineáris vektortranszformációt jelent, tehát tetszőleges $\underline{\mathbf{a}}^{(1)}$, $\underline{\mathbf{a}}^{(2)}$ vektorok esetén

$$\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}^{(i)}=\underline{\mathbf{b}}^{(i)} \quad (i=1,2) \qquad 17.7\text{-}1$$

és

$$\underline{\underline{\mathbf{A}}}(\alpha_1\underline{\mathbf{a}}^{(1)}+\alpha_2\underline{\mathbf{a}}^{(2)})=\alpha_1\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}^{(1)}+\alpha_2\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}^{(2)}. \qquad 17.7\text{-}2$$

Legyenek a $\mathcal{K}$ koordináta-rendszer alapvektorai az $\underline{\mathbf{f}}^{(k)}$ $(k=1,2,3)$ nem komplanáris vektorok.

Legyen

$$\underline{\mathbf{b}}=\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}} \qquad 17.7\text{-}3$$

és

$$\underline{\mathbf{a}}=\sum_k \bar{a}_k\underline{\mathbf{f}}^{(k)}. \qquad 17.7\text{-}4$$

17.7-1-ből és 17.7-2-ből következik, hogy

$$\underline{\mathbf{b}}=\underline{\underline{\mathbf{A}}}\underline{\mathbf{a}}=\sum_k \bar{a}_k(\underline{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)}). \qquad 17.7\text{-}5$$

17.7-5-öt balról beszorozva $\underline{\mathbf{f}}^{(l)}$-lel, 17.2-2 értelmében a $\underline{\mathbf{b}}$ vektor

$$b_l=\sum_k \bar{a}_k\underline{\mathbf{f}}^{(l)}(\underline{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)}) \qquad 17.7\text{-}6$$

kovariáns komponenseihez jutunk.[1] Bevezetve az $\mathbf{A}$ mátrixot, amelynek komponenseit az

[1] A továbbiakban $\underline{\mathbf{f}}^{(k)}(\underline{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)})$ helyett az egyszerűség kedvéért $\underline{\mathbf{f}}^{(l)}\underline{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)}$-t írunk.

$$A_{lk} = \underline{\mathbf{f}}^{(l)}\underline{\mathbf{A}}\underline{\mathbf{f}}^{(k)} \qquad 17.7\text{-}7$$

összefüggések határozzák meg, 17.7-6 a

$$b_l = \sum_k A_{lk}\bar{a}_k \qquad 17.7\text{-}8$$

alakot ölti. 17.7-8-at mátrixjelöléssel a

$$\mathbf{b} = \mathbf{A}\bar{\mathbf{a}} \qquad 17.7\text{-}9$$

alakban is felírhatjuk.

Az $\underline{\mathbf{A}}$ tenzor $\mathcal{K}^{(0)}$ ortogonális reprezentációjának komponenseit az

$$A_{lk}^{(0)} = \underline{\mathbf{e}}^{(k)}\underline{\mathbf{A}}\underline{\mathbf{e}}^{(l)} \qquad 17.7\text{-}10$$

formula határozza meg. A 17.7-10 által meghatározott kilenc komponens segítségével tetszőleges $\underline{\mathbf{a}}$ vektor $\underline{\mathbf{A}}\underline{\mathbf{a}}$ transzformáltjának komponensei meghatározhatók. 17.7-8-ból látható, hogy tökéletesen alkalmas erre a célra a 17.7-7-nek megfelelő kilenc komponens is. Az

$$\mathbf{A} = \mathcal{K}(\underline{\mathbf{A}})$$

mátrix az $\underline{\mathbf{A}}$ tenzor ferdeszögű reprezentációja. Mivel $\mathbf{A}$ előállítására az $\underline{\mathbf{f}}^{(l)}$ ($l = 1, 2, 3$) vektorokat használtuk, a vektorreprezentációkhoz hasonlóan $\mathbf{A} = \mathcal{K}(\underline{\mathbf{A}})$-t $\underline{\mathbf{A}}$ *kovariáns reprezentációjának tekintjük.*

Induljunk ki most az $\underline{\mathbf{a}}$ vektor

$$\underline{\mathbf{a}} = \sum_k a_k\bar{\underline{\mathbf{f}}}^{(k)} \qquad 17.7\text{-}11$$

előállításából. Ennek segítségével az előzőhöz hasonló gondolatmenettel a

$$\bar{b}_l = \sum_k a_k\bar{\underline{\mathbf{f}}}^{(l)}\underline{\mathbf{A}}\bar{\underline{\mathbf{f}}}^{(k)} \qquad 17.7\text{-}12$$

összefüggéshez jutunk. Bevezetve az $\bar{\mathbf{A}}$ mátrixot, amelynek elemeit az

$$\bar{A}_{lk} = \bar{\underline{\mathbf{f}}}^{(l)}\underline{\mathbf{A}}\bar{\underline{\mathbf{f}}}^{(k)} \quad (k, l = 1, 2, 3) \qquad 17.7\text{-}13$$

összefüggések adják, 17.7-12 a

$$\bar{b}_l = \sum \bar{A}_{lk}a_k, \qquad 17.7\text{-}14$$

ill. mátrixjelölésben a

$$\bar{\mathbf{b}} = \bar{\mathbf{A}}\mathbf{a} \qquad 17.7\text{-}15$$

alakot ölti.

Az $\bar{\mathbf{A}}$ mátrix az $\underline{\mathbf{A}}$ tenzor $\mathcal{K}$ koordináta-rendszerben vett *kontravariáns reprezentációja.* A kontravariáns reprezentáció jelölése félreértésre adhat okot, hiszen egy mátrix adjungáltját hasonló módon jelöltük. A következőkben ennek elkerülésére új jelölést vezetünk be.

17.8. A tenzorreprezentációk Einstein-féle jelölésmódja

A 17.7. fejezetben bevezettük az $\underline{\mathbf{A}}$ tenzor

$$A_{kl}=\mathbf{f}^{(k)}\underline{\mathbf{A}}\mathbf{f}^{(l)} \qquad 17.8\text{-}1$$

kovariáns és az

$$\bar{A}_{kl}=\bar{\mathbf{f}}^{(k)}\underline{\mathbf{A}}\bar{\mathbf{f}}^{(l)} \qquad 17.8\text{-}2$$

kontravariáns reprezentációját.

A kovariáns, ill. kontravariáns jelleg attól függ, hogy az $\underline{\mathbf{A}}$ tenzor $\underline{\mathbf{f}}^{(k)}$ ($k=1$, 2, 3) vagy $\bar{\underline{\mathbf{f}}}^{(l)}$ ($l=1, 2, 3$) vektorokkal vett szorzatait használjuk az $\underline{\mathbf{A}}$ tenzor jellemzésére. A mátrixjelölésben az A_{kl} mátrix indexei jelentik, hogy melyik alapvektorral szoroztunk, az eddigi jelölésben azonban az indexek nem tartalmaznak semmiféle utalást arra, hogy az $\underline{\mathbf{f}}^{(k)}$ vagy $\bar{\underline{\mathbf{f}}}^{(l)}$ vektorokat használtuk az adott komponens meghatározására.

Ennek jelölésére Einstein a következő konvenciót javasolta: Jelöljük *a kovariáns komponenseket alsó, a kontravariáns komponenseket* pedig *felső* indexszel. Így tehát az $\underline{\mathbf{a}}$ vektor reprezentációit az

$$\bar{\mathbf{a}}=a^1, a^2, a^3, \qquad 17.8\text{-}3$$

$$\mathbf{a}=a_1, a_2, a_3 \qquad 17.8\text{-}4$$

szimbólumokkal, a 17.8-1 és 17.8-2 tenzorkomponenseket pedig az

$$\underline{\mathbf{f}}^{(k)}\underline{\mathbf{A}}\underline{\mathbf{f}}^{(l)}=A_{kl} \qquad 17.8\text{-}5$$

és

$$\bar{\underline{\mathbf{f}}}^{(k)}\underline{\mathbf{A}}\bar{\underline{\mathbf{f}}}^{(l)}=A^{kl} \qquad 17.8\text{-}6$$

szimbólumokkal jelöljük.

Ennek megfelelően az $\underline{\mathbf{a}}$ vektort az

$$\underline{\mathbf{a}}=\sum_k a^k\underline{\mathbf{f}}^{(k)}=\sum_l a_l\bar{\underline{\mathbf{f}}}^{(l)} \qquad 17.8\text{-}7$$

alakban írhatjuk fel.

Az Einstein-féle jelölésmód segítségével az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok skaláris szorzata az

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\sum_k a_k b^k=\sum_l a^l b_l \qquad 17.8\text{-}8$$

alakban írható fel. A jelölés további egyszerűsítésére Einstein javasolta, hogy azokban a formulákban, ahol egy kovariáns és egy kontravariáns index megegyezik, a szumma jelet is hagyjuk el, s használjuk az

$$a_k b^k = \sum_k a_k b^k \qquad 17.8\text{-}9$$

jelölést. Ezt — bár a szakirodalomban meglehetősen elterjedt — nem használjuk, inkább arra törekszünk, hogy a formulákat mátrixjelölésben fogalmazzuk meg, ami az indexeket is feleslegessé teszi.

17.9. A G mátrix tulajdonságai

A **G** mátrix elemeit a

$$G_{kl} = \underline{\mathbf{f}}^{(k)} \underline{\mathbf{f}}^{(l)} \qquad 17.9\text{-}1$$

formulával definiáltuk, valamint bebizonyítottuk, hogy

$$\overline{\mathbf{G}} = \mathbf{G}^{-1}$$

és

$$G^{+}_{kl} = \bar{\underline{\mathbf{f}}}^{(k)} \bar{\underline{\mathbf{f}}}^{(l)}. \qquad 17.9\text{-}2$$

A 17.9-1 és 17.9-2 formulákat átírhatjuk a

$$G_{kl} = \underline{\mathbf{f}}^{(k)} \underline{\underline{\mathbf{E}}}\, \underline{\mathbf{f}}^{(l)} \qquad 17.9\text{-}3$$

és

$$G^{+}_{kl} = \bar{\underline{\mathbf{f}}}^{(k)} \underline{\underline{\mathbf{E}}}\, \bar{\underline{\mathbf{f}}}^{(l)} \qquad 17.9\text{-}4$$

alakba is, ahol $\underline{\underline{\mathbf{E}}}$ az egységtenzor. 17.9-3-at és 17.9-4-et a 17.8-6 formulákkal összehasonlítva, leolvasható, hogy **G** az $\underline{\underline{\mathbf{E}}}$ egységtenzor kovariáns, $\mathbf{G}^{-1}$ pedig kontravariáns reprezentációja. Ily módon az Einstein-féle konvenciónak megfelelően G^{+}_{kl} helyet a

$$G^{+}_{kl} = G^{kl} \qquad 17.9\text{-}5$$

jelölést is használhatjuk annak kifejezésére, hogy $\mathbf{G}^{-1}$ az egységtenzor kontravariáns reprezentációja.

A ferdeszögű reprezentáció kihangsúlyozására szokásos a $\underline{\underline{\mathbf{G}}}$ jelölés és a *metrikus tenzor* elnevezés.

17.10. Kevert reprezentációk

Nagyon hasznos, és az eddigiek alapján kézenfekvő a tenzorok ún. kevert reprezentációinak bevezetése az

$$A^{k}{}_{l} = \bar{\underline{\mathbf{f}}}^{(k)} \underline{\underline{\mathbf{A}}}\, \underline{\mathbf{f}}^{(l)}, \qquad 17.10\text{-}1$$

ill.

$$A_{k}{}^{l} = \underline{\mathbf{f}}^{(k)} \underline{\underline{\mathbf{A}}}\, \bar{\underline{\mathbf{f}}}^{(l)} \qquad 17.10\text{-}2$$

finíciók segítségével. 17.10-1-ben és 17.10-2-ben a tenzorkomponensek meghatározására az $\underline{\mathbf{f}}^{(k)}$ és $\underline{\bar{\mathbf{f}}}^{(l)}$ alapvektorokat együttesen használjuk. 17.10-1 az $\underline{\mathbf{A}}$ tenzor balról kontravariáns, jobbról kovariáns reprezentációja, 17.10-2 pedig a balról kovariáns, jobbról kontravariáns reprezentáció.

A kevert reprezentációk vektorok transzformáltjainak meghatározásakor hasznosak. A 17.7-8 és 17.7-14 formulák szerint ugyanis

$$b_l = \sum_k A_{lk} a^k, \qquad 17.10\text{-}3$$

és

$$b^l = \sum_k A^{lk} a_k, \qquad 17.10\text{-}4$$

vagyis az $\underline{\mathbf{A}}$ tenzor tiszta kovariáns reprezentációja segítségével $\underline{\mathbf{a}}$ kontravariáns reprezentációjából $\underline{\mathbf{b}}$ kovariáns reprezentációja határozható meg, $\underline{\mathbf{A}}$ tiszta kontravariáns előállítása pedig $\underline{\mathbf{a}}$ kovariáns komponenseiből $\underline{\mathbf{b}}$ kontravariáns komponenseit adja.

Ez a reprezentációváltás sok esetben kényelmetlen. A kevert reprezentációk jelentésének megértésére célszerű a kovariáns és kontravariáns reprezentációk bevezetésekor követett gondolatmenet rövid ismétlése.

Induljunk ki az

$$\underline{\mathbf{a}} = \sum a_k \underline{\bar{\mathbf{f}}}^{(k)}$$

előállításából. Ekkor

$$\underline{\mathbf{b}} = \sum a_k \underline{\mathbf{A}} \underline{\bar{\mathbf{f}}}^{(k)}. \qquad 17.10\text{-}5$$

17.10-5-ben $\underline{\mathbf{a}}$ kovariáns komponensei szerepelnek. Amennyiben $\underline{\mathbf{b}}$ kovariáns komponenseit kívánjuk meghatározni, 17.10-5-öt az $\underline{\mathbf{f}}^{(l)}$ ($l = 1, 2, 3$) vektorokkal kell megszorozni, tehát

$$b_l = \underline{\mathbf{b}} \underline{\mathbf{f}}^{(l)} = \sum_k a_k \underline{\mathbf{f}}^{(l)} \underline{\mathbf{A}} \underline{\bar{\mathbf{f}}}^{(k)}. \qquad 17.10\text{-}6$$

17.10-6 a 17.10-2 kevert reprezentáció felhasználásával a

$$b_l = \sum_k A_l^{\cdot k} a_k \qquad 17.10\text{-}7$$

alakot ölti.[1] Hasonló módon adódik, hogy a kontravariáns reprezentációk között a

$$b^l = \sum_k A_{\cdot k}^{l} a^k \qquad 17.10\text{-}8$$

formula létesít kapcsolatot.

Megjegyzés: Kevert reprezentációk esetén ügyelnünk kell az indexek sorrendjére, hiszen az $A_l^{\cdot k}$ és $A_{\cdot l}^{k}$ mátrixok nem azonosak.

[1] A formulákban szereplő pont az alsó és felső indexek sorrendjét szabja meg. Ezt a jelölést Einstein nem használta. A sorrendtévesztésekből adódó félreértések miatt azonban szükségesnek láttuk az Einstein-féle jelölés fenti kiegészítését.

17.11. A tenzorok kovariáns, kontravariáns és vegyes reprezentációi közötti összefüggés

Az előző fejezetben láttuk, hogy a kovariáns és kontravariáns reprezentációk mellett a vegyes reprezentációk is hasznos eszközei a tenzorokkal végzett konkrét számításoknak.

Hangsúlyozni kívánjuk azonban, hogy a reprezentáció megválasztása tisztán technikai kérdés, mert egy tenzor különböző reprezentációi mindig ugyanazt a fizikai mennyiséget jelentik. Ezért a fizikai összefüggéseket célszerű reprezentációtól független formában kifejezni.

A reprezentációk azonban nélkülözhetetlenek a konkrét számítások elvégzésekor, így feltétlenül szükséges, hogy ismerjük egy tenzor kovariáns, kontravariáns és vegyes reprezentációi közötti átszámítási szabályokat.

Mivel a különböző reprezentációk az $\underline{\bar{\mathbf{f}}}^{(k)}$ $(k=1, 2, 3)$ és $\underline{\mathbf{f}}^{(l)}$ $(l=1, 2, 3)$ vektorok segítségével határozhatók meg, az átszámítási szabályok a kétféle vektorhármas közötti összefüggésből határozhatók meg.

Állítsuk elő pl. az $\underline{\bar{\mathbf{f}}}^{(k)}$ reciprok vektorokat az $\underline{\mathbf{f}}^{(l)}$ $(l=1, 2, 3)$ vektorok lineáris kombinációjaként.

Legyen

$$\underline{\bar{\mathbf{f}}}^{(k)}=\sum_l P_{kl}\underline{\mathbf{f}}^{(l)}. \qquad 17.11\text{-}1$$

Szorozzuk be 17.11-1-et $\underline{\bar{\mathbf{f}}}^{(l)}$-lel. A

$$G^{+}_{kl}=G^{kl}=P_{kl} \qquad 17.11\text{-}2$$

eredményre jutunk. A két vektorhármas közötti összefüggés tehát az

$$\underline{\bar{\mathbf{f}}}^{(k)}=\sum G^{kl}\underline{\mathbf{f}}^{(l)}, \qquad 17.11\text{-}3$$

illetve

$$\underline{\mathbf{f}}^{(l)}=\sum G_{kl}\underline{\bar{\mathbf{f}}}^{(k)} \qquad 17.11\text{-}4$$

alakban írható fel.

17.11-3 és 17.11-4 felhasználásával egyszerűen meghatározhatjuk egy tenzor különböző reprezentációi között fennálló összefüggéseket.

Induljunk ki például az

$$A_{kl}=\underline{\mathbf{f}}^{(k)}\underline{\mathbf{A}}\underline{\mathbf{f}}^{(l)} \qquad 17.11\text{-}5$$

tiszta kovariáns reprezentációból. Helyettesítsük be 17.11-5-ben $\underline{\mathbf{f}}^{(k)}$ helyére a 17.11-4-ből adódó összefüggést. Az

$$A_{kl}=\sum_m G_{km}\underline{\bar{\mathbf{f}}}^{(m)}\underline{\mathbf{A}}\underline{\mathbf{f}}^{(l)} \qquad 17.11\text{-}6$$

formulához jutunk. 17.11-6-ban azonban $\underline{\mathbf{A}}$ vegyes reprezentációja szerepel, tehát

$$A_{kl}=\sum_{m} G_{km}A^{m}_{\cdot l}, \qquad 17.11\text{-}7$$

Amennyiben 17.11-5-ben $\underline{\mathbf{f}}^{(l)}$-t is kifejezzük az $\bar{\underline{\mathbf{f}}}^{(k)}$ $(k=1, 2, 3)$ reciprok vektorok segítségével, az

$$A_{kl}=\sum_{m,n} G_{km}G_{ln}\bar{\underline{\mathbf{f}}}^{(m)}\underline{\mathbf{A}}\bar{\underline{\mathbf{f}}}^{(n)}, \qquad 17.11\text{-}8$$

illetve

$$A_{kl}=\sum_{m,n} G_{km}G_{ln}A^{mn} \qquad 17.11\text{-}9$$

összefüggéshez jutunk. 17.11-9 az $\underline{\mathbf{A}}$ tenzor tisztán kovariáns és kontravariáns komponensei között létesít kapcsolatot.

Az

$$A^{kl}=\bar{\underline{\mathbf{f}}}^{(k)}\underline{\mathbf{A}}\bar{\underline{\mathbf{f}}}^{(l)}$$

kontravariáns reprezentációból kiindulva, a 17.11-3 összefüggés felhasználásával az

$$A^{kl}=\sum_{m} G^{km}A^{\cdot l}_{m}, \qquad 17.11\text{-}10$$

illetve

$$A^{kl}=\sum_{m,n} G^{km}G^{ln}A_{mn} \qquad 17.11\text{-}11$$

formulákhoz jutunk. Hasonló módon adódnak az

$$A^{\cdot l}_{k}=\sum_{m} G^{lm}A_{km} \qquad 17.11\text{-}12$$

és

$$A^{\cdot l}_{k}=\sum_{m} G_{km}A^{ml} \qquad 17.11\text{-}13$$

összefüggések is.

Megjegyzés: A 17.11-9 és 17.11-11 formulák felírásakor kihasználtuk a $\mathbf{G}$ mátrix szimmetriáját. A teljesség kedvéért felírjuk még a vektorok kovariáns és kontravariáns komponensei között fennálló 17.3-7 és 17.3-8 összefüggéseket az Einstein-féle jelölésmóddal:

$$\left.\begin{aligned} a_k&=\sum_{l} G_{kl}a^{l} \\ \text{és} \qquad a^{l}&=\sum_{m} G^{lm}a_{m}. \end{aligned}\right\} \qquad 17.11\text{-}14$$

A 17.11-7, 17.11-9, 17.11-10, 17.11-11, 17.11-12, 17.11-13 és 17.11-14 összefüggések mutatják, hogy a különböző reprezentációk között a $\underline{\mathbf{G}}\equiv\underline{\mathbf{E}}$ tenzor megfelelő $\mathbf{G}$ és $\mathbf{G}^{-1}$ reprezentációival való szorzás létesít kapcsolatot. A transzformációs szabályok az Einstein-féle jelölésmódban indexek „le- és felhúzásának" tűnnek.

A **G**-*vel való szorzás ugyanis mindig egy kontravariáns (felső) indexből képez kovariáns (alsó) indexet, a* $\mathbf{G}^{-1}$-*gyel való szorzás pedig egy kovariáns indexből képez kontravariáns indexet.*

Példaként alkalmazzuk a fenti szabályt a $\underline{\mathbf{G}}$ metrikus tenzorra. Határozzuk meg először a metrikus tenzor vegyes reprezentációját. 17.11-12 segítségével adódik, hogy

$$G_k^{\cdot l} = \sum_m G^{lm} G_{km}. \qquad 17.11\text{-}15$$

17.11-15 azonban 17.9-5 miatt éppen az egységmátrix, tehát

$$G_k^{\cdot l} = \delta_k^l. \qquad 17.11\text{-}16$$

Hasonló módon

$$G^l_{\cdot k} = \delta_k^l, \qquad 17.11\text{-}17$$

ebben az esetben tehát az indexek sorrendje tetszőleges.

Példaként alkalmazzuk még **G**-re a 17.11-11 összefüggést. A már ismert

$$\overline{\mathbf{G}} = \mathbf{G}^{-1}\mathbf{G}^{-1}\mathbf{G} = \mathbf{G}^{-1} \qquad 17.11\text{-}18$$

formulához jutunk.

17.12. A tenzorműveletek mátrixjelölése

Az előző fejezetben meghatároztuk a tenzorok kovariáns és kontravariáns reprezentációi közötti összefüggést. Megállapítottuk, hogy a **G**-vel való szorzás a kontravariáns mennyiségekből kovariáns mennyiséget, a $\mathbf{G}^{-1}$-gyel történő szorzás pedig a kovariáns mennyiségekből kontravariáns mennyiséget képez.

Az átszámítási szabály rendkívül egyszerű és jól definiált, így fölöslegesnek tűnik, hogy minden egyes formulába újra és újra belefoglaljuk az átszámítási törvényt is.

Az aláhúzott félkövér betűkkel reprezentációtól függetlenül magukat a fizikai mennyiségeket jelöltük. Az ily módon megfogalmazott törvények mindig a fizikai mennyiségek közötti kapcsolatot, tehát valamilyen természeti törvényt jelentenek.

Időnként szükséges azonban a formulákban is figyelmeztetni arra, hogy a számítást ferdeszögű koordináta-rendszerben kell elvégezni. Ebben az esetben javasoljuk, hogy tegyünk a szorzandó mennyiségek közé pontot. Az egyszerű jelölés arra figyelmeztet, hogy a műveletet az adott reprezentációban értelemszerűen kell elvégezni.

Tehát pl.

$$\underline{a}b = a \cdot b$$

azt jelenti, hogy az $\underline{a}$ és $\underline{b}$ vektorok skaláris szorzatát ferdeszögű reprezentációban kell meghatároznunk. Erre az

$$a \cdot b = \begin{cases} \bar{\underline{a}}\mathbf{G}\bar{b} \\ a\underline{b} \\ \underline{a}b \\ a\mathbf{G}^{-1}b \end{cases} \qquad 17.12\text{-}1$$

formulák közül bármelyik alkalmas. A „pont" tehát itt azt jelenti, hogy az $\underline{a}$ és $\underline{b}$ vektorok reprezentációi közé az $\underline{E}$ egységtenzorok megfelelő reprezentációját kell elhelyeznünk. (Kontravariáns vektorkomponensek közé a kovariáns **G** mátrixot, kovariáns vektorkomponensek közé a $\mathbf{G}^{-1}$ mátrixot kell tennünk.)

Vektorok és tenzorok szorzatának képzésekor hasonló módon járhatunk el:

$$a \cdot A \qquad 17.12\text{-}2$$

azt jelenti, hogy **E**-t, **G**-t vagy $\mathbf{G}^{-1}$-t kell elhelyeznünk a vektor- és tenzorreprezentáció között aszerint, hogy $\underline{a}$ és $\underline{A}$ mely reprezentációjában dolgozunk.

A fenti jelölést azért tartjuk célszerűnek, mert véleményünk szerint a számítások lényege homályosodik el, ha az általánosan érvényes formulákat a szorzási szabályok állandó részletezésével bonyolítjuk.

17.13. Koordináta-transzformációk

Végül a ferdeszögű koordináta-rendszerek transzformációjával foglalkozunk. Meghatározzuk tehát, hogy hogyan változik a vektorok és tenzorok reprezentációja, ha a $\mathcal{K}$ ferdeszögű koordináta-rendszerről a $\mathcal{K}'$ koordináta-rendszerre térünk át.

Legyenek a $\mathcal{K}$ rendszer alapvektorai az $\underline{f}^{(k)}$ $(k=1, 2, 3)$ vektorok, a $\mathcal{K}'$ rendszer alapvektorai pedig az $\underline{f}^{(l)'}$ $(l=1, 2, 3)$ vektorok. Mivel sem az $\underline{f}^{(k)}$, sem az $\underline{f}^{(l)'}$ vektorok nem komplanárisak, az $\underline{f}^{(l)'}$ vektorokat kifejezhetjük az $\underline{f}^{(k)}$ vektorok lineáris kombinációjaként.

Legyen

$$\underline{f}^{(l)'} = \sum_{k=1}^{3} S_{lk}\underline{f}^{(k)} \quad (l=1, 2, 3). \qquad 17.13\text{-}1$$

A 17.13-1 összefüggés megfordítását pedig az

$$\underline{\mathbf{f}}^{(l)} = \sum_{k=1} S^{+}_{lk} \underline{\mathbf{f}}^{(k)'} \quad (l=1, 2, 3) \qquad 17.13\text{-}2$$

formulák szolgáltatják.

Határozzuk meg most az $\underline{\mathbf{a}}$ vektor

$$\mathcal{K}(\underline{\mathbf{a}}) = \bar{\mathbf{a}} = a^1, a^2, a^3$$

és

$$\mathcal{K}'(\underline{\mathbf{a}}) = \bar{\mathbf{a}}' = a^{1'}, a^{2'}, a^{3'}$$

kontravariáns reprezentációi között fennálló összefüggést.

Tudjuk, hogy

$$\underline{\mathbf{a}} = \sum_k a^k \underline{\mathbf{f}}^{(k)} = \sum a^{l'} \underline{\mathbf{f}}^{(l)'}. \qquad 17.13\text{-}3$$

Behelyettesítve ide a 17.13-2 összefüggést, a

$$\sum_l a^{l'} \underline{\mathbf{f}}^{(l)'} = \sum_{k,l} a^k S^{+}_{kl} \underline{\mathbf{f}}^{(l)'} \qquad 17.13\text{-}4$$

formulához jutunk. 17.13-4-et $\bar{\underline{\mathbf{f}}}^{(m)'}$-vel végigszorozva, az

$$a'^m = \sum_k a^k S^{+}_{km} \qquad 17.13\text{-}5$$

összefüggés adódik.

17.13-5 mátrixjelölésben az

$$\bar{\mathbf{a}}' = \bar{\mathbf{a}}\mathbf{S}^{-1} = \tilde{\mathbf{S}}^{-1}\bar{\mathbf{a}} \qquad 17.13\text{-}6$$

alakot ölti. 17.13-6-ból a transzformáció

$$\bar{\mathbf{a}} = \bar{\mathbf{a}}'\mathbf{S} = \tilde{\mathbf{S}}\bar{\mathbf{a}}' \qquad 17.13\text{-}7$$

megfordítását **S**-sel való szorzással kaptuk. 17.13-6 mutatja, hogy a $\mathcal{K} \to \mathcal{K}'$ transzformáció során az a^l vektorkomponensek az $\tilde{\mathbf{S}}^{-1}$ mátrix segítségével transzformálhatók.

Az alapvektorok közötti kapcsolatot 17.13-1 szerint az **S** mátrix elemeivel adhatjuk meg.

A koordináták közötti 17.13-6, 7 transzformáció tehát a fordítottja az alapvektorok közötti 17.13-1, 2 kapcsolatnak. Ezért nevezzük az a^l vektorkomponenseket *kontravariáns* (ellentétesen változó) komponenseknek.

A metrikus tenzor komponenseinek

$$G_{kl} = \underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}$$

definíciója alapján egyszerűen megkaphatjuk a transzformáció után adódó $\underline{\underline{\mathbf{G}}}'$ tenzor elemeit. Definíció szerint

$$G'_{kl}=\mathbf{r}^{(k)'}\mathbf{r}^{(l)'}\,, \tag{17.13-8}$$

17.13-8-ba behelyettesítve a 17.13-1 formulát, a

$$G'_{kl}=\sum S_{kK}S_{lL}\mathbf{r}^{(K)}\mathbf{r}^{(L)} \tag{17.13-9}$$

eredményhez jutunk, tehát

$$\mathbf{G}'=\mathbf{SG\tilde{S}} \tag{17.13-10}$$

adja a metrikus tenzor transzformáltját.

Mivel

$$\mathbf{a}'=\mathbf{G}'\bar{\mathbf{a}}', \tag{17.13-11}$$

17.13-10 segítségével meghatározhatjuk az $\underline{\mathbf{a}}$ vektor kovariáns komponenseinek transzformációs szabályát. Írjuk be ugyanis 17.13-10-et és 17.13-6-ot 17.13-11-be, az

$$\mathbf{a}'=\mathbf{SG\tilde{S}\tilde{S}}^{-1}\bar{\mathbf{a}}=\mathbf{SG}\bar{\mathbf{a}}, \tag{17.13-12}$$

illetve

$$\mathbf{G}\bar{\mathbf{a}}=\mathbf{a}$$

felhasználásával az

$$\mathbf{a}'=\mathbf{Sa} \tag{17.13-13}$$

transzformációs szabályhoz jutunk.

Látható tehát, hogy a kovariáns komponensek az $\mathbf{S}$ mátrixszal transzformálhatók, tehát az alapvektorhoz hasonló módon változnak.

A vektorokra kapott eredmények tenzorokra is érvényesek. Mivel

$$\bar{\mathbf{a}}\mathbf{A}\bar{\mathbf{b}}=\alpha$$

skalár, és $\bar{\mathbf{a}}$, $\bar{\mathbf{b}}$ az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorok kontravariáns, $\mathbf{A}$ pedig az $\underline{\mathbf{A}}$ tiszta kovariáns reprezentációja, kell, hogy

$$\bar{\mathbf{a}}\mathbf{A}\bar{\mathbf{b}}=\bar{\mathbf{a}}'\mathbf{SA\tilde{S}}\bar{\mathbf{b}}' \tag{17.13-14}$$

teljesüljön; 17.13-14-ből leolvashatjuk, hogy az $\underline{\mathbf{A}}$ tenzor $\underline{\mathbf{A}}'$ transzformáltját az

$$\mathbf{A}'=\mathbf{SA\tilde{S}} \tag{17.13-15}$$

formulával határozhatjuk meg. A tenzorok kovariáns reprezentációi tehát az $\mathbf{S}$ mátrix segítségével transzformálhatók.

Az Einstein-féle jelölésmódot használva, a 17.13-15 szabály az

$$A'_{kl}=\sum S_{kK}S_{lL}A_{KL} \tag{17.13-16}$$

alakot ölti.

A skaláris szorzat

$$\underline{\mathbf{a}}\cdot\underline{\mathbf{A}}\cdot\underline{\mathbf{b}}=\alpha \tag{17.13-17}$$

alakjából az $\underline{\mathbf{A}}$ tenzor kontravariáns reprezentációjának transzformációs szabályát kapjuk. 17.13-13-at beírva 17.13-17-be, az

$$\mathbf{a}'\tilde{\mathbf{S}}^{-1}\cdot\mathbf{A}\cdot\mathbf{S}^{-1}\mathbf{b}=\alpha$$

összefüggéshez, s ennek alapján az

$$\bar{\mathbf{A}}'=\tilde{\mathbf{S}}^{-1}\bar{\mathbf{A}}\mathbf{S}^{-1} \qquad 17.13\text{-}18$$

transzformációs szabályhoz jutunk. 17.13-18 az Einstein-féle jelöléssel az

$$A^{kl'}=\sum S^{+}_{Kk}S^{+}_{Ll}A^{KL} \qquad 17.13\text{-}19$$

alakban írható fel.

Hasonló gondolatmenettel kapjuk az $\underline{\mathbf{A}}$ tenzor vegyes reprezentációinak

$$A^{k}_{\cdot l}=\sum S^{+}_{Kk}S_{lL}A^{K}_{\cdot L}\,, \qquad 17.13\text{-}20$$

és

$$A^{\cdot l}_{k}=\sum S_{kK}S^{+}_{Ll}A^{\cdot L}_{K} \qquad 17.13\text{-}21$$

transzformációs szabályait.

18. Több dimenziós tenzorok

Az eddigiekben egydimenziós (vektorok) és kétdimenziós tenzorokkal foglalkoztunk. A dimenziót úgy értelmezzük, hogy egy tenzor dimenziószámát a tenzort reprezentáló mátrix indexeinek száma adja meg. Így a skalárokat nulladimenziós tenzornak is tekinthetjük. A következőkben a kettőnél több indexes mátrixokkal reprezentálható tenzorokkal foglalkozunk.

Megjegyezzük még, hogy a tárgyalást továbbra is a háromdimenziós térre korlátozzuk, a tenzorokat reprezentáló mátrixok indexei tehát most is csak az 1, 2, 3 értékeket vehetik fel. Figyelmeztetünk még arra, hogy a fenti kifejezésmódot használva, a „tenzor dimenziószáma" és a tér „dimenziószáma" nem azonos jelentésűek, a két fogalom összetévesztése komoly hibákat okozhat.

18.1. A több dimenziós tenzor definíciója

A kétdimenziós tenzorokat homogén lineáris vektortranszformációként, tehát az eggyel alacsonyabb dimenziójú tenzorok segítségével értelmeztük.

Ugyanezzel a módszerrel értelmezhetjük a több dimenziós tenzorokat is.

Az $\overset{(n)}{\underline{\mathbf{A}}}$ n dimenziós tenzor olyan homogén lineáris operátor, amely tetszőleges $\underline{\mathbf{a}}$ vektort $n-1$ dimenziós tenzorrá képez le.

Tehát

$$\overset{(n)}{\underline{\mathbf{A}}}\underline{\mathbf{a}} = \overset{(n-1)}{\underline{\mathbf{B}}}. \qquad 18.1\text{-}1$$

Az operáció homogén lineáris tulajdonságából következik, hogy ha

$$\underline{\mathbf{a}} = \sum_k a^k \underline{\mathbf{f}}^{(k)},$$

akkor

$$\overset{(n)}{\underline{\mathbf{A}}}\underline{\mathbf{a}} = \sum_k a^k (\underline{\mathbf{A}}\underline{\mathbf{f}}^{(k)}). \qquad 18.1\text{-}2$$

Amennyiben az $\underline{\mathbf{f}}^{(k)}$ vektorok nem komplanárisak, akkor általában egy ferdeszögű koordináta-rendszert határoznak meg. 18.1-2-ből látható, hogy az

$$\overset{(n-1)}{\underline{\mathbf{A}}_k} = \overset{(n)}{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)} \quad (k=1, 2, 3) \qquad 18.1\text{-}3$$

$(n-1)$ dimenziós operátorok segítségével tetszőleges $\underline{\mathbf{a}}$ vektor leképzése meghatározható, tehát egy $\overset{(n)}{\underline{\mathbf{A}}}$ n dimenziós tenzor három $\overset{(n-1)}{\underline{\mathbf{A}}_k}$ $(n-1)$ dimenziós tenzorral egyértelműen jellemezhető.

Ha n helyett $n-1$-et írunk, akkor adódik, hogy egy $(n-1)$ dimenziós $\overset{(n-1)}{\underline{\mathbf{A}}_k}$ tenzor egyértelműen jellemezhető az

$$\overset{(n-2)}{\underline{\mathbf{A}}_{kl}} = \overset{(n-1)}{\underline{\mathbf{A}}_k}\underline{\mathbf{f}}^{(l)} = (\overset{(n)}{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)})\underline{\mathbf{f}}^{(l)} \quad (k, l=1, 2, 3) \qquad 18.1\text{-}4$$

$(n-2)$ dimenziós tenzorokkal.

A gondolatmenetet folytatva, az

$$A_{kl\dots n} = \overset{(n)}{\underline{\mathbf{A}}}\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}\dots\underline{\mathbf{f}}^{(n)} \quad (k, l, \dots, n=1, 2, 3) \qquad 18.1\text{-}4$$

skalár értékekhez juthatunk. 18.1-4-ben az egyszerűség kedvéért nem tettük ki a zárójeleket, természetes azonban, hogy most az asszociatív törvény nem érvényes, és a leírás sorrendje egyben a műveletek sorrendjét is jelenti.

A 18.1-4 által definiált 3^n darab skalárt egy n dimenziós mátrix elemeinek tekinthetjük, s ez a mátrix az $\underline{\mathbf{f}}^{(k)}$ vektorok által definiált koordináta-rendszerben egyértelműen meghatározza az $\overset{(n)}{\underline{\mathbf{A}}}$ operátort. Írhatjuk tehát, hogy

$$\mathcal{K}(\overset{(n)}{\underline{\mathbf{A}}}) = \overset{(n)}{\mathbf{A}} = A_{kl\ldots n} \quad (k, l, \ldots, n = 1, 2, 3). \qquad 18.1\text{-}5$$

A tenzorműveletek reprezentációkban történő kifejtése mátrixműveletekkel történik, a mátrixműveleteket pedig már tárgyaltuk.

18.2. Háromdimenziós tenzorok

Az általános definíció jobb megértése céljából foglalkozzunk kissé részletesebben a háromdimenziós tenzorokkal.

Legyen $\overset{(3)}{\underline{\mathbf{A}}}$ egy háromdimenziós tenzor. Határozzuk meg a tenzort reprezentáló mátrix elemeit az $\underline{\mathbf{f}}^{(k)}$ $(k=1, 2, 3)$ vektorok által definiált ferdeszögű koordináta-rendszerben.

Az általános definíciók alapján az $\overset{(3)}{\underline{\mathbf{A}}}$ tenzor a három

$$\overset{(2)}{\underline{\mathbf{A}}}_k = \overset{(3)}{\underline{\mathbf{A}}}\,\underline{\mathbf{f}}^{(k)} \quad (k=1, 2, 3) \qquad 18.2\text{-}1$$

kétdimenziós, ill. kilenc

$$\overset{(1)}{\underline{\mathbf{A}}}_{kl} = \overset{(3)}{\underline{\mathbf{A}}}\,\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)} \quad (k, l=1, 2, 3) \qquad 18.2\text{-}2$$

egydimenziós tenzorral, ill. a huszonhét

$$\overset{(0)}{\underline{\mathbf{A}}}_{klm} = \overset{(3)}{\underline{\mathbf{A}}}\,\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}\underline{\mathbf{f}}^{(m)} \quad (k, l, m=1, 2, 3) \qquad 18.2\text{-}3$$

skalárral jellemezhető. 18.2-3 adja $\overset{(3)}{\underline{\mathbf{A}}}$ $\mathcal{K}(\overset{(3)}{\underline{\mathbf{A}}}) = \overset{(3)}{\mathbf{A}}$ mátrixreprezentációját.

Az

$$A_{klm} = \overset{(3)}{\underline{\mathbf{A}}}\,\underline{\mathbf{f}}^{(k)}\underline{\mathbf{f}}^{(l)}\underline{\mathbf{f}}^{(m)} \quad (k, l, m=1, 2, 3) \qquad 18.2\text{-}4$$

előállítást a kétdimenziós eset általánosításaként az $\overset{(3)}{\underline{\mathbf{A}}}$ tenzor kovariáns előállításának nevezzük.

$\overset{(3)}{\underline{\mathbf{A}}}$ kontravariáns reprezentációja az

$$A^{klm} = \overset{(3)}{\underline{\mathbf{A}}}\,\bar{\underline{\mathbf{f}}}^{(k)}\bar{\underline{\mathbf{f}}}^{(l)}\bar{\underline{\mathbf{f}}}^{(m)} \quad (k, l, m=1, 2, 3) \qquad 18.2\text{-}5$$

formulával határozható meg. 18.2-5-ben az $\bar{\underline{\mathbf{f}}}^{(k)}$ $(k=1, 2, 3)$ vektorok az $\underline{\mathbf{f}}^{(l)}$ $(l=1, 2, 3)$ alapvektorok reciprok rendszerét jelentik.

A kétdimenziós eset analógiájára képezhetjük a különböző kevert reprezentációkat is. Tehát például

$$A^{k\,;\,m}_{\;\,l} = \overset{(3)}{\underline{\mathbf{A}}}\,\bar{\underline{\mathbf{f}}}^{(k)}\underline{\mathbf{f}}^{(l)}\bar{\underline{\mathbf{f}}}^{(m)}. \qquad 18.2\text{-}6$$

18.2.1. A háromdimenziós tenzorok transzformációja

Határozzuk meg az $\overset{(3)}{\underline{\mathbf{A}}}$ háromdimenziós tenzor

$$\mathcal{K}(\overset{(3)}{\underline{\mathbf{A}}}) = \overset{(3)}{\mathbf{A}} \quad \text{és} \quad \mathcal{K}'(\overset{(3)}{\underline{\mathbf{A}}}) = \overset{(3)\prime}{\mathbf{A}}$$

reprezentációi közötti összefüggést. Jelöljük a $\mathcal{K}$ és $\mathcal{K}'$ koordináta-rendszerek alapvektorait rendre $\underline{\mathbf{f}}^{(k)}$-val és $\underline{\mathbf{f}}^{(l)\prime}$-vel ($k, l = 1, 2, 3$), és legyen

$$\underline{\mathbf{f}}^{(k)\prime} = \sum_{l=1}^{3} S_{kl}\underline{\mathbf{f}}^{(l)}, \qquad 18.2\text{-}7$$

ill. az inverz transzformáció

$$\underline{\mathbf{f}}^{(k)} = \sum_{l} S^{+}_{kl}\underline{\mathbf{f}}^{(l)\prime}. \qquad 18.2\text{-}8$$

18.2-8-at behelyettesítve 18.2-4-be, és felhasználva $\overset{(3)}{\underline{\mathbf{A}}}$ homogén lineáris tulajdonságát, az

$$A_{klm} = \sum_{K,L,M=1}^{3} S^{+}_{kK}S^{+}_{lL}S^{+}_{mM}\overset{(3)}{\underline{\mathbf{A}}}\,\underline{\mathbf{f}}^{(K)\prime}\underline{\mathbf{f}}^{(L')}\underline{\mathbf{f}}^{(M)\prime} \qquad 18.2\text{-}9$$

formulát kapjuk. Mivel

$$\overset{(3)}{\underline{\mathbf{A}}}\,\underline{\mathbf{f}}^{(K)\prime}\underline{\mathbf{f}}^{(L)\prime}\underline{\mathbf{f}}^{(M)\prime} = A'_{KLM},$$

az

$$A_{klm} = \sum_{L,M} S^{+}_{kK}S^{+}_{lL}S^{+}_{mM}A_{KLM} \qquad 18.2\text{-}10$$

transzformációs szabályhoz jutunk. Az inverz transzformációt pedig az

$$A'_{klm} = \sum_{K,L,M} S_{kK}S_{lL}S_{mM}A_{KLM} \qquad 18.2\text{-}11$$

alakban írhatjuk fel. Az A'_{klm} tiszta kovariáns reprezentáció, tehát valóban ugyanazzal az **S** mátrixszal transzformálható, amely az alapvektorok transzformációját megszabja.

Határozzuk meg most az

$$A^{klm} = \overset{(3)}{\underline{\mathbf{A}}}\,\bar{\underline{\mathbf{f}}}^{(k)}\bar{\underline{\mathbf{f}}}^{(l)}\bar{\underline{\mathbf{f}}}^{(m)} \qquad 18.2\text{-}12$$

kontravariáns reprezentáció transzformációs szabályát. Most

$$A'^{klm} = \overset{(3)}{\underline{\mathbf{A}}}\,\bar{\underline{\mathbf{f}}}^{(k)'}\bar{\underline{\mathbf{f}}}^{(l)'}\bar{\underline{\mathbf{f}}}^{(m)'}, \qquad 18.2\text{-}13$$

így az $\bar{\underline{\mathbf{f}}}^{(k)'}$-t $(k=1, 2, 3)$ kell kifejeznünk az $\bar{\underline{\mathbf{f}}}^{(k)}$ vektorok segítségével. Az $\underline{\mathbf{f}}^{(k)}$ $(k=1, 2, 3)$ vektorok transzformációs szabályát 18.2-7-tel előírtuk. Mivel

$$\bar{\underline{\mathbf{f}}}^{(k)'} = \sum_{l} G^{kl'}\underline{\mathbf{f}}^{(l)'}, \qquad 18.2\text{-}14$$

és 17.13-19 szerint

$$G^{kl'} = \sum_{K,L} S^{+}_{Kk} S^{+}_{Ll} G^{KL}, \qquad 18.2\text{-}15$$

18.2-7 és 18.2-15 18.2-14-be való behelyettesítésével adódik, hogy

$$\bar{\underline{\mathbf{f}}}^{(k)'} = \sum_{K,L,l,j} S^{+}_{Kk} S^{+}_{Ll} S_{lj} G^{KL}\, \underline{\mathbf{f}}^{(j)}.$$

Ebből pedig, ha a jobb oldalon az l-re vonatkozó összegezést elvégezzük,

$$\bar{\underline{\mathbf{f}}}^{(k)'} = \sum_{K,L} S^{+}_{Kk} G^{KL} \underline{\mathbf{f}}^{(L)} = \sum_{K} S^{+}_{Kk}\, \bar{\underline{\mathbf{f}}}^{(K)} \qquad 18.2\text{-}16$$

adódik.

18.2-16-ot 18.2-13-ba beírva és 18.2-5-öt felhasználva, az

$$A'^{klm} = \sum_{K,L,M} S^{+}_{Kk} S^{+}_{Ll} S^{+}_{Mm} A^{KLM}$$

transzformációs szabályhoz jutunk.

Hasonló gondolatmenettel kaphatjuk az $\underline{\mathbf{A}}$ tenzor vegyes reprezentációinak transzformációs szabályait.

A háromdimenziós tenzorok esetén nyert összefüggések minden további nélkül általánosíthatók több dimenzióra, tehát pl. kontravariáns reprezentáció esetén

$$A'^{kl\ldots n} = \sum_{K,L,\ldots,N} S^{+}_{kK} S^{+}_{lL} \ldots S^{+}_{nN} A^{KL\ldots N}. \qquad 18.2\text{-}17$$

19. Különleges operátorok

Az operátorokat az eddigiekben általában fizikai jelentésükön keresztül és nem adott koordináta-rendszerben vett reprezentációik segítségével igyekeztünk bevezetni. Néhány operátor esetén azonban érdemes a reprezentációkból kiindulni. A következőkben először összefoglaljuk a zérus- és egységoperátor reprezentációinak sajátságait, majd egy speciális háromdimenziós operátorral az $\overset{(1)}{\underline{\mathbf{E}}}$ tenzorral foglalkozunk.

19.1. A zérus- és egységoperátor

A zérusoperátor tetszőleges reprezentációban csupa

$$0_{kl}=0$$

elemekből áll.

Az $\underline{\mathbf{E}}$ egységoperátor

$$E_{kl}=\delta_{kl} \tag{19.1-1}$$

elemei minden ortogonális reprezentációban azonosak.

Ferdeszögű reprezentációkban azonban — mint már láttuk —

$$\mathcal{K}(\underline{\mathbf{E}})=\mathbf{G} \tag{19.1-2}$$

a kovariáns, és $\mathbf{G}^{-1}$ a kontravariáns előállítás. A kevert reprezentáció azonban most is

$$G_k^l=G_l^k=\delta_{lk} \tag{19.1-3}$$

Az egységoperátor $\mathcal{K}$ és $\mathcal{K}'$ koordináta-rendszerben vett reprezentációja között a

$$\mathbf{G}'=\mathbf{S}\mathbf{G}\tilde{\mathbf{S}} \tag{19.1-4}$$

összefüggés teremt kapcsolatot ($\mathbf{S}$ a $\mathcal{K}\to\mathcal{K}'$ transzformációt leíró mátrix). A 19.1-4 összefüggés jobb és bal oldalának determinánsát képezve, a

$$\det\mathbf{G}'=(\det\mathbf{S})^2\det\mathbf{G} \tag{19.1-5}$$

egyenlőséget kapjuk. Bevezetve a szokásos $G=\det\mathbf{G}$ és $G'=\det\mathbf{G}'$ jelölést, 19.1-5-ből a transzformációs mátrix determinánsa és a metrikus tenzorok determinánsai között a

$$\det \mathbf{S} = \pm \sqrt{\frac{G'}{G}} \qquad 19.1\text{-}6$$

összefüggést kapjuk. 19.1-6-ban a négyzetgyököt pozitív előjellel kell venni, ha a $\mathcal{K}$ és $\mathcal{K}'$ koordináta-rendszerek sodrása azonos; a negatív előjel érvényes, ha a sodrás különböző.

19.2. Az $\overset{(3)}{\underline{\mathbf{E}}}$ operátor

A következőkben egy speciális, az egységoperátorhoz nagyon hasonló háromdimenziós operátorral foglalkozunk.

$\overset{(3)}{\underline{\mathbf{E}}}$ definíció szerint az a háromdimenziós tenzor, amelynek egy adott $\mathcal{K}$-beli kovariáns reprezentációja

$$\mathcal{K}(\overset{(3)}{\underline{\mathbf{E}}}) = \overset{(3)}{\mathrm{E}}_{klm} = \alpha\varepsilon_{klm}, \qquad 19.2\text{-}1$$

ahol α egy később meghatározandó állandó, ε_{klm} pedig a Levi—Civita-szimbólum.

Transzformációs formulák segítségével 19.2-1-ből $\overset{(3)}{\underline{\mathbf{E}}}$ reprezentációját tetszőleges $\mathcal{K}'$ reprezentációban meg tudjuk adni. A $\mathcal{K}'(\overset{(3)}{\underline{\mathbf{E}}})$ kovariáns reprezentációt 18.2-11 szerint az

$$\overset{(3)'}{E}_{klm} = \sum_{K,L,M} S_{kK} S_{lL} S_{mM} \overset{(3)}{E}_{KLM} \qquad 19.2\text{-}2$$

képlettel határozhatjuk meg. Beírva ide 19.2-1-et, és felhasználva a determináns definícióját, kapjuk, hogy

$$\overset{(3)'}{E}_{klm} = \alpha\varepsilon_{klm} (\det \mathbf{S}). \qquad 19.2\text{-}3$$

19.1-6 felhasználásával 19.2-3 az

$$\overset{(3)'}{E}_{klm} = \alpha \sqrt{\frac{G'}{G}}\, \varepsilon_{klm}$$

alakot ölti. Amennyiben $\alpha = \sqrt{G}$, akkor

$$\overset{(3)}{E}_{klm} = \sqrt{G}\varepsilon_{klm} \quad \text{és} \quad \overset{(3)'}{E}_{KLM} = \sqrt{G'}\varepsilon_{KLM}. \qquad 19.2\text{-}4$$

19.2-5-ből látjuk, hogy ha az $a=\sqrt{G}$ választással élünk, akkor a 19.2-1 által definiált tenzor kovariáns reprezentációja bármilyen vonatkoztatási rendszerben hasonló formában jelenik meg. Tehát az $\overset{(3)}{\underline{\mathbf{E}}}$ háromdimenziós tenzor kovariáns reprezentációja tetszőleges rendszerben

$$\mathcal{K}(\overset{(3)}{\underline{\mathbf{E}}})=\overset{(3)}{E}_{klm}=\sqrt{G}\varepsilon_{klm}. \qquad 19.2\text{-}5$$

A definíciókat felhasználva, $\overset{(3)}{\underline{\mathbf{E}}}$ kontravariáns reprezentációját az

$$\overset{(3)}{E}{}^{klm}=\sum_{KLM} G^{+}_{kK}G^{+}_{lL}G^{+}_{mM}\overset{(3)}{E}_{KLM}=\frac{\overset{(3)}{E}_{klm}}{G}$$

formulából kapjuk meg. Tehát

$$\overset{(3)}{E}{}^{klm}=\frac{\varepsilon_{klm}}{\sqrt{G}}.$$

Megfontolásunk tehát arra az eredményre vezetett, hogy ha 19.2-5 érvényes egy koordinátareprezentációban, akkor tetszőleges reprezentációban is fennáll.

A következőkben az $\overset{(3)}{\underline{\mathbf{E}}}$ tenzor által létrehozott vektor—tenzor és tenzor—vektor leképezések tulajdonságait vizsgáljuk.

Legyen $\underline{\mathbf{c}}$ egy vektor, és képezzük az

$$\overset{(3)}{\underline{\mathbf{E}}}\underline{\mathbf{c}}=\underline{\mathbf{T}} \qquad 19.2\text{-}6$$

transzformációt. Belátjuk, hogy $\underline{\mathbf{T}}$ antiszimmetrikus tenzor. A 19.2-6 összefüggést egy $\mathcal{K}$ reprezentációban a

$$\sqrt{G}\sum_{m}\varepsilon_{klm}c^{m}=T_{kl} \qquad 19.2\text{-}7$$

formula reprezentálja. Mivel $\varepsilon_{klm}=-\varepsilon_{lkm}$,

$$T_{kl}=-T_{lk}, \qquad 19.2\text{-}8$$

azaz $\underline{\mathbf{T}}$ valóban antiszimmetrikus. Írjuk fel 19.2-7 alapján a $\mathbf{T}$ mátrix elemeit részletesebben. A három független mátrixelem a következő:

$$\begin{aligned} T_{12}&=-T_{21}=\sqrt{G}c^{3},\\ T_{23}&=-T_{32}=\sqrt{G}c^{1}, \qquad 19.2\text{-}9\\ T_{31}&=-T_{13}=\sqrt{G}c^{2}. \end{aligned}$$

19.2-9 mutatja, hogy a $\underline{c}$ vektor kontravariáns komponenseit $\sqrt{G}$-vel szorozva, a $\underline{T}$ tenzor kovariáns komponenseit kapjuk.

Hasonlóan

$$T^{ki}=\frac{1}{\sqrt{G}}\sum_m \varepsilon_{klm}c_m, \tag{19.2-10}$$

tehát

$$\left.\begin{aligned} T^{12}&=\frac{1}{\sqrt{G}}\,c_3,\\ T^{23}&=\frac{1}{\sqrt{G}}\,c_2,\\ T^{31}&=\frac{1}{\sqrt{G}}\,c_1. \end{aligned}\right\} \tag{19.2-11}$$

Ortogonális reprezentáció esetén

$$T_{12}=c_3,\quad T_{23}=c_2,\quad T_{31}=c_1. \tag{19.2-12}$$

Mivel minden antiszimmetrikus tenzor felírható a 19.2-6 alakban, megállapítható, hogy az antiszimmetrikus operátorok reprezentációinak három független elemét egy vektor komponenseinek is tekinthetjük. A $\underline{c}$ vektort a $\underline{T}$ operátor *vektorinvariánsának* nevezzük.

Megfordított összefüggés is létezik!

Képezzük az

$$\left(\left(\overset{(3)}{\underline{E}}\,\underline{T}\right)\right)=\underline{c} \tag{19.2-13}$$

tenzor—vektor leképezést, ahol a kettős zárójel arra utal, hogy 19.2-13-at egy $\mathcal{K}$ reprezentációban felírva, két indexpárra kell összegezni. Tehát

$$c_k=\sum_{l,m}\overset{(3)}{E}_{klm}T^{ml}, \tag{19.2-14}$$

vagyis

$$c_k=\sqrt{G}\sum \varepsilon_{klm}T^{ml}. \tag{19.2-15}$$

19.2-15 komponensenként kiírva:

$$c_1=\sqrt{G}(T^{23}-T^{32}),$$

$$c_2=\sqrt{G}(T^{31}-T^{13}),$$

$$c_3=\sqrt{G}(T^{12}-T^{21}).$$

A 19.2-13 leképezést természetesen a $\underline{\mathbf{T}}$ tenzor kovariáns komponensei segítségével is felírhatjuk. Ekkor a

$$c^1=\frac{1}{\sqrt{G}}(T_{23}-T_{32}),$$

$$c^2=\frac{1}{\sqrt{G}}(T_{31}-T_{13}), \tag{19.2-16}$$

$$c^3=\frac{1}{\sqrt{G}}(T_{12}-T_{21})$$

összefüggéseket kapjuk.

Megállapítható, hogy ha $\underline{\mathbf{T}}$ antiszimmetrikus tenzor, akkor a

$$\underline{\mathbf{c}}=\left|\left(\overset{(3)}{\underline{\mathbf{E}}}\,\underline{\mathbf{T}}\right)\right|$$

vektorra alkalmazva az $\overset{(3)}{\underline{\mathbf{E}}}$ tenzort, az eredeti tenzor kétszeresét kapjuk, tehát

$$\left(\overset{(3)}{\underline{\mathbf{E}}}\,\underline{\mathbf{c}}\right)=2\underline{\mathbf{T}}, \tag{19.2-17}$$

azaz a $\underline{\mathbf{c}}$ vektor a $\underline{\mathbf{T}}$ tenzor vektorinvariánsának kétszerese.

19.3. Az $\overset{(3)}{\underline{\mathbf{E}}}$ tenzor és a vektoriális szorzat

Az $\overset{(3)}{\underline{\mathbf{E}}}$ tenzor segítségével a vektoriális szorzat mélyebb megértéséhez juthatunk.

A 19.2-13 leképezést alkalmazzuk a

$$\underline{\mathbf{T}}=\underline{\mathbf{a}}\circ\underline{\mathbf{b}} \tag{19.3-1}$$

tenzorra, ahol $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ tetszőleges vektorok. A 19.2-15 formula szerint a kapott eredmény azonos a

$$\sqrt{G}(\bar{\mathbf{a}}\times\bar{\mathbf{b}})=\mathbf{c} \tag{19.3-2}$$

vektoriális szorzattal. Szimmetrikusabb formát kapunk, ha bevezetjük az

$$[\underline{\mathbf{a}}\times\underline{\mathbf{b}}]=\underline{\mathbf{a}}\circ\underline{\mathbf{b}}-\underline{\mathbf{b}}\circ\underline{\mathbf{a}} \tag{19.3-3}$$

antiszimmetrikus tenzort, és összehasonlítjuk az

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=\underline{\mathbf{c}} \qquad 19.3\text{-}4$$

vektoriális szorzattal.

Megállapítható, hogy

$$\overset{(3)}{\underline{\mathbf{E}}}(\underline{\mathbf{a}}\times\underline{\mathbf{b}})=[\underline{\mathbf{a}}\times\underline{\mathbf{b}}] \qquad 19.3\text{-}5$$

és

$$\frac{1}{2}\left((\overset{(3)}{\underline{\mathbf{E}}}[\underline{\mathbf{a}}\times\underline{\mathbf{b}}])\right)=\underline{\mathbf{a}}\times\underline{\mathbf{b}}. \qquad 19.3\text{-}6$$

Az $\underline{\mathbf{a}}\times\underline{\mathbf{b}}$ vektor tehát a

$$\underline{\mathbf{T}}=\underline{\mathbf{a}}\circ\underline{\mathbf{b}}-\underline{\mathbf{b}}\circ\underline{\mathbf{a}} \qquad 19.3\text{-}7$$

tenzor vektorinvariánsa.

Összefoglalva a fentieket, azt mondhatjuk, hogy az $[\underline{\mathbf{a}}\times\underline{\mathbf{b}}]$ antiszimmetrikus tenzor a vektoriális szorzat tenzorreprezentációja, a szokásos $\underline{\mathbf{a}}\times\underline{\mathbf{b}}$ szorzat pedig a vektorreprezentáció. Az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektorokhoz a vektoriális szorzat egyenértékű módon rendel hozzá egy vektort és egy antiszimmetrikus tenzort. A két objektum matematikailag különbözik, azonban ugyanazon mennyiség kétféle leírásáról van szó.

Az irodalomban gyakran találkozunk az ún. „poláris" és „axiális" vektor megkülönböztető elnevezéssel. A vektoriális szorzatok eredményét poláris vektornak nevezik.

Véleményünk szerint az elnevezés fogalmi zavar terméke! A valóságban bármilyen mennyiséget, amelyet vektorral írhatunk le, éppen úgy antiszimmetrikus tenzorral is leírhatunk. Mint kimutattuk, két ekvivalens reprezentációról van szó. Az ún. „poláris", illetve „axiális" vektorok egyszerűen ugyanannak a mennyiségnek két reprezentációját jelentik. A két reprezentáció hasonló szerepet játszik a leírásban, mint az, hogy egy és ugyanazon vektormennyiséget kovariáns, illetve kontravariáns reprezentációban is megadhatunk.

A

$$\underline{\mathbf{c}}=\underline{\mathbf{a}}\times\underline{\mathbf{b}}$$

vektor általános tulajdonságait illetően nem különbözik az $\underline{\mathbf{a}}$ és $\underline{\mathbf{b}}$ vektoroktól.

A következőkben ismertetjük azt a gondolatmenetet, amelynek helytelen magyarázatából az említett félreértések származnak.

Képezzük le az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$, ... vektorokat egy ortogonális $\underline{\mathbf{S}}$ operátor segítségével az

$$\underline{\mathbf{a}}^*=\underline{\mathbf{S}}\underline{\mathbf{a}},\quad \underline{\mathbf{b}}^*=\underline{\mathbf{S}}\underline{\mathbf{b}},\quad \underline{\mathbf{c}}^*=\underline{\mathbf{S}}\underline{\mathbf{c}},\ldots \qquad 19.3\text{-}8$$

vektorokra.

Amennyiben

$$\det \mathbf{S}\neq 0,$$

akkor bizonyos összefüggések, amelyek az eredeti vektorokra érvényesek, változatlan formában fennállnak a leképezett vektorok között is. Így pl., ha

$$\alpha\underline{\mathbf{a}}+\beta\underline{\mathbf{b}}+\gamma\underline{\mathbf{c}}+\ldots=\underline{\mathbf{g}},$$

akkor 19.3-9

$$\alpha\underline{\mathbf{a}}^*+\beta\underline{\mathbf{b}}^*+\gamma\underline{\mathbf{c}}^*+\ldots=\underline{\mathbf{g}}^*,$$

valamint, ha

$$\underline{\mathbf{a}}\underline{\mathbf{b}}=\alpha,$$

akkor 19.3-10

$$\underline{\mathbf{a}}^*\underline{\mathbf{b}}^*=\alpha.$$

Az ortogonális transzformáció tehát vektorok lineáris kombinációját a transzformált vektorok lineáris kombinációjára képezi le, a skaláris szorzat értéke pedig nem változik a transzformáció során.

Legyen

$$\underline{\mathbf{a}}\times\underline{\mathbf{b}}=\underline{\mathbf{c}}.$$

Megállapítható, hogy

$$\underline{\mathbf{c}}^*=\det \mathbf{S}(\underline{\mathbf{a}}^*\times\underline{\mathbf{b}}^*). \tag{19.3-11}$$

Vektoriális szorzat transzformáltja tehát csak akkor egyezik meg a tényezők transzformáltjainak vektoriális szorzatával, ha

$$\det \mathbf{S}=+1.$$

Ha

$$\det \mathbf{S}=-1,$$

akkor

$$\underline{\mathbf{c}}^*=-(\underline{\mathbf{a}}^*\times\underline{\mathbf{b}}^*),$$

azaz

$$(\underline{\mathbf{a}}\times\underline{\mathbf{b}})^*=\underline{\mathbf{b}}^*\times\underline{\mathbf{a}}^*. \tag{19.3-12}$$

$\det \mathbf{S}=-1$ esetén az $\underline{\mathbf{S}}$ operátor által definiált leképezés *tükrözést* is tartalmaz. A 19.3-10, 19.3-11 és 19.3-12 összefüggések mutatják, hogy a transzformáció során egy pontrendszer egyes tulajdonságai megmaradnak, mások azonban megváltoznak.

Ez szemléletesen is könnyen érthetővé válik, ha pl. tükörből próbálunk újságot olvasni. A betűkből álló „pontrendszer” elolvasás szempontjából lényeges tulajdonságait a tükrözés erősen megváltoztatja.

Az a tény tehát, hogy a vektoriális szorzat nem invariáns a tükrözéssel szemben, nem azt jelenti, hogy a $\underline{\mathbf{c}}=\underline{\mathbf{a}}\times\underline{\mathbf{b}}$ „rendhagyó” vektor, hanem arra utal, hogy az $\underline{\mathbf{a}}\times\underline{\mathbf{b}}$-vel jellemzett fizikai összefüggés a vektorok tükrözése után megváltozik.

20. A négydimenziós tér

Az eddigiekben a háromdimenziós térben értelmezett tenzor- és vektormennyiségekkel foglalkoztunk. A tenzorok mátrixreprezentációjában ez azt jelentette, hogy a mátrixok egyes indexei három értéket vehetnek fel.

A következőkben, kifejezetten adott reprezentációból kiindulva, négy komponenssel jellemezhető mennyiségekkel foglalkozunk. Ezeket a mennyiségeket egyelőre formálisan négyesvektoroknak nevezzük, és a későbbiekben belátjuk, hogy a vektorfogalom olyan logikus általánosításáról van szó, amely fizikai összefüggések leírására alkalmas.

A következőkben a vektor- és tenzorfogalom alkalmazásaként a relativitáselméletben alapvető szerepet játszó Lorentz-transzformációval foglalkozunk.

20.1. A vonatkoztatási rendszer

20.1.1. A mozgás térbeli és időbeli jellemzése

Egy P pont mozgását az

$$\underline{\mathbf{r}} = \underline{\mathbf{r}}(t) \qquad 20.1\text{-}1$$

vektor-skalár függvénnyel írhatjuk le. $\underline{\mathbf{r}}(t)$ megadja a P pont helyvektorát az idő függvényében. 20.1-1 helyett paraméteres reprezentációt is használhatunk. A P pont helyzetének időbeli változását egy független p paraméter segítségével is megadhatjuk az

$$\underline{\mathbf{r}} = \underline{\mathbf{r}}(p) \quad \text{és} \quad t = t(p) \qquad 20.1\text{-}2$$

függvények segítségével. A 20.1-2 függvények a p paraméter adott értéke mellett meghatározzák az összetartozó idő- és helykoordinátákat.

$$\underline{\mathbf{r}}^{(k)} = \underline{\mathbf{r}}(p_k) \quad \text{és} \quad t^{(k)} = t(p_k) \qquad 20.1\text{-}3$$

tehát azt jelenti, hogy a mozgó pont a $t = t^{(k)}$ időpillanatban az $\underline{\mathbf{r}} = \underline{\mathbf{r}}^{(k)}$ helyen halad át.

Kényelmes, ha a $t(p)$ függvényt monoton növekedőnek választjuk, mert ekkor növekvő p értékeket választva, a P pont pályájának egymást követő pontjaihoz jutunk.

A pálya paraméteres leírását négyesvektorok segítségével is elvégezhetjük. Vezessük be az

$$\mathbf{x} = \mathbf{x}(p) \qquad 20.1\text{-}4$$

függvényt, ahol

és

$$\left.\begin{aligned} \mathbf{x} &= x_1, x_2, x_3, x_4, \\ \mathbf{r} &= x_1, x_2, x_3, \\ t &= x_4. \end{aligned}\right\} \qquad 20.1\text{-}5$$

Természetesen 20.1-5-ben az x_ν koordináták a p paraméter függvényei, tehát

$$x_\nu = x_\nu(p). \qquad 20.1\text{-}6$$

A következőkben a hármas- és négyesvektorok egyszerűbb megkülönböztetésére latin betűkkel jelöljük azokat az indexeket, amelyek az 1, 2, 3 értékeken futnak végig, görög betűkkel pedig azokat, amelyek az 1, 2, 3, 4 értékeket veszik fel.

20.1.2. Az idő mérése

A négyesvektorok bevezetésével a három helykoordináta mellé negyedikként az időt vezettük be egy mozgó pont pályájának leírására. A mozgás időbeli leírásához szükséges, hogy abban a tartományban, amelyben a pályát meghatározzuk, egyértelmű időmértékkel rendelkezzünk.

Az időmérés problémájának részletesebb vizsgálata azt mutatja, hogy egy egész tartományra nem tudunk egyértelmű időskálát definiálni. Látni fogjuk, hogy a problémát a kérdéses tartomány különböző pontjaiban elhelyezett órák megfelelő pontosságú összehangolása okozza. Tűrhető pontosságú időskála vezethető be azonban fénysugarak segítségével (a fény igen nagy terjedési sebessége miatt) a következő módon.

Helyezzünk el a kérdéses tartomány egy tetszőleges O_0 pontjában egy órát. Válasszuk ezt a pontot koordináta-rendszerünk origójának, tehát legyen $\mathbf{r}(O_0) = \underline{0}$. Válasszuk időmértékként az egész tartományban az O_0-ban elhelyezett (a továbbiakban egyszerűen O_0) óra által mutatott értéket. A mozgó pont $\mathbf{r} = \mathbf{r}(t)$ pályájának hely—idő függését ekkor úgy kaphatjuk, hogy a pálya egyes pontjaiból leolvassuk az O_0 óra által mutatott időt.

A fény véges terjedési sebessége miatt azonban az O_0 óra által mutatott időt a $P(r)$ pontból $\Delta t = \dfrac{r}{c}$ késéssel észleljük. Itt r a P pont origótól mért távolsága, c pedig a fénysebesség. A pálya pontosabb leírásához jutunk tehát, ha az

$$\underline{\mathbf{r}} = \underline{\mathbf{r}}\left(t^{(0)} + \frac{r}{c}\right) \qquad 20.1\text{-}7$$

függvényt határozzuk meg, ahol $t^{(0)}$ az O_0 óráról leolvasott érték, $\dfrac{r}{c}$ pedig a fény véges terjedési sebessége miatt alkalmazott korrekció. A korrekció 20.1-7-ben alkalmazott formája tartalmazza azt a feltevést, hogy a fény az adott tartományban izotrop módon terjed. Az alkalmazott leírás viszonylag lassan mozgó objektumok esetén nem okoz problémákat. Gyorsan mozgó testek esetén azonban — különösen akkor, ha a fény izotrop terjedésére vonatkozó feltevésünket is ellenőrizni kívánjuk — lényeges nehézségek merülnek fel.

A mozgás konzekvens leírásához juthatunk akkor is, ha a kiterjedt tartomány egyes pontjaiban — elég sűrűn — órákat helyezünk el. A mozgó pont pályájának időbeli változását ebben az esetben úgy állapítjuk meg, hogy a pálya egyes pontjaiban mindig az adott ponthoz közeli órán látható időt olvassuk le. Az

$$\underline{\mathbf{r}}^{(k)}=\underline{\mathbf{r}}(t^{(k)}) \qquad 20.1\text{-}$$

függvény tehát most azt jelenti, hogy amikor a mozgó test az $\underline{\mathbf{r}}^{(k)}$ helyvektorú pontban volt, akkor a pont közvetlen közelében fekvő óra a $t=t^{(k)}$ időt mutatta. (A $t^{(k)}$ érték esetleg több közeli óra által mutatott időmérték közötti interpoláció eredménye is lehet.)

Egy koordináta-rendszert, amelynek pontjaiban elég sűrűn órákat helyeztünk el, *vonatkoztatási rendszernek* nevezünk.[1] A vonatkoztatási rendszerekben az $\underline{\mathbf{r}}=\underline{\mathbf{r}}(t)$ összefüggés pontosan értelmezhető.

20.1.3. A vonatkoztatási rendszer meghatározása

Vonatkoztatási rendszerként használható tehát minden olyan koordináta-rendszer, melynek pontjaiban elég sűrűn órákat helyeztünk el. A pálya egyértelmű leírása egymástól teljesen függetlenül viselkedő órák esetén is lehetséges. Az összefüggések azonban csak olyan vonatkoztatási rendszerekben írhatók le, amelyekben az órák járása megfelelően összehangolt.

A rendszer különböző pontjaiban elhelyezett órákat tehát szinkronizálni kell. A szinkronizáció csak valamilyen fizikai folyamat segítségével végezhető el.

Az órák összehangolását célszerű fénysugarak segítségével elvégezni, és az eljárás lényegesen egyszerűsödik, ha feltételezzük, hogy a fény izotrop módon terjed. A szinkronizáció ennek alapján valóban lehetséges, a módszert most nem ismertetjük, részletes leírása megtalálható Jánossy L. „Relativitáselmélet a fizikai valóság alapján" c. könyvében.

Vizsgáljunk most egy fényjelet, amely az $\mathbf{r}^{(1)}$ pontból a $t^{(1)}$ időpillanatban indul és a $t^{(2)}$ időpillanatban érkezik az $\mathbf{r}^{(2)}$ pontba. Pontosabban az $\mathbf{r}^{(1)}$ pontból akkor indul egy fényjel, amikor a közvetlen közelében elhelyezett óra $t^{(1)}$-et mutat, s az $\mathbf{r}^{(2)}$ pontba akkor érkezik meg, amikor az $\mathbf{r}^{(2)}$ közelében elhelyezett óra $t^{(2)}$-t mutat. A fény izotrop terjedése folytán azt várjuk, hogy $t^{(2)}>t^{(1)}$ legyen, valamint teljesüljön az

[1] Vonatkoztatási rendszernek nevezzük a koordináta-rendszert akkor is, ha rendelkezünk olyan módszerrel, amelynek segítségével tetszőleges pontban egyértelmű időmértéket tudunk meghatározni.

$$(\mathbf{r}^{(2)}-\mathbf{r}^{(1)})^2-c^2(t^{(2)}-t^{(1)})^2=0 \qquad 20.1\text{-}9$$

egyenlet. A fenti összefüggést

$$\mathbf{x}^{(1)}=\mathbf{r}^{(1)},\ t^{(1)}, \qquad 20.1\text{-}10$$

$$\mathbf{x}^{(2)}=\mathbf{r}^{(2)},\ t^{(2)} \qquad 20.1\text{-}11$$

és

$$\mathbf{x}=\mathbf{x}^{(2)}-\mathbf{x}^{(1)}=\mathbf{r}^{(2)}-\mathbf{r}^{(1)},\quad t^{(2)}-t^{(1)} \qquad 20.1\text{-}12$$

négyesvektorok segítségével az

$$\mathbf{x}\Gamma\mathbf{x}=0 \qquad 20.1\text{-}13$$

alakban írhatjuk fel, ahol

$$\Gamma=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -c^2\end{pmatrix}. \qquad 20.1\text{-}14$$

Az $\mathbf{x}$ négyesvektort az $\mathbf{x}^{(2)}$ és $\mathbf{x}^{(1)}$ vektorokhoz tartozó pontok *négyes távolságának* nevezhetjük.

Négyes koordinátákban tehát a fény terjedését a 20.1-13 egyenlet írja le.

20.2. A Lorentz-transzformáció

Felvetődik a kérdés, hogy ha egy $\mathcal{K}$ vonatkoztatási rendszer mértékeiben érvényes a 20.1-13 egyenlet, tehát a fény izotrop módon terjed, akkor $\mathcal{K}$-t kitüntetett rendszernek kell-e tekintenünk, vagy léteznek más olyan koordinátarendszerek is, amelyekben a fény izotrop módon terjed? A kérdésfelvetést indokolják például a hang terjedésére vonatkozó tapasztalatok. A $\mathbf{v}$ sebességgel terjedő hanghullámok esetén azt tapasztaljuk, hogy az

$$(\mathbf{r}^{(2)}-\mathbf{r}^{(1)})^2-v^2(t^{(2)}-t^{(1)})^2=0$$

$$t_2>t_1 \qquad 20.2\text{-}1$$

egyenlet csak a levegőhöz képest rögzített vonatkoztatási rendszer mértékeiben írja le helyesen a hang terjedési törvényét.

Ennek alapján azt hihetnénk, hogy a fény terjedését leíró 20.1-9, ill. 20.1-13 törvény is csak a fény hordozójához képest — ezt a feltételezett közeget éternek szokás nevezni — nyugvó vonatkoztatási rendszerben érvényes. A következőkben belátjuk, hogy a fény a hanggal ellentétben nem tüntet ki egyetlen vonatkoztatási rendszert sem.

Legyen az négyesvektorok egy lineáris transzformációja

$$\mathbf{x}^{(k)}=\Lambda\mathbf{x}^{(k)}+\lambda \qquad 20.2\text{-}2$$

alakú, ahol $\mathbf{\Lambda}$ egy 4×4-es mátrix, λ pedig egy négykomponensű mennyiség. A felső index arra utal, hogy a vizsgált tartomány $P_1, \dots, P_n$ pontjait írjuk le a $\mathcal{K}$, ill. $\mathcal{K}'$ vonatkoztatási rendszerben.

Legyen

és

$$\left.\begin{aligned} \mathbf{x}^{(k)}-\mathbf{x}^{(l)}&=\mathbf{x}, \\ \mathbf{x}^{(k)'}-\mathbf{x}^{(l)'}&=\mathbf{x}', \end{aligned}\right\} \qquad 20.2\text{-}3$$

azaz $\mathbf{x}$ és $\mathbf{x}'$ a $\mathcal{K}$, ill. $\mathcal{K}'$ vonatkoztatási rendszerben reprezentálja a P_k-ból P_l-be mutató négyesvektort.

Ha $\underline{\mathbf{x}}$ egy fényjel pályájának két pontját összekötő négyesvektor, akkor

$$\mathbf{x}\mathbf{\Gamma}\mathbf{x}=0. \qquad 20.2\text{-}4$$

20.2-2. és 20.2-3. felhasználásával adódik, hogy

$$\mathbf{x}'=\mathbf{\Lambda}\mathbf{x}. \qquad 20.2\text{-}5$$

Amennyiben létezik a $\mathbf{\Lambda}^{-1}$ inverz mátrix, akkor

$$\mathbf{x}=\mathbf{\Lambda}^{-1}\mathbf{x}'=\mathbf{x}'\tilde{\mathbf{\Lambda}}^{-1}. \qquad 20.2\text{-}6$$

20.2-4-be 20.2-6-ot beírva, az

$$\mathbf{x}'(\tilde{\mathbf{\Lambda}}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\mathbf{x}'=0 \qquad 20.2\text{-}7$$

összefüggéshez jutunk.

Ha tehát

$$\tilde{\mathbf{\Lambda}}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1}=\mathbf{\Gamma}, \qquad 20.2\text{-}8$$

ill.

$$\tilde{\mathbf{\Lambda}}\mathbf{\Gamma}\mathbf{\Lambda}=\mathbf{\Gamma}, \qquad 20.2\text{-}9$$

akkor 20.2-7 az

$$\mathbf{x}'\mathbf{\Gamma}\mathbf{x}'=0 \qquad 20.2\text{-}10$$

alakot ölti, vagyis a $\mathcal{K}'$ vonatkoztatási rendszer mértékeiben is izotrop fényterjedési törvény érvényes.

A 20.2-9 összefüggést kielégítő mátrixokat *Lorentz-mátrixoknak*, a 20.2-2 által Lorentz-mátrixok segítségével definiált transzformációkat pedig *Lorentz-transzformációknak* nevezzük. Megmutatjuk, hogy számos 20.2-9-nek eleget tevő mátrix létezik.

Látjuk tehát, hogy amennyiben a fény egy $\mathcal{K}$ rendszer mértékeiben izotrop módon terjed, akkor bármely, a $\mathcal{K}$-ból Lorentz-transzformációval keletkező $\mathcal{K}'$ rendszerben is izotrop törvény áll fenn.

Amennyiben tehát létezik olyan $\mathcal{K}$ rendszer, amelyben a fény izotrop módon terjed, akkor több, $\mathcal{K}$-val egyenértékű vonatkoztatási rendszer is létezik.

A későbbiekben a 20.2-9 egyenlet megoldásait vizsgálva, meghatározzuk a Lorentz-mátrixok konkrét alakját.

20.2.1. Az „időtranszformáció" jelentése

A 20.2-5 formulával meghatározott Lorentz-transzformáció négyesvektorokra vonatkozik. A transzformáció fizikai jelentés szempontjából az eddigi transzformációktól (pl. forgatások) abban különbözik, hogy nemcsak a helykoordinátákat, hanem az időmértékeket is megváltoztatja.

Az „időtranszformáció" könnyebb megértése céljából írjuk fel a transzformációt az idő- és helykoordináták szétválasztásával.

Legyen tehát

$$\mathbf{x} = r_1, r_2, r_3, t \tag{20.2-11}$$

és

$$\mathbf{x}' = r'_1, r'_2, r'_3, t'. \tag{20.2-12}$$

A 20.2-5 transzformációt a 20.2-11 és 20.2-12 komponensek segítségével az

$$\left.\begin{aligned} r'_k &= \sum_{l=1}^{3} \Lambda_{kl} r_l + \Lambda_{k4} t \quad (k = 1, 2, 3), \\ t' &= \sum_{l=1}^{3} \Lambda_{4l} r_l + \Lambda_{44} t \end{aligned}\right\} \tag{20.2-13}$$

összefüggésekkel írhatjuk fel.

A 20.2-13 egyenletek azt jelentik, hogy a $\mathcal{K}$ rendszerben $\mathbf{r}$ = állandó helyvektorral rendelkező pont helyvektora a $\mathcal{K}'$ rendszerben az idő lineáris függvénye. A $P(\mathbf{r})$ pont tehát a $\mathcal{K}'$ rendszerhez képest egyenletes mozgást végez. Ez azt jelenti, hogy a $\mathcal{K}$ és $\mathcal{K}'$ rendszerek is egyenletes mozgást végeznek egymáshoz képest. Megjegyezzük, hogy $\mathcal{K}$ origójának, tehát az $\mathbf{r} = \mathbf{0}$ pontnak

$$r'_k = \Lambda_{k4} t, \tag{20.2-13a}$$

$$t' = \Lambda_{44} t \tag{20.2-13b}$$

a transzformáltja, tehát $v = r'/t'$ $\mathcal{K}$-nak a sebessége $\mathcal{K}'$ mértékeiben, és

$$v_k = \Lambda_{k4}/\Lambda_{44}. \tag{20.2-14a}$$

Hasonló módon, ha 20.2-13, 20.2-14 inverz transzformációjából indulunk ki, azt kapjuk, hogy $\mathcal{K}'$ origójának tehát az $\mathbf{r}' = \mathbf{0}$ pontnak a sebessége $\mathcal{K}$-ból nézve $\mathbf{w}$, ahol $\mathcal{K}$ mértékeiben

$$w_k = \Lambda^{+}_{k4}/\Lambda^{+}_{44}. \tag{20.2-14b}$$

A 20.2-13b formula tartalmát szokás azzal jellemezni, hogy ezzel az összefüggéssel a $\mathcal{K} \to \mathcal{K}'$ transzformáció során bekövetkező „időtranszformációt" írja le.

Az „időtranszformáció" kifejezés azonban nem egészen pontos, 20.2-13b valójában az *időmértékek* transzformációját adja meg. A $\mathcal{K}$ és $\mathcal{K}'$ rendszerben elhelyezett $O^{(k)}$ és $O^{(k)'}$ órák ugyanis különbözőképpen vannak szinkronizálva. Így ha az egymáshoz képest mozgó $O^{(k)}$ és $O^{(k)'}$ órák találkoznak, a találkozás pillanatában t és t' időmértékek általában különböznek. A 20.2-13b egyenlet éppen a t és t' mértékszámok közötti kapcsolatot határozza meg.

20.2.2. A Lorentz-csoport

Mielőtt a Lorentz-mátrixok meghatározására áttérnénk, bebizonyítjuk, hogy a Lorentz-mátrixok a mátrixszorzásra vonatkozóan csoportot alkotnak. A bizonyítást a csoportaxiómák teljesülésének bizonyításával végezzük el.

1. $\mathbf{\Lambda}=\mathbf{E}$ Lorentz-mátrix, hiszen

$$\tilde{\mathbf{E}}\mathbf{\Gamma}\mathbf{E}=\mathbf{\Gamma}$$

nyilvánvalóan teljesül.

2. Tegyük fel, hogy $\mathbf{\Lambda}^{(1)}$ és $\mathbf{\Lambda}^{(2)}$ két Lorentz-mátrix. Vizsgáljuk meg a

$$\tilde{\mathbf{\Lambda}}^{(3)}\mathbf{\Gamma}\mathbf{\Lambda}^{(3)} \qquad 20.2\text{-}15$$

szorzatot, ahol

$$\mathbf{\Lambda}^{(3)}=\mathbf{\Lambda}^{(1)}\mathbf{\Lambda}^{(2)}. \qquad 20.2\text{-}16$$

20.2-16-ot a szorzat transzformáltjára vonatkozó azonosság felhasználásával beírva 20.2-15-be, a

$$\tilde{\mathbf{\Lambda}}^{(2)}\tilde{\mathbf{\Lambda}}^{(1)}\mathbf{\Gamma}\mathbf{\Lambda}^{(1)}\mathbf{\Lambda}^{(2)}=\tilde{\mathbf{\Lambda}}^{(2)}(\tilde{\mathbf{\Lambda}}^{(1)}\mathbf{\Gamma}\mathbf{\Lambda}^{(1)})\mathbf{\Lambda}^{(2)}$$

formulához jutunk. Mivel $\mathbf{\Lambda}^{(1)}$ és $\mathbf{\Lambda}^{(2)}$ Lorentz-mátrix,

$$\tilde{\mathbf{\Lambda}}^{(1)}\mathbf{\Gamma}\mathbf{\Lambda}^{(1)}=\mathbf{\Gamma},$$

illetve

$$\tilde{\mathbf{\Lambda}}^{(2)}\mathbf{\Gamma}\mathbf{\Lambda}^{(2)}=\mathbf{\Gamma},$$

adódik, hogy

$$\tilde{\mathbf{\Lambda}}^{(3)}\mathbf{\Gamma}\mathbf{\Lambda}^{(3)}=\mathbf{\Gamma}.$$

Beláttuk tehát, hogy két Lorentz-mátrix szorzata is Lorentz-mátrix.

3. Képezzük a

$$\tilde{\mathbf{\Lambda}}\mathbf{\Gamma}\mathbf{\Lambda}=\mathbf{\Gamma} \qquad 20.2\text{-}17$$

egyenlőség mindkét oldalának determinánsát. A

$$(\det\mathbf{\Lambda})^2\det\mathbf{\Gamma}=\det\mathbf{\Gamma}$$

összefüggéshez jutunk, és minthogy

$$\det \boldsymbol{\Gamma} = -c^2 \neq 0,$$

és $\quad$ 20.2-18

$$\det \boldsymbol{\Lambda} = \pm 1$$

eredmény adódik. Létezik tehát a $\boldsymbol{\Lambda}$ és $\tilde{\boldsymbol{\Lambda}}$ mátrix inverze. Szorozzuk be a 20.2-17 mátrixegyenletet balról $\tilde{\boldsymbol{\Lambda}}^{-1}$-gyel, jobbról $\boldsymbol{\Lambda}^{-1}$-gyel. A

$$\boldsymbol{\Gamma} = \tilde{\boldsymbol{\Lambda}}^{-1}\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1} \qquad 20.2\text{-}19$$

összefüggéshez jutunk. Beláttuk tehát, hogy ha $\boldsymbol{\Lambda}$ Lorentz-mátrix, akkor $\boldsymbol{\Lambda}^{-1}$ létezik, és szintén Lorentz-mátrix.

4. Minthogy a Lorentz-mátrixok szorzása a szokásos mátrixszorzás segítségével történik, teljesül az asszociatív törvény. Tehát

$$(\boldsymbol{\Lambda}^{(1)}\boldsymbol{\Lambda}^{(2)})\boldsymbol{\Lambda}^{(3)} = \boldsymbol{\Lambda}^{(1)}(\boldsymbol{\Lambda}^{(2)}\boldsymbol{\Lambda}^{(3)}). \qquad 20.2\text{-}20$$

A Lorentz-mátrixok által alkotott csoportot *Lorentz-csoportnak* nevezzük.

20.2.3. A Lorentz-transzformációk explicit előállítása

20.2-5 és 20.2-8 szerint a Lorentz-transzformációk olyan

$$\mathbf{x}' = \boldsymbol{\Lambda}\mathbf{x}$$

homogén lineáris koordinátatranszformációk, amelyekre teljesül a

$$\tilde{\boldsymbol{\Lambda}}\boldsymbol{\Gamma}\boldsymbol{\Lambda} = \boldsymbol{\Gamma} \qquad 20.2\text{-}21$$

összefüggés.[1] A 20.2-21 formula 4×4-es mátrixok között létesít kapcsolatot, így 16 egyenletet jelent. Transzponálva az egyenlet mindkét oldalát, a formula önmagába megy át. Ez azt jelenti, hogy a 16 egyenlet között legfeljebb 10 független.

A következőkben megmutatjuk, hogy a 20.2-21-nek léteznek megoldásai, azután pedig független paraméterek segítségével elő is állítjuk ezeket a megoldásokat.

A megoldás meghatározása céljából hasznosabb 20.2-21-nek egy átalakított formájából kiindulni. Szorozzuk 20.2-21-et jobbról $\boldsymbol{\Lambda}^{-1}$-gyel, balról $\boldsymbol{\Gamma}^{-1}$-gyel; akkor a

$$\boldsymbol{\Lambda}^{-1} = \boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{\Lambda}}\boldsymbol{\Gamma} \qquad 20.2\text{-}22$$

formulához jutunk. 20.2-22 kényelmes módon adja meg $\boldsymbol{\Lambda}$ reciprokát $\tilde{\boldsymbol{\Lambda}}$ segítségével, és minthogy 20.2-22 egyenértékű 20.2-21-gyel, így $\boldsymbol{\Lambda}$ definíciójának is tekinthetnénk.

[1] Az inhomogén

$$\mathbf{x}' = \boldsymbol{\Lambda}\mathbf{x} + \boldsymbol{\lambda}$$

transzformációt Poincaré-transzformációnak nevezzük.

Írjuk ki a 20.2-22 összefüggést a

$$\Gamma_{\nu\mu} = \gamma_\nu \delta_{\nu\mu},$$

$$\gamma_1 = \gamma_2 = \gamma_3 = 1 \quad \text{és} \quad \gamma_4 = -c^2$$

formulák felhasználásával komponensekre bontva. A

$$\Lambda^+_{\nu\mu} = \gamma_\nu^{-1} \Lambda_{\mu\nu} \gamma_\mu \quad (\mu, \nu = 1, 2, 3, 4) \qquad 20.2\text{-}23$$

képlethez, ill. részletezve a

$$\Lambda^+_{kl} = \Lambda_{lk} \quad (k, l = 1, 2, 3), \qquad 20.2\text{-}24$$

$$\Lambda^+_{k4} = -c^2 \Lambda_{4k} \quad (k = 1, 2, 3), \qquad 20.2\text{-}25$$

$$\Lambda^+_{4l} = -\frac{1}{c^2} \Lambda_{l4} \quad (l = 1, 2, 3), \qquad 20.2\text{-}26$$

$$\Lambda^+_{44} = \Lambda_{44} \qquad 20.2\text{-}27$$

formulákhoz jutunk.

A kétféle típusú — a négyes indexet tartalmazó, ill. nem tartalmazó — elem könnyebb megkülönböztetésére a 20.2-14 formulák felhasználásával a

$$\left.\begin{aligned} &\Lambda_{kl} = G_{kl} \quad (k, l = 1, 2, 3), \\ &\Lambda_{l4} = -Bv_l \quad (l = 1, 2, 3), \\ &\Lambda_{4k} = -\frac{Bw_k}{c^2} \quad (k = 1, 2, 3), \\ &\Lambda_{44} = B \end{aligned}\right\} \qquad 20.2\text{-}28$$

jelöléseket vezetjük be. A 20.2-28 jelölés segítségével a 20.2-24, 20.2-25, 20.2-26, 20.2-27 egyenletek rendre a

$$\Lambda^+_{kl} = G_{lk}, \qquad 20.2\text{-}29$$

$$\Lambda^+_{k4} = Bw_k, \qquad 20.2\text{-}30$$

$$\Lambda^+_{4l} = \frac{Bv_l}{c^2}, \qquad 20.2\text{-}31$$

$$\Lambda^+_{44} = B \qquad 20.2\text{-}32$$

alakot öltik.

A 20.2-28 jelölésekkel a $\mathbf{\Lambda}$ Lorentz-mátrix és a $\mathbf{\Lambda}^{-1}$ reciprok mátrix a

$$\mathbf{\Lambda} = \begin{bmatrix} G_{11} & G_{12} & G_{13} & -Bv_1 \\ G_{21} & G_{22} & G_{23} & -Bv_2 \\ G_{31} & G_{32} & G_{33} & -Bv_3 \\ -\dfrac{Bw_1}{c^2} & -\dfrac{Bw_2}{c^2} & -\dfrac{Bw_3}{c^2} & B \end{bmatrix} \qquad 20.2\text{-}33$$

és

$$\mathbf{A}^{-1}=\begin{pmatrix} G_{11} & G_{21} & G_{31} & Bw_1 \\ G_{12} & G_{22} & G_{32} & Bw_2 \\ G_{13} & G_{23} & G_{33} & Bw_3 \\ \frac{Bv_1}{c^2} & \frac{Bv_2}{c^2} & \frac{Bv_3}{c^2} & B \end{pmatrix} \qquad 20.2\text{-}34$$

alakban írható fel. A 20.2-33 és 20.2-34 mátrixok felírása sokkal tömörebbé válik, ha bevezetjük a G_{kl} elemekkel definiált 3×3-as $\mathbf{G}$ mátrixot, valamint a v_k és w_l elemekkel definiált háromkomponensű $\underline{\mathbf{v}}$ és $\underline{\mathbf{w}}$ vektorokat.

$\mathbf{G}$, $\mathbf{v}$ és $\mathbf{w}$ felhasználásával a 20.2-33 és 20.2-34 mátrixok a

$$\mathbf{A}=\begin{pmatrix} \mathbf{G} & -B\mathbf{v} \\ -\frac{B\mathbf{w}}{c^2} & B \end{pmatrix} \qquad 20.2\text{-}35$$

és

$$\mathbf{A}^{-1}=\begin{pmatrix} \mathbf{G} & B\mathbf{w} \\ \frac{B\mathbf{v}}{c^2} & B \end{pmatrix} \qquad 20.2\text{-}36$$

alakot öltik. 20.2-35-ben értelemszerűen $\underline{\mathbf{w}}$ elemeit oszlopba, $\underline{\mathbf{v}}$ elemeit pedig sorba rendezve kell elképzelnünk.

A 20.2-35 és 20.2-36 formulákban szereplő mátrixokat *hipermátrixoknak* is szokás nevezni, mert a mátrix egyes elemei maguk is mátrixok. A hipermátrixok tulajdonságainak tárgyalásával most nem foglalkozunk, a fogalmat pusztán egyszerűsítő jelölésként használjuk. Annyit azonban megjegyzünk, hogy két mátrix szorzata nem változik, ha az eredeti mátrixokat hipermátrix formában szorozzuk össze.

Ennek megfelelően 20.2-35 és 20.2-36 segítségével az

$$\mathbf{E}=\mathbf{A}\,\mathbf{A}^{-1}$$

szorzat az

$$\mathbf{E}=\mathbf{A}\,\mathbf{A}^{-1}=\begin{pmatrix} \mathbf{G}\widetilde{\mathbf{G}}-B^2\frac{\mathbf{v}\circ\mathbf{v}}{c^2} & B\mathbf{G}\mathbf{w}-B^2\mathbf{v} \\ \frac{-B\mathbf{w}\widetilde{\mathbf{G}}+B^2\mathbf{v}}{c^2} & -B^2\frac{w^2}{c^2}+B^2 \end{pmatrix} \qquad 20.2\text{-}37$$

alakban írható fel.

Győződjünk meg példaként, hogy a mátrixok 20.2-33 és 20.2-34 eredeti alakjának összeszorzásával és 20.2-37 szerint $(\mathbf{A}\,\mathbf{A}^{-1})_{23}$-ra ugyanaz az eredmény adódik.

Az eredeti alakok felhasználásával

$$(\Lambda\Lambda^{-1})_{23}=G_{21}G_{31}+G_{22}G_{32}+G_{23}G_{33}-\frac{B^2 v_2 v_3}{c^2} \qquad 20.2\text{-}38$$

adódik. Ez az elem 20.2-37 bal felső sarokban elhelyezkedő blokkjában található, s könnyen meggyőződhetünk róla, hogy valóban megegyezik 20.2-38-cal.

Mivel 20.2-37-tel az egységmátrixot állítottuk elő, kell, hogy

$$\mathbf{G}\tilde{\mathbf{G}}-B^2\frac{\mathbf{v}\circ\mathbf{v}}{c^2}=\mathbf{E} \qquad 20.2\text{-}39$$

és

$$B^2\left(1-\frac{w^2}{c^2}\right)=1 \qquad 20.2\text{-}40$$

legyen,

ahol $\mathbf{E}$ a 3×3-as egységmátrix. Továbbá a nem diagonális elemek eltűnése miatt

$$B\mathbf{G}\mathbf{w}-B^2\mathbf{v}=\mathbf{0}. \qquad 20.2\text{-}41$$

Hasonló módon, ha a

$$\Lambda^{-1}\Lambda=\mathbf{E}$$

formulából indulunk ki, a

$$\tilde{\mathbf{G}}\mathbf{G}-B^2\frac{\mathbf{w}\circ\mathbf{w}}{c^2}=\mathbf{E}, \qquad 20.2\text{-}42$$

$$B\tilde{\mathbf{G}}\mathbf{v}-B^2\mathbf{w}=\mathbf{0}, \qquad 20.2\text{-}43$$

és

$$B^2\left(1-\frac{v^2}{c^2}\right)=1 \qquad 20.2\text{-}44$$

összefüggésekhez jutunk.

20.2-40 és 20.2-44 összehasonlítása mutatja, hogy amennyiben B valós, akkor

$$v=w<c, \qquad 20.2\text{-}45$$

továbbá

$$B=\frac{\varepsilon}{\sqrt[+]{1-\frac{w^2}{c^2}}} \quad (\varepsilon=\pm1). \qquad 20.2\text{-}46$$

Mivel $\mathbf{v}$ és $\mathbf{w}$ abszolút értéke megegyezik, bevezethetünk egy $\mathbf{O}$ ortogonális mátrixot úgy, hogy

$$\mathbf{v}=\varepsilon\mathbf{O}\mathbf{w} \quad (\varepsilon=\pm1). \qquad 20.2\text{-}47$$

20.2-47 biztosítja, hogy $v=w$ legyen, az ε szorzót kényelmi szempontok miatt vezettük be.

20.2-41-et és 20.2-43-at B-vel elosztva és átrendezve, adódik, hogy

$$\left.\begin{array}{l}\mathbf{G}\mathbf{w} = B\mathbf{v},\\ \widetilde{\mathbf{G}}\mathbf{v} = B\mathbf{w}.\end{array}\right\} \qquad 20.2\text{-}48$$

Célszerű a $\mathbf{G}$ mátrixot az $\mathbf{O}$ ortogonális mátrix segítségével a

$$\mathbf{G} = \mathbf{O} + \boldsymbol{\alpha} \qquad 20.2\text{-}49$$

alakban felírni.

20.2-49-et beírva 20.2-48-ba, és felhasználva 20.2-47-et, az

$$\boldsymbol{\alpha}\mathbf{w} = (B-\varepsilon)\mathbf{v}$$

és az

$$\widetilde{\boldsymbol{\alpha}}\mathbf{v} = (B-\varepsilon)\mathbf{w} \qquad 20.2\text{-}50$$

formulákhoz jutunk. A 20.2-50 egyenletek szerint az $\boldsymbol{\alpha}$ mátrix egy a $\mathbf{w}$ vektort $\mathbf{v}$ irányú vektorba, az $\widetilde{\boldsymbol{\alpha}}$ mátrix a $\mathbf{v}$ vektort $\mathbf{w}$ irányú vektorba viszi át. Ezt a feltételt teljesíti a $\mathbf{v}\circ\mathbf{w}$ mátrix is. Így $\boldsymbol{\alpha}$-t az

$$\boldsymbol{\alpha} = (B-\varepsilon)\frac{\mathbf{v}\circ\mathbf{w}}{w^2} \qquad 20.2\text{-}51$$

alakban vehetjük fel.

Mivel

$$B-\varepsilon = \varepsilon\left(\frac{1}{\sqrt{1-\frac{w^2}{c^2}}}-1\right), \qquad 20.2\text{-}52$$

látható, hogy $\boldsymbol{\alpha}\to\mathbf{0}$, ha $v = w\to 0$, tehát ez esetben $\mathbf{G}\to\mathbf{O}$. Ez az eredmény a 20.2-49 felbontás célszerűségét mutatja, hiszen kis sebességek esetén $\mathbf{G}$ majdnem ortogonális mátrix, s $\boldsymbol{\alpha}$ éppen az ortogonálistól vett eltérést adja.

20.2-51 felhasználásával 20.2-49 a

$$\mathbf{G} = \mathbf{O} + (B-\varepsilon)\frac{\mathbf{v}\circ\mathbf{w}}{w^2} \qquad 20.2\text{-}53$$

alakot ölti.

Bizonyítanunk kell még, hogy 20.2-53 kielégíti a még fel nem használt 20.2-39 és 20.2-42 egyenleteket is.

20.2-53-ból következik, hogy

$$\mathbf{G}\widetilde{\mathbf{G}} = \mathbf{O}\widetilde{\mathbf{O}} + (B-\varepsilon)\mathbf{O}\,\frac{\mathbf{w}\circ\mathbf{v}}{w^2} + (B-\varepsilon)\,\frac{\mathbf{v}\circ\mathbf{w}}{w^2}\,\widetilde{\mathbf{O}} + (B-\varepsilon)^2\,\frac{\mathbf{v}\circ\mathbf{w}}{w^2}\,\frac{\mathbf{w}\circ\mathbf{v}}{w^2}. \qquad 20.2\text{-}54$$

20.2-47 felhasználásával adódik, hogy

$$\mathbf{O}\frac{\mathbf{w}\circ\mathbf{v}}{w^2}=\varepsilon\frac{\mathbf{v}\circ\mathbf{v}}{w^2}$$

és

$$\frac{\mathbf{v}\circ\mathbf{w}}{w^2}\tilde{\mathbf{O}}=\varepsilon\frac{\mathbf{v}\circ\mathbf{v}}{w^2},$$

valamint tudjuk, hogy

$$\frac{\mathbf{v}\circ\mathbf{w}}{w^2}\,\frac{\mathbf{w}\circ\mathbf{v}}{w^2}=\frac{\mathbf{v}\circ\mathbf{v}}{w^2}.$$

Ennek megfelelően 20.2-54 a

$$\mathbf{G}\tilde{\mathbf{G}}=\mathbf{E}+2\varepsilon(B-\varepsilon)\frac{\mathbf{v}\circ\mathbf{v}}{w^2}+(B-\varepsilon)^2\frac{\mathbf{v}\circ\mathbf{v}}{w^2},$$

ill.

$$\mathbf{G}\tilde{\mathbf{G}}=E+(B^2-1)\frac{\mathbf{v}\circ\mathbf{v}}{w^2} \qquad 20.2\text{-}55$$

alakra egyszerűsödik.

Minthogy

$$\frac{B^2-1}{w^2}=\frac{B^2}{c^2},$$

20.2-55-ből következik, hogy

$$\mathbf{G}\tilde{\mathbf{G}}-\frac{B^2}{c^2}(\mathbf{v}\circ\mathbf{v})=\mathbf{E}. \qquad 20.2\text{-}56$$

20.2-56 megegyezik 20.2-39-cel, tehát állításunkat bebizonyítottuk.

Hasonló módon látható be, hogy **G** 20.2-53 alakja kielégíti 20.2-42-t is.

Összefoglalva tehát a

$$\mathbf{\Lambda}=\begin{pmatrix}\mathbf{G} & -B\mathbf{v}\\ -\dfrac{B\mathbf{w}}{c^2} & B\end{pmatrix} \qquad 20.2\text{-}57$$

alakú mátrix, amelyben

$$\mathbf{G}=\mathbf{O}+(B-\varepsilon)\frac{\mathbf{v}\circ\mathbf{w}}{w^2},$$

$$\mathbf{v}=\varepsilon\mathbf{O}\mathbf{w},$$

és

$$B=\frac{\varepsilon}{\sqrt{1-\dfrac{w^2}{c^2}}}\qquad(\varepsilon=\pm1)$$

valóban Lorentz-mátrix.

Kimutatható, de ezzel most nem foglalkozunk, hogy minden Lorentz-mátrix előállítható a 20.2-57 alakban, tehát 20.2-57 a 20.2-21 definíciós egyenlet általános megoldása.

A 20.2-57 mátrixban $\mathbf{O}$ 3×3-as ortogonális mátrix, tehát három független paraméterrel jellemezhető. Ugyancsak három független paramétert adnak $\mathbf{w}$ komponensei is, hiszen a $w<c$ feltétel nem jelent határozott összefüggést $\mathbf{w}$ komponensei között. A $\mathbf{\Lambda}$ Lorentz-mátrixot hat független paraméter segítségével állítottuk elő.

Az $\mathbf{O}$ forgatást pl. három szög segítségével megadva, valamint a $\mathbf{w}$ komponenseit előírva, $\mathbf{\Lambda}$-t még nem határoztuk meg egészen egyértelműen, hiszen az $s=\pm1$ lehetőség miatt B pozitív és negatív értéket vehet fel. Ugyancsak pozitív és negatív lehet det $\mathbf{O}$ is, attól függően, hogy tiszta elforgatásról vagy forgatás és tükrözés együtteséről van szó.

Így B és det $\mathbf{O}$ különböző előjelű értékei még négyféle lehetőséget tesznek lehetővé $\mathbf{\Lambda}$ előállításában.

Ha mind B, mind det $\mathbf{O}$ pozitív, akkor *valódi Lorentz-transzformációról* beszélünk. Belátható — de ezzel most nem foglalkozunk —, hogy a valódi Lorentz-transzformációk is csoportot alkotnak.

20.2.4. A Lorentz-mátrix komponenseinek fizikai jelentése

Az előzőek a Lorentz-mátrix komponenseinek fizikai jelentését már tartalmazzák, célszerűnek tartjuk azonban az ezzel kapcsolatos eredmények összefoglalását, a valódi Lorentz-transzformációkra vonatkozóan. Legyen

$$\mathbf{\Lambda}=\begin{pmatrix} \mathbf{O}+(B-1)\dfrac{\mathbf{v}\circ\mathbf{w}}{w^2} & -B\mathbf{v} \\ -\dfrac{B}{c^2}\mathbf{w} & B \end{pmatrix}$$

valódi Lorentz-mátrix.

A transzformáció linearitásából következik — amint ezt a 20.2.1. fejezetben már megállapítottuk —, hogy $-\mathbf{v}$ az eredeti $\mathcal{K}$ koordináta-rendszer origójának sebességét jelenti a $\mathcal{K}'$ rendszer mértékeiben kifejezve. Hasonló megfontolásokból adódott, hogy $-\mathbf{w}$ a $\mathcal{K}'$ koordináta-rendszer origójának sebessége a $\mathcal{K}$ rendszer mértékeiben kifejezve.

Megállapítottuk azt is, hogy a

$$\mathbf{G}=\mathbf{O}+(B-1)\frac{\mathbf{v}\circ\mathbf{w}}{w^2}$$

mátrix a $\mathcal{K}'$ koordináta-rendszer tengelyeinek orientációját jellemzi $\mathcal{K}$ tengelyeihez képest. Mivel $\mathbf{G}$ elemei kicsiny v értékek esetén alig térnek el $\left(\frac{v^2}{c^2}\right.$ nagyságrendben$\left.\right)$ egy ortogonális mátrix elemeitől, megállapítható, hogy a $\mathcal{K}'$ rendszer tengelyei $\mathcal{K}$ mértékeiben kifejezve majdnem — de nem pontosan — ortogonálisak.

20.2.5. A Lorentz-transzformáció néhány speciális esete

a) Legyen $\mathbf{v}=\mathbf{w}=\mathbf{0}$. Ekkor a $\mathcal{K}$ és $\mathcal{K}'$ rendszer nem mozog egymáshoz képest. A Lorentz-transzformáció ekkor egyszerű forgatásba megy át. A Lorentz-mátrixok ebben az esetben a

$$\mathbf{\Lambda}^{(0)}=\begin{pmatrix}\mathbf{O} & \mathbf{0}\\ \mathbf{0} & 1\end{pmatrix} \qquad 20.2\text{-}63$$

háromparaméteres sokaság elemei. Tehát

$$\Lambda_{kl}^{(0)}=O_{kl} \quad (k, l=1, 2, 3),$$

$$\Lambda_{4k}=\Lambda_{l4}=0, \quad \Lambda_{44}=1.$$

b) Egyszerű transzformációhoz jutunk az $\mathbf{O}=\mathbf{E}$ esetben. Ekkor $\mathcal{K}$ és $\mathcal{K}'$ megfelelő tengelyei párhuzamosak egymással. A Lorentz-mátrixok ismét háromparaméteres sokaságot alkotnak. A három paramétert $\mathbf{v}$ komponensei jelentik. Ekkor

$$\mathbf{v}=\varepsilon\mathbf{w},$$

$$\mathbf{\Lambda}^{(v)}=\begin{pmatrix}\mathbf{E}+(B-1)\dfrac{\mathbf{v}\circ\mathbf{v}}{v^2} & -B\mathbf{v}\\ -\dfrac{B}{c^2}\mathbf{v} & B\end{pmatrix}, \qquad 20.2\text{-}64$$

ahol

$$B=\frac{\varepsilon}{\sqrt[+]{1-\dfrac{v^2}{c^2}}}. \qquad 20.2\text{-}65$$

Megjegyezzük, hogy a 20.2-63 és 20.2-64 típusú mátrixok szorzata hatparaméteres sokaságot alkot, s meg lehet mutatni, hogy minden Lorentz-mátrix előállítható ilyen típusú mátrixok szorzataként.

c) Gyakran használjuk azt az egyparaméteres mátrixsokaságot, amelyben $\mathbf{O}=\mathbf{E}$ és $\mathbf{v}$ párhuzamos az egyik koordinátatengellyel, így pl. $\mathbf{v}=v, 0, 0$.

Ebben az esetben

$$\Lambda^{(v)}=\begin{pmatrix} B & 0 & 0 & -Bv \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{Bv}{c^2} & 0 & 0 & B \end{pmatrix}. \qquad 20.2\text{-}66$$

A transzformáció explicit alakját pedig az $\varepsilon=+1$ esetben az

$$x_1'=\frac{x_1-vt}{\sqrt{1-\frac{v^2}{c^2}}}, \quad x_2'=x_2, \quad x_3'=x_3, \qquad 20.2\text{-}67$$

$$t'=\frac{t-\frac{x_1 v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

egyenletek adják meg.

A 20.2-67 transzformációt igen gyakran „a Lorentz-transzformációnak" nevezik, valójában azonban ez az egyparaméteres transzformáció a hat paraméterrel jellemezhető általános Lorentz-transzformációknak csak egy speciális esete.

20.3. A Lorentz-deformáció

Az

$$\mathbf{x}'=\Lambda\mathbf{x} \qquad 20.3\text{-}1$$

Lorentz-transzformáció a $\mathcal{K}$ és $\mathcal{K}'$ vonatkoztatási rendszerben vett méretek között állapít meg összefüggést. Megállapítottuk, hogy a

$$\tilde{\Lambda}\Gamma\Lambda=\Gamma \qquad 20.3\text{-}2$$

egyenletnek eleget tevő mátrixok olyan koordináta-rendszerekben vett mértékek között határoznak meg összefüggést, amelyekben a fényterjedés törvénye

$$\mathbf{x}\Gamma\mathbf{x}=0 \qquad 20.3\text{-}3$$

alakú. A transzformációk által összekapcsolt rendszerekben tehát izotrop fényterjedési törvény áll fenn.

A Lorentz-transzformáció azonban a koordináta-transzformációval kapcsolatos jelentésen túl mélyebb tartalommal is rendelkezik:

Legyen

$$\mathbf{x} = \mathbf{x}(p) \qquad 20.3\text{-}4$$

egy P pont pályája paraméteres reprezentációban. 20.3-4, mint már említettük, azt jelenti, hogy a $t = t(p)$ időpillanatban a mozgó pont az $\mathbf{r} = \mathbf{r}(p)$ helyvektorú ponton halad át. 20.3-4-ben az egyes komponensek jelentése a következő:

$$\mathbf{r}(p) = x_1(p),\ x_2(p),\ x_3(p),$$

és

$$t(p) = x_4(p). \qquad 20.3\text{-}5$$

A 3.4 és 3.5 függvények alkalmasak a pálya leírására, ha $x_4(p)$ a p paraméter monoton növekedő függvénye.

Egy $\mathbf{\Lambda}$ Lorentz-mátrix segítségével a 20.3-4 pályát egy másik pályára képezhetjük le. Legyen ugyanis

$$\mathbf{x}^*(p) = \mathbf{\Lambda x}(p). \qquad 20.3\text{-}6$$

Az $\mathbf{x}^*(p)$ négyesvektorokat a $\mathcal{K}$ koordináta-rendszerben egy P^* pont pályájának tekinthetjük. Amennyiben $\mathbf{\Lambda} \neq \mathbf{E}$, akkor $\mathbf{x}^*(p)$ nem azonos $\mathbf{x}(p)$-vel. Mondhatjuk tehát, hogy a 20.3-6 transzformáció a P pont pályáját egy másik, P^* pont pályájára képezte le. Ebben az értelmezésben ne felejtsük el, hogy $\mathbf{x}(p)$ és $\mathbf{x}^*(p)$ ugyanazon vonatkoztatási rendszerben jelenti két különböző, P, illetve P^* pont négyes koordinátáit.

A fenti transzformációt pontrendszerekre is alkalmazhatjuk.

Jelentsék az

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k)}(p) \quad (k = 1, 2, \ldots)$$

függvények egy $\mathbf{P}$ pontrendszer különböző pontjainak pályáit. Az

$$\mathbf{x}^{(k)*}(p) = \mathbf{\Lambda x}^{(k)}(p) \qquad 20.3\text{-}7$$

négyesvektorok egy $\mathbf{P}^*$ pontrendszer pontjait jellemzik. A $\mathbf{P}^*$ pontrendszert a $\mathbf{P}$ pontrendszer *Lorentz-deformáltjának* nevezzük.

A 20.3-6, ill. 20.3-7 által definiált Lorentz-deformáció a pontrendszer méretváltozását adott $\mathcal{K}$ rendszerben fejezi ki. Megmutatjuk azonban, hogy a Lorentz-deformáció koordináta-rendszertől függetlennek tekinthető.

Legyen ugyanis $\mathcal{K}'$ egy vonatkoztatási rendszer, amely $\mathcal{K}$-ból a $\bar{\mathbf{\Lambda}}$ transzformációval keletkezik. Az $\mathbf{x}$ négyesvektorral jellemzett pont koordinátáit a $\mathcal{K}'$ rendszerben az

$$\mathbf{x}' = \bar{\mathbf{\Lambda}}\mathbf{x} \qquad 20.3\text{-}8$$

egyenlet adja meg. Adjuk meg a 20.3-7 deformáció leírását a $\mathcal{K}'$ rendszerben:

$$\mathbf{x}^{*\prime}(p) = \bar{\Lambda}\mathbf{x}^{*}(p) = \bar{\Lambda}\Lambda\mathbf{x}(p). \qquad 20.3\text{-}9$$

20.3-9-be beírva az

$$\mathbf{x} = \bar{\Lambda}^{-1}\mathbf{x}'$$

inverz transzformációt, 20.3-9 az

$$\mathbf{x}^{*\prime}(p) = \bar{\Lambda}\Lambda\bar{\Lambda}^{-1}\mathbf{x}'(p) \qquad 20.3\text{-}10$$

alakot ölti. 20.3-10 helyett az

$$\mathbf{x}^{*\prime}(p) = \Lambda'\mathbf{x}'(p) \qquad 20.3\text{-}11$$

összefüggést is felírhatjuk, ahol

$$\Lambda' = \bar{\Lambda}\Lambda\bar{\Lambda}^{-1}. \qquad 20.3\text{-}12$$

A Lorentz-mátrixok csoporttulajdonságából következik, hogy Λ' is Lorentz-mátrix. A $\mathcal{K}$ koordináta-rendszerben felírt 20.3-7 transzformáció a $\mathcal{K}'$ koordináta-rendszerben tehát a 20.3-11 alakban jelenik meg.

Megállapíthatjuk tehát, hogy a $\mathcal{K}$-ban definiált Lorentz-deformáció $\mathcal{K}'$-ben is Lorentz-deformációként fogható fel, csak a $\mathcal{K}$-ban Λ által képviselt deformáció $\mathcal{K}'$-ben Λ' segítségével állítható elő.

Koordináta-rendszertől függetlenül írhatjuk tehát, hogy

$$\underline{\mathbf{x}}^{*} = \underline{\Lambda}\underline{\mathbf{x}} \qquad 20.3\text{-}13$$

és

$$\mathbf{x} = \mathcal{K}(\underline{\mathbf{x}}), \quad \mathbf{x}^{*} = \mathcal{K}(\underline{\mathbf{x}}^{*}), \quad \Lambda = \mathcal{K}(\underline{\Lambda}), \qquad 20.3\text{-}14$$

illetve

$$\mathbf{x}' = \mathcal{K}'(\underline{\mathbf{x}}), \quad \mathbf{x}^{*\prime} = \mathcal{K}'(\underline{\mathbf{x}})^{*}, \quad \Lambda' = \mathcal{K}'(\underline{\Lambda}) \qquad 20.3\text{-}15$$

mátrixok a 20.3-13-ban szereplő valódi mennyiségek $\mathcal{K}$, ill. $\mathcal{K}'$ rendszerben vett reprezentációi.

20.3.1. A Lorentz-deformáció explicit formája

A $\underline{\mathbf{P}}$ pontrendszer $\underline{\Lambda}$ Lorentz-deformációját formálisan a

$$\underline{\mathbf{P}}^{*} = \underline{\Lambda}\underline{\mathbf{P}} \qquad 20.3\text{-}16$$

képlettel jelölhetjük, ahol

$$\underline{\mathbf{P}} = \underline{\mathbf{x}}^{(1)}(p), \underline{\mathbf{x}}^{(2)}(p), \dots, \underline{\mathbf{x}}^{(n)}(p),$$

$$\underline{\mathbf{P}}^{*} = \underline{\mathbf{x}}^{(1)*}(p), \underline{\mathbf{x}}^{(2)*}(p), \dots, \underline{\mathbf{x}}^{(n)*}(p) \qquad 20.3\text{-}17$$

és

$$\underline{\mathbf{x}}^{(k)*} = \underline{\Lambda}\underline{\mathbf{x}}^{(k)}(p) \quad (k = 1, \dots, n). \qquad 20.3\text{-}18$$

20.3-18 egy $\mathcal{K}$ rendszerbeli reprezentációját jelenti az

$$\mathbf{x}^{(k)*}(p) = \Lambda \mathbf{x}^{(k)}(p) \qquad 20.3\text{-}19$$

összefüggés.

20.3-19-be különböző $\mathbf{x}^{(k)}(p)$ értékeket helyettesítve, meghatározhatjuk a deformált rendszer alakját és mozgását.

A Λ Lorentz-mátrixot írjuk fel most a

$$\Lambda = \begin{pmatrix} \mathbf{O} + (B-1)\dfrac{\mathbf{v} \circ \mathbf{w}}{v^2} & B\mathbf{v} \\ \dfrac{B\mathbf{w}}{c^2} & B \end{pmatrix} \qquad 20.3\text{-}20$$

alakban. A deformációk esetén azért célszerűbb 20.3-20-at használni, mert ez a transzformáció ekkor egy nyugvó pályát egy $\mathcal{K}$-hoz képest $\mathbf{v}$ és nem $-\mathbf{v}$ sebességgel mozgó pályára képez le. 20.3-20-ban továbbá $B>0$-t választottunk, tehát a továbbiakban az $\varepsilon = +1$ esetre szorítkozunk.

Válasszuk szét 20.3-19-ben a hely- és időkoordinátákat. A szokásos jelölésekkel az

$$\mathbf{r}^{(k)*}(p) = \mathbf{G}\mathbf{r}^{(k)}(p) + B\mathbf{v}t^{(k)}(p), \qquad 20.3\text{-}21$$

$$t^{(k)*}(\mathrm{p}) = \mathrm{B}\left[\frac{\mathbf{w}\mathbf{r}^{(k)}(p)}{c^2} + t^{(k)}(p)\right] \qquad 20.3\text{-}22$$

egyenletekhez jutunk.

20.3-20 és 20.3-21 segítségével a deformált rendszer tetszőleges P_k^* pontjának helyzetét tetszőleges τ időpontban meghatározhatjuk. Legyen például

$$t^{(k)*}(p) = \tau. \qquad 20.3\text{-}23$$

Beírva ezt 20.3-22-be, a

$$B[\mathbf{w}\mathbf{r}^{(k)}(p) + t^{(k)}(p)] = \tau \qquad 20.3\text{-}24$$

egyenletből megállapíthatjuk a megfelelő $p(\tau)$ paraméterértéket, hiszen $\mathbf{r}^{(k)}(p)$ és $t^{(k)}(p)$ adott függvények. Így 20.3-24-ből a

$$p = p(\tau) \qquad 20.3\text{-}25$$

függvény adódik. 20.3-25-öt 20.3-21-be behelyettesítve, az

$$\mathbf{r}^{(k)*}(p_\tau) = \mathbf{G}\mathbf{r}^{(k)}(p_\tau) + B\mathbf{v}t^k(p_\tau) \qquad 20.3\text{-}26$$

formulát kapjuk. 20.3-25 és 20.3-26 $\mathbf{r}^{(k)*}$-ot a τ idő függvényeként adja meg. Ez az összefüggés általában bonyolult, mert 20.3-25 többnyire nem lineáris egyenlet.

Látható azonban, hogy a $\underline{\mathbf{P}}^*$ rendszer pontjait a $\underline{\mathbf{P}}$ rendszer pontjai meghatározzák. Megállapítható az is, hogy a P_k^* pontok helyzetét a τ időpillanatban általában P_k pontok más időpillanatban vett helyzetei határozzák meg.

20.3.2 A sebesség-összeadási törvény

A Lorentz-deformáció az eredeti rendszer igen bonyolult leképezésének eredménye. Speciális esetekben azonban viszonylag egyszerű összefüggésekhez juthatunk. Mozogjon az eredeti rendszer egyenletes $\mathbf{V}$ sebességgel, és paraméterként válasszuk az időt. Ekkor

$$\left.\begin{aligned} \mathbf{r}(t) &= \mathbf{V}t, \\ t &= p. \end{aligned}\right\} \qquad 20.3\text{-}27$$

20.3-27-et beírva 20.3-21-be és 20.3-22-be, a rendszer Lorentz-deformáltját a

$$\mathbf{r}^*(t) = (\mathbf{G}\mathbf{V} + \mathbf{v}B)t,$$

$$t^*(t) = B\left(\frac{\mathbf{w}\mathbf{V}}{c^2}\right) + 1 \qquad 20.3\text{-}28$$

összefüggések adják. 20.3-28-ból meghatározhatjuk az $\mathbf{r}^*(t^*)$ függvényt. Mivel

$$t = \frac{t^*}{B\left(1 + \frac{\mathbf{w}\mathbf{V}}{c^2}\right)}, \qquad 20.3\text{-}29$$

$$\mathbf{r}^*(t^*) = \frac{\mathbf{G}\mathbf{V} + \mathbf{v}B}{B\left(1 + \frac{\mathbf{w}\mathbf{V}}{c^2}\right)}, \qquad 20.3\text{-}30$$

$\mathbf{r}^*(t^*)$ tehát az

$$\mathbf{r}^*(t^*) = \mathbf{V}^* t^* \qquad 20.3\text{-}31$$

alakban is felírható, ahol

$$\mathbf{V}^* = \frac{\mathbf{G}\mathbf{V} + \mathbf{v}B}{B\left(1 + \frac{\mathbf{w}\mathbf{V}}{c^2}\right)}. \qquad 20.3\text{-}32$$

20.3-31 és 20.3-32 mutatja, hogy a $\mathbf{V}$ sebességgel mozgó rendszer Lorentz-deformáltja $\mathbf{V}^*$ sebességgel halad. 20.3-32 az ún. Einstein-féle sebesség-összeadási törvény általános alakja. A formulát áttekinthetőbb alakra hozhatjuk, ha felhasználjuk a $\mathbf{G}$ mátrix

$$\mathbf{G} = \mathbf{O} + (B - 1)\frac{\mathbf{v} \circ \mathbf{w}}{v^2} \qquad 20.3\text{-}33$$

alakját és a

$$\mathbf{v} = \mathbf{O}\mathbf{w}$$

összefüggést. Beírva ezt 20.3-33-ba és felhasználva, hogy $v^2 = w^2$, a formula a

$$\mathbf{G} = \mathbf{O}\left[\mathbf{E} + (B-1)\frac{\mathbf{w} \circ \mathbf{w}}{w^2}\right] \qquad 20.3\text{-}34$$

alakot ölti. Írjuk fel még az eredeti rendszer **V** sebességét **w**-vel párhuzamos $\mathbf{V}_1$ és **w**-re merőleges $\mathbf{V}_2$ vektorok összegeként. Legyen tehát

$$\mathbf{V} = \mathbf{V}_1 + \mathbf{V}_2, \qquad 20.3\text{-}35$$

ahol

$$\mathbf{w}\mathbf{V} = \mathbf{w}\mathbf{V}_1 \quad \text{és} \quad \mathbf{w}\mathbf{V}_2 = 0. \qquad 20.3\text{-}36$$

20.3-36 felhasználásával

$$(B-1)\frac{\mathbf{w} \circ \mathbf{w}}{w^2}V = (B-1)\frac{\mathbf{w}(\mathbf{w}V)}{w^2} = (B-1)\mathbf{V}_1, \qquad 20.3\text{-}37$$

következésképpen

$$\mathbf{G}\mathbf{V} = \mathbf{O}(\mathbf{V}_1 + \mathbf{V}_2 + (B-1)\mathbf{V}_1) = \mathbf{O}(\mathbf{V}_2 + B\mathbf{V}_1). \qquad 20.3\text{-}38$$

20.3-38 felhasználásával a 20.3-32 Einstein-féle sebesség-összeadási törvény a

$$\mathbf{V}^* = \frac{\mathbf{O}(\mathbf{V}_2 + B\mathbf{V}_1) + Bv}{B\left(1 + \frac{\mathbf{V}\mathbf{w}}{c^2}\right)} \qquad 20.3\text{-}39$$

alakot ölti.

Foglalkozzunk most az $\mathbf{O} = \mathbf{E}$ esettel. Ekkor $\mathbf{v} = \mathbf{w}$. Így

$$\mathbf{V}^* = \frac{\mathbf{V}_1 + \frac{1}{B}\mathbf{V}_2 + \mathbf{v}}{1 + \frac{\mathbf{v}\mathbf{V}}{c^2}}. \qquad 20.3\text{-}40$$

Amennyiben $\mathbf{V}_2 = 0$, tehát $\mathbf{v} = \mathbf{w}$ párhuzamos **V**-vel, akkor $\mathbf{V}^*$ is ebbe az irányba mutat, és a sebesség-összeadási formula az egyszerű

$$V^* = \frac{V + v}{1 + \frac{vV}{c^2}} \qquad 20.3\text{-}41$$

alakra redukálódik. 20.3-41-ben a vektorjeleket elhagytuk, mert a formulában minden sebesség iránya azonos. Nagyon gyakran ezt a formulát nevezik az „Einstein-féle sebesség-összeadási törvénynek”. Ez azonban — amint a fentiekben megmutattuk — egy általános összefüggésnek csak speciális esete.

A következőkben a 20.3-41 formula két alkalmazását tárgyaljuk.

1. Vizsgáljuk egy fényjel Lorentz-deformáltját, tehát legyen $V=c$. Ekkor

$$V^*=\frac{c+v}{1+\frac{cv}{c^2}}=c, \tag{20.3-42}$$

vagyis a Lorentz-deformáció a fényjelek sebességét nem változtatja.

2. Legyen $V=\frac{c}{n}$, ahol V egy $n>1$ törésmutatójú anyagban terjedő fényhullám fázissebessége. Amennyiben az adott anyag a $\mathcal{K}$ rendszerhez képest nyugalomban van, a rendszer Lorentz-deformáltja egy v sebességgel áramló folyadék, amelyben a fényhullám

$$V^*=\frac{\frac{c}{n}+v}{1+\frac{v}{nc}} \tag{20.3-43}$$

fázissebességgel halad. 20.3-43-at $\frac{v}{c}$ szerint sorba fejtve, azt kapjuk, hogy

$$V^*=\frac{c}{n}+v\left(1-\frac{1}{n^2}\right)+\text{magasabb rendű tagok.} \tag{20.3-44}$$

20.3-44 egyenlet megegyezik a mozgó folyadékban terjedő hullámok fázissebességére Fizeau által kísérletileg megállapított összefüggéssel.

20.3.3. A Lorentz-kontrakció

Legyen **P** egy $\mathcal{K}$ rendszerhez képest egyenletes sebességgel mozgó pontrendszer. Tehát

$$\mathbf{r}^{(k)}(t)=\mathbf{r}^{(k)}(0)+\mathbf{V}t \quad (k=1, 2, \ldots, n). \tag{20.3-45}$$

Legyen továbbá a rendszer $\mathbf{r}^{(k)}$-adik és $\mathbf{r}^{(l)}$-edik pontjának távolsága

$$\mathbf{r}^{(k)}(t)-\mathbf{r}^{(l)}(t)=\mathbf{r}^{(k)}(0)-\mathbf{r}^{(l)}(0)=\mathbf{R}. \tag{20.3-46}$$

Vizsgáljuk meg, hogyan változik **R** a rendszer Lorentz-deformációja során. Az **R** távolság **R*** Lorentz-deformáltja az

$$\mathbf{r}^{(k)*}=\mathbf{G}\mathbf{r}^{(k)}+B\mathbf{v}t^{(k)}, \tag{20.3-47}$$

$$t^{(k)*}=B\left(\frac{\mathbf{w}\mathbf{r}^{(k)}}{c^2}+t^{(k)}\right) \quad (k=1,\ldots,n) \tag{20.3-48}$$

formulák alapján határozható meg. 20.3-47 szerint

$$\mathbf{R}^* = \mathbf{G}\mathbf{R} + \mathbf{v}B(t^{(k)} - t^{(l)}). \qquad 20.3\text{-}49$$

Határozzuk meg az $\mathbf{R}^*$ vektort egy $t = t^{(k)*} = t^{(l)*}$ időpillanatban. 20.3-48 szerint ekkor

$$0 = t^{(k)*} - t^{(l)*} = B\left(\frac{\mathbf{w}\mathbf{R}}{c^2} + t^{(k)} - t^{(l)}\right), \qquad 20.3\text{-}50$$

azaz

$$t^{(k)} - t^{(l)} = -\frac{\mathbf{w}\mathbf{R}}{c^2}. \qquad 20.3\text{-}51$$

20.3-51-et behelyettesítve 20.3-49-be, az

$$\mathbf{R}^* = \left[G - \frac{B}{c^2}(\mathbf{v} \circ \mathbf{w})\right]\mathbf{R} \qquad 20.3\text{-}52$$

eredmény adódik. Felhasználva a

$$\mathbf{G} = \mathbf{O}\left[E + (B-1)\frac{\mathbf{w} \circ \mathbf{w}}{w^2}\right] \qquad 20.3\text{-}53$$

(20.3-34) formulát, 20.3-52 az

$$\mathbf{R}^* = \mathbf{O}\left[\mathbf{E} + (B-1)\frac{\mathbf{w} \circ \mathbf{w}}{w^2} - \frac{B}{c^2}(\mathbf{w} \circ \mathbf{w})\right]\mathbf{R} \qquad 20.3\text{-}54$$

alakot ölti. Vonjuk össze a szögletes zárójelben levő második két tagot:

$$\left(B - 1 - B\frac{w^2}{c^2}\right)\frac{\mathbf{w} \circ \mathbf{w}}{w^2} = \left(\frac{1}{B} - 1\right)\frac{\mathbf{w} \circ \mathbf{w}}{w^2}. \qquad 20.3\text{-}55$$

Itt felhasználtuk, hogy

$$1 - \frac{w^2}{c^2} = \frac{1}{B^2}.$$

20.3-55 felhasználásával 20.3-54

$$\mathbf{R}^* = \mathbf{O}\left[R + \left(\frac{1}{B} - 1\right)\frac{\mathbf{w}(\mathbf{w}R)}{w^2}\right] \qquad 20.3\text{-}56$$

alakban írható fel. Mivel

$$\frac{\mathbf{w}(\mathbf{w}R)}{w^2} = \mathbf{R}_1 \qquad 20.3\text{-}57$$

az $\mathbf{R}$ vektor $\mathbf{w}$-vel párhuzamos komponense, célszerű $\mathbf{R}$-et az

$$\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2 \qquad 20.3\text{-}58$$

alakban felírni, ahol $\mathbf{w}\mathbf{R}_2 = 0$, tehát $\mathbf{R}_2$ az $\mathbf{R}$ vektor $\mathbf{w}$-re merőleges komponense.

A 20.3-58 felbontással az

$$\mathbf{R}^* = \mathbf{O}\left[\mathbf{R}_2 + \frac{1}{B}\mathbf{R}_1\right] \qquad 20.3\text{-}59$$

összefüggéshez jutunk.

20.3-59 szerint a Lorentz-deformáció az $\mathbf{R}$ vektor $\mathbf{w}$-re merőleges komponensén nem változtat, a $\mathbf{w}$-vel párhuzamos $\mathbf{R}_1$ összetevő azonban

$$\frac{1}{B} = \sqrt{1 - \frac{v^2}{c^2}}$$

-szeresére változik. 20.3-59-ből $\mathbf{O}\tilde{\mathbf{O}} = \mathbf{E}$ és $\mathbf{R}_1\mathbf{R}_2 = 0$ felhasználásával következik, hogy

$$R^{*2} = R_2^2 + \left(1 - \frac{v^2}{c^2}\right)R_1^2. \qquad 20.3\text{-}60$$

A 20.3-60 formula azt a Lorentz és Fitzgerald által feltételezett jelenséget fejezi ki, amit Lorentz-kontrakciónak is nevezünk.

20.4. Koordináta-transzformációk és Lorentz-deformációk

Az előzőekben láttuk, hogy a Lorentz-transzformációknak kettős jelentése van. Hasonló kettősséggel már a háromdimenziós esetben is találkoztunk. Az $\underline{\mathbf{r}}_1, \underline{\mathbf{r}}_2, \ldots, \underline{\mathbf{r}}_n$ helyvektorokkal jellemzett $P_1, \ldots, P_n$ pontrendszert egy elmozgatással az $\underline{\mathbf{r}}_1^* \ldots \underline{\mathbf{r}}_n^*$ helyvektorú $P_1^* \ldots P_n^*$ pontrendszerbe vihetjük át. Ha a pontrendszer egy merev test pontjait jelentette, akkor az elmozgatást egy $\underline{\mathbf{O}}$ elforgatás és egy $\underline{\mathbf{a}}$ vektorral jellemzett eltolás együttesével adhatjuk meg. Tehát

$$\underline{\mathbf{r}}^{(k)*} = \underline{\mathbf{O}}\,\underline{\mathbf{r}}^{(k)} + \underline{\mathbf{a}}. \qquad 20.4\text{-}1$$

A 20.4-1 formulát egy $\mathcal{K}$ reprezentációban

$$\mathcal{K}(\underline{\mathbf{O}}) = \mathbf{O}, \quad \mathcal{K}(\underline{\mathbf{r}}^{(k)}) = \mathbf{r}^{(k)}, \quad \mathcal{K}(\underline{\mathbf{r}}^{(k)*}) = \mathbf{r}^{(k)*} \qquad 20.4\text{-}2$$

mennyiségekkel az

$$\mathbf{r}^{(k)*} = \mathbf{O}\mathbf{r}^{(k)} + \mathbf{a} \quad (k = 1, \ldots, n) \qquad 20.4\text{-}3$$

alakban fejezhetjük ki.

Mivel egy merev test pontjait transzformáltuk, az egyes pontok

$$\underline{\mathbf{r}}^{(k)} - \underline{\mathbf{r}}^{(l)} = \underline{\mathbf{R}} \qquad 20.4\text{-}4$$

17 42235/I. Vektorszámítás

távolságai 20.4-3-nak megfelelően az

$$\underline{\mathbf{R}}^* = \underline{\mathbf{O}}\,\underline{\mathbf{R}} \qquad 20.4\text{-}5$$

formula szerint változnak. 20.4-5 a $\mathcal{K}$ reprezentációban az

$$\mathbf{R}^* = \mathbf{O}\mathbf{R} \qquad 20.4\text{-}6$$

képlettel adható meg.

A fenti matematikai kifejezések nemcsak a $P_1 \dots P_n$ pontokból álló merev test elmozdulásaként értelmezhetők. Amennyiben ugyanis a $\mathcal{K}$ koordináta-rendszer tengelyeit egy merev testnek tekintjük és ezt a testet $\underline{\mathbf{O}}$-val elforgatjuk, akkor egy $\mathcal{K}'$ koordináta-rendszert kapunk. A $P_1, \dots, P_n$ pontokból álló merev test pontjait összekötő vektorok $\mathcal{K}$ és $\mathcal{K}'$-beli reprezentációja között, mint tudjuk, az

$$\mathbf{R}' = \tilde{\mathbf{O}}\mathbf{R} \qquad 20.4\text{-}7$$

összefüggés áll fenn.

Az $\underline{\mathbf{O}}$ ortogonális operátor tehát egyrészt egy merev pontrendszer elmozdulásának, másrészt egy koordinátareprezentációbeli változásnak a leírására is alkalmas.

Ez a kettősség a Lorentz-transzformáció értelmezésében is fennáll.

$$\underline{\mathbf{x}}^* = \underline{\mathbf{\Lambda}}\,\underline{\mathbf{x}} \qquad 20.4\text{-}8$$

azt a változást adja meg, amit egy fizikai rendszer külső beavatkozások hatására elszenved.

A változás egy $\mathcal{K}$ rendszerben az

$$\mathbf{x}^* = \mathbf{\Lambda}\mathbf{x} \qquad 20.4\text{-}9$$

mátrixegyenlettel jellemezhető.

20.4-9-hez formálisan hasonló az

$$\mathbf{x}' = \mathbf{\Lambda}^{-1}\mathbf{x} \qquad 20.4\text{-}10$$

egyenlet, amely azt fejezi ki, hogy $\underline{\mathbf{x}}$ $\mathcal{K}$-beli $\mathbf{x}$ és $\mathcal{K}'$-beli $\mathbf{x}'$ reprezentációja között milyen összefüggés áll fenn.

20.4-6 és 20.4-9 fizikailag összefüggő, de nem merev rendszerek tulajdonságait írja le. 20.4-7 és 20.4-10 azért ad hasonló eredményt, mert ez esetben a koordináta-rendszert kezeljük fizikai rendszerként. Az összefüggések értelmezése a háromdimenziós forgatások értelmezéséhez hasonlóan történhet, ennek részleteire azonban most nem térünk ki.

Minthogy a matematikai eljárás a koordináta-transzformáció és az elmozgatás esetén azonos, könnyen összetéveszthetjük őket.

A probléma illusztrációjaként röviden foglalkozzunk még a sebesség-összeadási törvény értelmezésével.

Legyen $\mathbf{x}'(p)$ egy pont pályája a $\mathcal{K}'$ rendszer mértékeiben kifejezve. A pont sebessége:

$$\mathbf{V}=\frac{dr'(p)}{dt}=\frac{dr'(p)}{dp}\frac{dp}{dt},$$

Amennyiben a $\mathcal{K}'$ rendszer a $\mathcal{K}$ rendszerhez viszonyítva **w** sebességgel mozog, akkor 20.3-32 szerint a pont sebességét $\mathcal{K}$ mértékeiben a

$$\mathbf{V}=\frac{dr(p)}{dt}=\frac{\tilde{\mathbf{G}}\mathbf{V}'+\mathbf{w}B}{B\left(1+\frac{\mathbf{V}'\mathbf{w}}{c^2}\right)} \qquad 20.4\text{-}11$$

formula adja.

A $\mathcal{K}$-hoz viszonyított sebesség nem a $\mathbf{V}'$ és $\mathbf{w}$ sebességek egyszerű vektori összege — hiszen $\mathbf{V}'$ és $\mathbf{w}$ más mértékrendszerben fejezik ki a megfelelő sebességeket —, hanem 20.4-11, amit szimbolikusan a

$$\mathbf{V}=\mathbf{V}'\hat{+}\mathbf{w}$$

jelöléssel fejezünk ki.

Abban az egyszerű esetben, amikor $\mathbf{O}=\mathbf{E}$, valamint $\mathbf{V}'$ és $\mathbf{w}$ párhuzamosak,

$$\mathbf{V}'\hat{+}\mathbf{w}=\frac{V+w}{1+\frac{Vw}{c^2}}. \qquad 20.4\text{-}12$$

A 20.4-11, ill. 20.4-12 formulák a különböző koordináta-rendszerekben megadott sebességmértékek esetén adják meg a sebességösszeg-értéket az egyik koordináta-rendszer mértékeiben.

20.5. A négyesvektorok

Vonatkoztatási rendszerekben egy esemény az $\mathbf{x}=x_1, x_2, x_3, x_4$ számnégyessel jellemezhető, ahol $\mathbf{x}$ első három komponense az esemény helyét, x_4 pedig időpontját adja meg a $\mathcal{K}$ rendszer mértékeiben.

A 20.2. fejezetben láttuk, hogy amennyiben olyan vonatkoztatási rendszerekre szorítkozunk, amelyek mértékeiben a fény izotrop módon terjed, akkor egy esemény $\mathcal{K}$ és $\mathcal{K}'$ rendszerbeli jellemzőit az

$$\mathbf{x}'=\mathbf{\Lambda}\mathbf{x}+\boldsymbol{\lambda} \qquad 20.5\text{-}1$$

transzformáció köti össze. 20.5-1-ben $\mathbf{\Lambda}$ Lorentz-mátrix, amelyre

$$\tilde{\mathbf{\Lambda}}\mathbf{\Gamma}\mathbf{\Lambda}=\mathbf{\Gamma}, \qquad 20.5\text{-}2$$

$\boldsymbol{\lambda}$ pedig tetszőleges négykomponensű mennyiség. Két esemény távolságát az

$$\mathbf{X}=\mathbf{x}^{(2)}-\mathbf{x}^{(1)}$$

formulával definiáltuk. A 20.5-1 transzformációs törvény szerint a távolságok mértékei az

$$\mathbf{X}' = \Lambda\mathbf{X} \qquad 20.5\text{-}3$$

egyenletnek megfelelően transzformálódnak. Mivel mind az $\mathbf{X}'$, mind az $\mathbf{X}$ ugyanazon két esemény tér- és időbeli eltérését jellemzi, bevezethetjük a

$$\mathcal{K}(\underline{\mathbf{X}}) = \mathbf{X} \quad \text{és} \quad \mathcal{K}'(\underline{\mathbf{X}}) = \mathbf{X}'$$

jelölést, hiszen $\mathbf{X}$ és $\mathbf{X}'$ valóban ugyanazon fizikai objektum leírására szolgál.

20.5.1. A négyesvektorok tulajdonságai

A fenti módon definiált négyesvektorok a hármasvektorokhoz hasonló tulajdonságokkal rendelkeznek.

Legyenek $\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \ldots$ négyesvektorok, melyek a $\mathcal{K}$ reprezentációban rendre az $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ négykomponensű mennyiségekkel adhatók meg. Képezzük az

$$\mathbf{y} = \sum_k \alpha_k \mathbf{x}^{(k)} \qquad 20.5\text{-}4$$

lineáris kombinációt, majd transzformáljuk a Λ Lorentz-mátrixszal. A homogén lineáris tulajdonsága miatt

$$\mathbf{y}' = \Lambda\mathbf{y} = \sum_k \alpha_k \Lambda\mathbf{x}^{(k)} = \sum_k \alpha_k \mathbf{x}'^{(k)}, \qquad 20.5\text{-}5$$

vagyis ha 20.5-4 egy adott reprezentációban érvényes, akkor tetszőleges reprezentációban fennáll. Írhatjuk tehát, hogy

$$\sum_k \alpha_k \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{y}}. \qquad 20.5\text{-}6$$

Négyesvektorok lineáris kombinációja tehát szintén négyesvektor.

20.5.2. Négyesvektorok skaláris szorzata

Két négyesvektor skaláris szorzatát az

$$\mathbf{x}\mathbf{y} = \mathbf{x}\,\Gamma\mathbf{y} \qquad 20.5\text{-}7$$

formulával definiálhatjuk. Belátjuk ugyanis, hogy az $\mathbf{x}\,\Gamma\mathbf{y}$ mennyiség reprezentációtól független.

Mivel

$$\mathbf{x}' = \Lambda\mathbf{x} = \mathbf{x}\,\widetilde{\Lambda},$$

$$\mathbf{y}' = \Lambda \mathbf{y}, \qquad 20.5\text{-}8$$

$$\mathbf{x}'\Gamma\mathbf{y}' = \mathbf{x}'(\tilde{\Lambda}\Gamma\Lambda)\mathbf{y}' = \mathbf{x}\Gamma\mathbf{y},$$

tehát 20.5-7 valóban invariáns mennyiség. A 20.5-7 összefüggés emlékeztet a háromdimenziós esetben a ferdeszögű koordináta-rendszerekben kifejezett

$$\underline{\mathbf{a}}\underline{\mathbf{b}} = \bar{\mathbf{a}}\mathbf{G}\bar{\mathbf{b}} \qquad 20.5\text{-}9$$

skaláris szorzatra, ahol $\bar{\mathbf{a}}$ és $\bar{\mathbf{b}}$ az $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ vektorok kontravariáns, $\mathbf{G}$ pedig a metrikus tenzor kovariáns reprezentációja.

20.5-9 analógiájára $\mathbf{x}$-et és $\mathbf{y}$-t az $\underline{\mathbf{x}}$ és $\underline{\mathbf{y}}$ négyesvektorok *kontravariáns* reprezentációjának tekinthetjük, ha Γ a mértéktenzor kovariáns reprezentációja. Tehát a továbbiakban $\mathbf{x}$ és $\mathbf{y}$ helyett $\bar{\mathbf{x}}$ és $\bar{\mathbf{y}}$-t írunk.

20.5-7 és 20.5-9 között azonban lényeges különbség van, mert míg a háromdimenziós esetben

$$\det \mathbf{G} > 0,$$

Γ definíciójából következik, hogy

$$\det \Gamma = -c^2. \qquad 20.5\text{-}10$$

Ezt úgy fejezzük ki, hogy a $\underline{\Gamma}$ mértéktenzor *indefinit* metrikát képvisel. A $\det \mathbf{G} > 0$ esetben ugyanis

$$\underline{\mathbf{a}}^2 > 0, \quad \text{ha} \quad \underline{\mathbf{a}} \neq 0,$$

az indefinit esetben azonban ez nem teljesül.

A kísérleti tapasztalatok, amelyek elemzésére itt nem térhetünk ki, a feltevést alátámasztják.

Az indefinit metrika ellenére a 20.5-7 által definiált skaláris szorzat elfogadható és hasznosnak bizonyul.

A hármasvektorok analógiájára bevezethetjük a négyesvektorok kovariáns reprezentációját is az

$$\mathbf{x} = \Gamma\bar{\mathbf{x}} \qquad 20.5\text{-}11$$

definíciós egyenlettel, ahol $\bar{\mathbf{x}}$ az $\underline{\mathbf{x}}$ kontravariáns, $\mathbf{x}$ pedig a kovariáns reprezentációját jelenti. 20.5-11-et komponensekben kifejezve, az

$$\begin{aligned} x_k &= x^k, \\ x_4 &= -c^2 x^4 \end{aligned} \quad (k = 1, 2, 3) \qquad 20.5\text{-}12$$

egyenleteket kapjuk. A kétféle reprezentáció tehát csak a negyedik koordinátában tér el.

Meghatározzuk még a kovariáns reprezentáció koordináta-transzformációjának szabályát.

A kontravariáns reprezentáció definíció szerint az

$$\bar{\mathbf{x}}' = \mathbf{\Lambda}\bar{\mathbf{x}} \qquad 20.5\text{-}13$$

formula szerint transzformálódik. Beírva ide 20.5-11 inverzét,

$$\bar{\mathbf{x}}' = \mathbf{\Lambda}\mathbf{\Gamma}^{-1}\mathbf{x}$$

és

$$\mathbf{x}' = \mathbf{\Gamma}\bar{\mathbf{x}}' = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}^{-1}\mathbf{x} \qquad 20.5\text{-}14$$

adódik. 20.5-2-ből következik azonban, hogy

$$\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}^{-1} = \tilde{\mathbf{\Lambda}}^{-1}, \qquad 20.5\text{-}15$$

tehát a kovariáns reprezentáció az

$$\mathbf{x}' = \tilde{\mathbf{\Lambda}}^{-1}\mathbf{x} \qquad 20.5\text{-}16$$

szabály szerint transzformálódik.

20.5.3. Példa a skaláris szorzás alkalmazására

Bizonyítás nélkül közöljük, hogy egy pontszerű részecske **p** impulzusa és E energiája egy kovariáns négyesvektor négy komponensét alkotja. Itt $E=mc^2$, az Einstein-féle formulának megfelelően.

Tehát

$$\mathbf{P} = \mathbf{p},\ mc^2. \qquad 20.5\text{-}17$$

Abban a $\mathcal{K}$ koordináta-rendszerben, amelyben a részecske nyugalomban van, $\mathbf{p}=0$ és $E=m_0c^2$. Tehát

$$\mathbf{P} = \mathcal{K}(\underline{\mathbf{P}}) = 0,\ 0,\ 0,\ m_0c^2. \qquad 20.5\text{-}18$$

Alkalmazzuk a

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{E} + (B-1)\dfrac{\mathbf{v}\circ\mathbf{v}}{v^2} & B\mathbf{v} \\ \dfrac{B\mathbf{v}}{c^2} & B \end{pmatrix} \qquad 20.5\text{-}19$$

Lorentz-transzformációt. Minthogy **P** kovariáns vektor, **P′** a

$$\tilde{\mathbf{\Lambda}}^{-1} = \begin{pmatrix} \mathbf{E} + (B-1)\dfrac{\mathbf{v}\circ\mathbf{v}}{v^2} & \dfrac{B\mathbf{v}}{c^2} \\ B\mathbf{v} & B \end{pmatrix} \qquad 20.5\text{-}20$$

mátrix segítségével határozható meg. A

$$\mathbf{P}' = \tilde{\Lambda}^{-1}\mathbf{P}$$

formulából

$$\mathbf{P}' = m_0 B\mathbf{v}, \quad B m_0 c^2 \tag{20.5-21}$$

adódik. Helyettesítsük be ide az

$$m_o B = \frac{m_o}{\sqrt{1-\frac{v^2}{c^2}}} = m \tag{20.5-22}$$

összefüggést. A

$$\mathbf{P}' = m\mathbf{v},\ mc^2 \tag{20.5-23}$$

formulához jutunk. A 20.5-23 összefüggés mutatja, hogy 20.5-17 a koordinátatranszformáció során megtartja eredeti alakját, ha feltételezzük, hogy a részecske tömege 20.5-22-nek megfelelően változik a sebességgel.

Érdemes még a **P** vektor négyzetét is meghatározni. Definíció szerint

$$\mathbf{P}^2 = \mathbf{P}\,\Gamma^{-1}\mathbf{P}. \tag{20.5-24}$$

Beírva ide **P** értékét, a

$$\mathbf{P}^2 = -m_0^2 c^2$$

eredmény adódik. A Λ által meghatározott $\mathcal{K}'$ rendszerbeli reprezentációt használva,

$$\mathbf{P}^2 = \mathbf{p}^2 - \frac{E^2}{c^2}.$$

Minthogy a skaláris szorzat invariáns,

$$\mathbf{p}^2 - \frac{E^2}{c^2} = -m_0^2 c^2.$$

Innen

$$E = \pm c\sqrt{p^2 + m_0^2 c^2}. \tag{20.5-25}$$

20.5-25 a relativisztikus energia — impulzus kapcsolatot adja meg.

20.6. A tér empirikus dimenziószáma

Megállapítottuk, hogy a fényterjedés leírásakor célszerű a négykomponensű mennyiségek használata. A helyvektorok meghatározására azonban általában háromkomponensű mennyiségeket használunk. A következőkben részletesebben vizsgáljuk, hogy mit jelent az, hogy a tér háromdimenziós.

A megállapítás, mint látni fogjuk, kísérleti tapasztalat eredménye. A tér pontjait helyvektorokkal jellemezhetjük, mérni azonban csak a tér pontjai közötti távolságokat lehet. Felvetődhet tehát a kérdés, hogy adott $P_0, P_1, \ldots, P_N$ pontok közötti r_{kl} távolságok ismeretében hogyan határozhatjuk meg az egyes pontok helyvektorait. (r_{kl} a P_k és P_l pontok távolságát jelenti.) A $\underline{\mathbf{G}}$ metrikus tenzorral jellemzett ferdeszögű koordináta-rendszerben a P_k pontok $\underline{\mathbf{r}}^{(k)}$ helyvektorai és az r_{kl} távolságértékek között az

$$(\mathbf{r}^{(k)}-\mathbf{r}^{(l)})\mathbf{G}(\mathbf{r}^{(k)}-\mathbf{r}^{(l)})=r_{kl}^2 \quad (k, l=0, 1, 2, \ldots, N) \qquad 20.6\text{-}1$$

összefüggésnek kell fennállnia.

20.6-1 G és r_{kl} $(k, l=0, 1, \ldots, N)$ ismeretében egy egyenletrendszert szolgáltat a helyvektorok meghatározására. Az egyenletrendszer csak akkor oldható meg, ha a mérések azt mutatják, hogy

$$r_{kl}=r_{lk}.$$

A tapasztalat azt mutatja, hogy az egyes helyvektorok három komponenssel jellemezhetők. Így elegendően nagy N esetén a 20.6-1 egyenletrendszer túlhatározott, hiszen az $\frac{N(N+1)}{2}$ egyenlet csak $3(N+1)$ ismeretlent tartalmaz. Ez arra mutat, hogy a mért r_{kl} távolságértékek között meghatározott kapcsolatoknak kell fennállni. Minthogy pontosan ezek a kapcsolatok teszik lehetővé, hogy a tér pontjait háromkomponensű mennyiségekkel jellemezzük, ezek az összefüggések jellemzik a tér háromdimenziós voltát is.

A jellegzetes összefüggések pontosabb megismerésére általánosítsuk a 20.6-1 összefüggést. Teljesen formálisan tételezzük fel, hogy **G** $N\times N$-es mátrix és az $\mathbf{r}^{(k)}$ mennyiségek N komponensűek.

Használjuk tehát 20.6-1 helyett az

$$(\mathbf{x}^{(k)}-\mathbf{x}^{(l)})\mathbf{G}(\mathbf{x}^{(k)}-\mathbf{x}^{(l)})=r_{kl}^2 \qquad 20.6\text{-}2$$

összefüggést, ahol

$$\mathbf{x}^{(k)}=x_1^{(k)}, x_2^{(k)}, \ldots, x_N^{(k)} \quad (k=0, 1, \ldots, N)$$

komponensű mennyiséget jelöl.

20.6-2 komponensekben kiírva az

$$\sum_{m,n} (x_m^{(k)} - x_m^{(l)}) G_{mn} (x_n^{(k)} - x_n^{(l)}) = r_{kl}^2 \qquad 20.6\text{-}3$$

alakot ölti.

Vezessük be az

$$S_{km} = x_m^{(k)} \qquad 20.6\text{-}4$$

mátrixot, és minthogy 20.6-2-ben csak különbségek szerepelnek, tételezzük fel, hogy

$$\mathbf{x}^{(0)} = \mathbf{0}, \quad \text{azaz} \quad x_m^{(0)} = 0 \quad (m = 1, \ldots, N). \qquad 20.6\text{-}5$$

20.6-4 és 20.6-5 felhasználásával 20.6-3-at az

$$(\mathbf{SG}\tilde{\mathbf{S}})_{kk} + (\mathbf{SG}\tilde{\mathbf{S}})_{ll} - 2(\mathbf{SG}\tilde{\mathbf{S}})_{kl} = r_{kl}^2 \qquad 20.6\text{-}6$$

alakban írhatjuk fel. Mivel

$$(\mathbf{SG}\tilde{\mathbf{S}})_{kk} = r_{k0}^2,$$

20.6-6 egyenértékű az

$$\mathbf{SG}\tilde{\mathbf{S}} = \mathbf{R} \qquad 20.6\text{-}7$$

formulával, ahol

$$R_{kl} = \frac{1}{2}(r_{k0}^2 + r_{l0}^2 - r_{kl}^2). \qquad 20.6\text{-}8$$

20.6-8 egy szimmetrikus mátrixot definiál, amelynek elemei méréssel, tehát empirikusan határozhatók meg.

Válasszunk most S-ként egy tetszőleges $N \times N$-es mátrixot, amelyre $\det \mathbf{S} \neq 0$. 20.6-7-ből ekkor a

$$\mathbf{G} = \mathbf{S}^{-1}\mathbf{R}\tilde{\mathbf{S}}^{-1} \qquad 20.6\text{-}9$$

összefüggés adódik. Az **S** mátrix elemei a 20.6-9-nek megfelelő $\underline{\mathbf{G}}$ metrikus tenzorral meghatározott koordináta-rendszerben adják meg az egyes helyvektorok koordinátáit. N komponensű helyvektorokkal tehát mindig kielégíthetjük a mért távolságértékekre és a helyvektorokra vonatkozó 20.6-2 összefüggést.

Vizsgáljuk meg most, hogy mi a feltétele annak, hogy a helyvektorokat kevesebb, pl. 3 komponensű mennyiségekkel jellemezhessük. Megállapítható, hogy a tér pontjait többféle helyvektorral is jellemezhetjük. Amennyiben ugyanis az $\mathbf{x}^{(k)}$ vektorok kielégítik a 20.6-2 összefüggést, akkor az $\mathbf{x}^{(k)\prime}$ vektorok is kielégítik, ha

$$\mathbf{x}^{(k)}\mathbf{G}\mathbf{x}^{(l)} = \mathbf{x}^{(k)\prime}\mathbf{G}\mathbf{x}^{(l)\prime}. \qquad 20.6\text{-}10$$

20.6-10 az

$$(\mathbf{x}^{(k)} - \mathbf{x}^{(k)\prime})\mathbf{G} = 0 \qquad 20.6\text{-}11$$

esetben teljesül.

Amennyiben det $\mathbf{G} \neq 0$, akkor a 20.6-11 egyenletrendszernek csak a triviális megoldása létezik, tehát

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k)\prime},$$

vagyis a helyvektorok megadása egyértelmű.

Ha azonban det $\mathbf{G} = 0$, akkor 20.6-11 nem triviális megoldásokkal is rendelkezik. A lineáris egyenletrendszerek elméletéből tudjuk, hogy

$$\mathcal{R}(\mathbf{G}) = N - n \qquad 20.6\text{-}12$$

esetén az

$$\mathbf{y}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k)\prime}$$

vektorok n komponensét szabadon választhatjuk.

Legyen a szabadon választott komponensek értéke

$$y_\nu^{(k)} = x_\nu^{(k)} \quad (\nu = 1', 2', \ldots, n').$$

Ekkor 20.6-2 $\mathbf{x}^{(k)\prime}$ megoldásrendszerének megfelelő n komponense eltűnik, vagyis a helyvektorokat meghatározó

$$\mathbf{x}^{(k)\prime} = x_1^{(k)\prime}, \ldots, x_N^{(k)\prime} \qquad 20.6\text{-}13$$

komponensű mennyiségekben csak $N-n$ érték nem zérus. Elhagyva a teljesen felesleges azonosan zérus komponenseket, $N-n$ komponensű vektorokat kapunk. Ekkor természetesen a $\mathbf{G}$ mátrix is redukálható a megfelelő $(N-n) \times (N-n)$-es mátrixszá.

Amennyiben

$$\mathcal{R}(\mathbf{G}) = N - n = 3,$$

akkor a 20.6-1 egyenlet kielégíthető, ami azt jelenti, hogy a tér háromdimenziós.

A $\mathbf{G}$ mátrix rangjának megállapítása az empirikus adatok alapján elég bonyolult lenne. 20.6-7 értelmében azonban

$$\mathbf{SG\tilde{S}} = \mathbf{R},$$

és mivel det $\mathbf{S} \neq 0$, a 11.14-1-ben tárgyalt tétel értelmében $\mathbf{G}$ és $\mathbf{R}$ rangja megegyezik. $\mathbf{R}$ elemei empirikusan meghatározhatók, így $\mathbf{R}$ rangjának meghatározásával $\mathbf{G}$ rangja is megállapítható.

A fentiekből következik, hogy a tér dimenziószáma a 20.6-8 mátrix rangjának meghatározásával empirikusan megállapítható.

A tapasztalat azt mutatja, hogy

$$\mathcal{R}(\mathbf{R}) = 3.$$

Egyes konkrét esetekben a mért értékek alapján

$$\mathcal{R}(\mathbf{R}) = 2$$

és

$$\mathcal{R}(\mathbf{R}) = 1$$

is adódhat. Ezekben az esetekben a vizsgált pontok egy síkban, ill. egy egyenesen helyezkednek el. Ekkor a vizsgált pontok helyvektorai két-, ill. egykomponensű mennyiségekkel jellemezhetők.

FÜGGELÉK

Komplex számok

1. Bevezetés

Az 1. fejezetben röviden ismertettük a számfogalom kialakulását, leírtuk a valós számtest felépítését.

Megállapítottuk, hogy a számfogalom bővítése mindig akkor vált szükségessé, amikor egyes műveleteket nem tudtunk elvégezni. A valós számok körében a négy alapművelet mindig elvégezhető.

Egyes algebrai egyenletek megoldásakor azonban még a valós számok halmaza is szűknek bizonyul. A következőkben ezen hiányosság kiküszöbölése céljából tovább bővítjük a számfogalmat.

A bővítés alapelve most is az eddigi műveleti szabályok lehetőség szerinti érvényben tartása lesz. Érvényben akarjuk tehát tartani az 1.2-1—1.2-10 szabályokat.

2. Az imaginárius egység

Az

$$1+x^2=0 \tag{2-1}$$

egyenletnek nincs megoldása a valós számok között. A 2-1 egyenlet megoldását jelöljük

$$\sqrt{-1}=i \tag{2-2}$$

-vel. Nyilvánvaló, hogy a 2-2 szám definíciója tulajdonképpen a 2-1 egyenlet. Az előzőek értelmében tegyük fel, hogy az így definiált i mennyiség eleget tesz az 1.2-1—1.2-10 szabályoknak. Így az ún. komplex számok algebrájához jutunk.

Megjegyezzük, hogy 2-2-ből következik, hogy

$$i^2=-1,$$

tehát mivel $(-1)\, i = -i$,

$$(-i)^2 = -1,$$

vagyis amennyiben 2-1-nek van megoldása, akkor két megoldással kell, hogy rendelkezzen. Ezeket $\pm i$-vel jelöljük.

Szorzással és összeadással i-t és a természetes számokat felhasználva

$$z = a + ib \qquad 2\text{-}3$$

típusú számokhoz jutunk. A 2-3 alakú számokat komplex számoknak nevezzük. A komplex számok egy valós (a) és egy képzetes vagy imaginárius (ib) részből állnak.

3. Komplex számok összege és szorzata

Legyen

$$z_1 = a_1 + ib_1, \quad z_2 = a_2 + ib_2; \qquad 3\text{-}1$$

az asszociatív és disztributív törvényt alkalmazva:

$$z_1 + z_2 = (a_1 + ib_1) + (a_2 + ib_2) = a_1 + a_2 + i(b_1 + b_2), \qquad 3\text{-}2$$

vagyis két komplex szám összegét úgy kapjuk, hogy a reális részeket és az imaginárius részeket külön-külön összeadjuk, és így az összeg reális és imaginárius részéhez jutunk. Hasonló módon adódik a szorzat

$$z_1 z_2 = (a_1 + ib_1)(a_2 + ib_2) = a_1 a_2 - b_1 b_2 + i(a_1 b_2 + a_2 b_1) \qquad 3\text{-}3$$

kifejezése.

A komplex szám konjugáltja

Legyen

$$z = a + ib. \qquad 3\text{-}4$$

A

$$z^* = a - ib \qquad 3\text{-}5$$

komplex számot a z komplex szám konjugáltjának nevezzük. Egy komplex szám és konjugáltja között fennállnak a

$$z + z^* = 2a, \quad zz^* = a^2 + b^2 \qquad 3\text{-}6$$

összefüggések, tehát mind az összeg, mind a szorzat valós számra vezet.

4. Komplex számok osztása

Legyen

$$z_1 z_2 = Z, \qquad 4\text{-}1$$

akkor definíció szerint

$$z_1 = \frac{Z}{z_2}.$$

Ha z_1 és z_2 3-1-nek megfelelő alakú, és Z-t a

$$Z = A + iB \qquad 4\text{-}2$$

alakban keressük, akkor

3-3, 4-1 és 4-2 összefüggésekből következik, hogy

$$A = a_1 a_2 - b_1 b_2,$$
$$B = a_1 b_2 + a_2 b_1. \qquad 4\text{-}3$$

Tekintsük 4-3-ban az a_1-et és b_1-et ismeretlennek. Rendezve az egyenletrendszert:

$$Aa_2 + Bb_2 = a_1(a_2^2 + b_2^2),$$

$$-Ab_2 + Ba_2 = b_1(a_2^2 + b_2^2).$$

Ebből

$$a_1 = \frac{Aa_2 + Bb_2}{a_2^2 + b_2^2}, \quad b_1 = \frac{Ba_2 - Ab_2}{a_2^2 + b_2^2}, \qquad 4\text{-}4$$

ha $a_2^2 + b_2^2 > 0$, vagyis ha $z_2 \neq 0$. Így

$$z_1 = \frac{Aa_2 + Bb_2}{a_2^2 + b_2^2} + i\,\frac{Ba_2 - Ab_2}{a_2^2 + b_2^2}, \qquad 4\text{-}5$$

tehát a $\dfrac{Z}{z_2}$ mennyiséget komplex számként kifejeztük.

Megjegyzés: Minthogy $z_1 z_2 = z_2 z_1$, ezért a fenti kifejezésnél nem kell különbséget tenni, hogy Z-t balról vagy jobbról osztjuk z_2-vel. Különleges eset, ha $Z = 1$. Ebben az esetben $A = 1$, $B = 0$. A

$$z_2 = z = a + ib$$

jelöléssel

$$\frac{1}{z} = \frac{a}{a^2 + b^2} - i\,\frac{b}{a^2 + b^2}. \qquad 4\text{-}6$$

4-6-ot röviden az

$$\frac{1}{z}=\frac{a-ib}{a^2+b^2} \tag{4-7}$$

alakban is kifejezhetjük.

Az osztás műveletét a továbbiakban $\frac{1}{z}$-vel való szorzással is kifejezhetjük.

5. Gyökvonás

Felmerülhet a gyanú, hogy amint $\sqrt{-1}=i$ az imaginárius számokra, $\sqrt{i}$ is újabb számtípusra vezet. Ez azonban nem így van. Legyen

$$z=a+ib, \quad Z=A+iB$$

$$z=\sqrt{Z}, \quad z^2=Z, \tag{5-1}$$

vagyis

$$(a+ib)^2=A+iB,$$

tehát

$$a^2-b^2+2abi=a+iB,$$

és

$$a^2-b^2=A, \quad 2ab=B. \tag{5-2}$$

A fenti egyenletrendszer valós megoldása

$$a=\sqrt{\frac{+\sqrt{A^2+B^2}+A}{2}}, \quad b=\sqrt{\frac{+\sqrt{A^2+B^2}-A}{2}},$$

ahol a külső gyököket egyforma előjellel kell választani.

Tehát

$$z=\pm\left\{\sqrt{\frac{+\sqrt{A^2+B^2}+A}{2}}+i\sqrt{\frac{+\sqrt{A^2+B^2}-A}{2}}\right\}.$$

Ha például $Z=i$, akkor $A=0$, $B=1$ és

$$\sqrt{i}=\pm\frac{1+i}{\sqrt{2}}. \tag{5-3}$$

Ellenőrzésképpen emeljük négyzetre az egyenlet jobb oldalát,

$$\left(\frac{1+i}{\sqrt{2}}\right)^2=\frac{1+2i-1}{2}=i$$

adódik. (Amint várható is.)

Az

$$ax^2+bx+c=0 \qquad 6\text{-}1$$

valós együtthatójú másodfokú egyenlet megoldóképlete,

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \qquad 6\text{-}2$$

a komplex számok bevezetése után a

$$b^2<4ac$$

esetben is értelmet nyer.

Ebben az esetben az egyenletnek két *komplex* megoldása van. 6-2 azonban akkor is megoldást ad, ha a, b komplex számok.

Általánosan kimutatható, hogy az n-ed fokú komplex együtthatós

$$x^n+a_1x^{n-1}+\ldots+a_n=0$$

egyenlet n komplex gyökkel rendelkezik. A gyökök között persze előfordulhatnak többszörös gyökök is. Ezt a tételt nevezik az algebra alaptételének.

7. A komplex számsík

A komplex számokat egy sík pontjaival szemléltetjük. Legyen $\mathcal{K}$ egy derékszögű koordináta-rendszer. A z számot annak a P pontnak feleltetjük meg, amelynek abszcisszája z valós része, ordinátája pedig z imaginárius része. A komplex számokhoz tartozó pontok halmazát komplex számsíknak nevezzük.

A z számot a komplex síkban vektorral ábrázoljuk. Ábrázolva a

$$z=a+ib$$

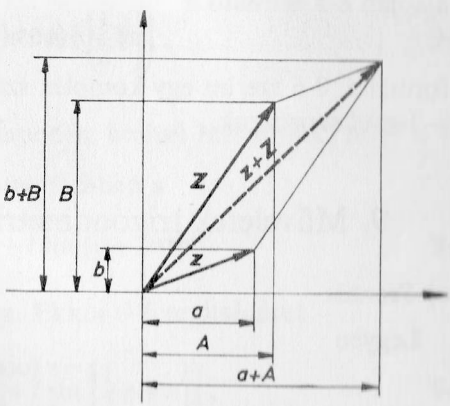
és

$$Z=A+iB \qquad 7\text{-}1$$

komplex számokat, a

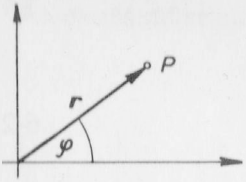
$$z+Z=(a+A)+i(b+B) \qquad 7\text{-}2$$

összegvektorról megállapítható, hogy geometriailag a vektorok összegezésére vonatkozó paralelogrammaszabálynak tesz eleget *(F.1. ábra)*.



F.1. ábra

Ábrázoljuk a



$$z = a + ib \tag{8-1}$$

számot a komplex síkon. Jelöljük P-vel a z-nek megfelelő vektor végpontját. A P pont origótól mért távolságát az *F. 2. ábra* alapján az

F.2. ábra

$$r = \sqrt{zz^*} = \sqrt{a^2 + b^2} \tag{8-2}$$

formulával számíthatjuk ki. Az OP vektor a valós tengellyel φ szöget zár be, melyre

$$\cos\varphi = \frac{a}{r} \quad \sin\varphi = \frac{b}{r}. \tag{8-3}$$

Megjegyzés: A szögnek mind cosinusát, mind sinusát meg kell adni az egyértelmű jellemzéshez.

A z számot 8-2 és 8-3 alapján a

$$z = r(\cos\varphi + i\sin\varphi) \tag{8-4}$$

formulával is megadhatjuk. r-et z modulusának, φ-t pedig fázisának is nevezzük.

8-4 a z komplex szám ún. trigonometrikus alakja.

A z komplex szám konjugáltját definíció szerint a

$$z^* = r(\cos\varphi - i\sin\varphi) \tag{8-5}$$

képlettel határozhatjuk meg. A trigonometrikus függvények tulajdonságai alapján 8-5 átírható a

$$z^* = r(\cos(-\varphi) + i\sin(-\varphi)) \tag{8-6}$$

formába. 8-6 szerint egy komplex szám konjugáltját úgy kapjuk, hogy a fázist -1-gyel szorozzuk.

9. Műveletek trigonometrikus alakban adott számokkal

a) *Szorzás*

Legyen

$$z = r(\cos\varphi + i\sin\varphi),$$

$$Z = R(\cos\Phi + i\sin\Phi), \tag{9-1}$$

ahol természetesen r-et, ill. φ-t, és R-et, ill. Φ-t úgy választjuk, hogy

$$r \cos \varphi = a, \quad r \sin \varphi = b$$

$$R \cos \Phi = A, \quad R \sin \Phi = B.$$

9-1 felhasználásával

$$zZ = rR\,[\cos \varphi \cos \Phi - \sin \varphi \sin \Phi + i\,(\sin \varphi \cos \Phi + \cos \varphi \sin \Phi)],$$

tehát

$$zZ = rR\,(\cos(\varphi + \Phi) + i \sin(\varphi + \Phi)). \qquad 9\text{-}2$$

A trigonometrikus alak segítségével, mint az 9-2-ből látható, a szorzás igen egyszerűen elvégezhető, hiszen a modulusok szorzódnak, a fázisok pedig egyszerűen összeadódnak.

b) *Hatványozás*

A szorzásra vonatkozó 9-2 szabály alapján

$$z^2 = r^2\,(\cos 2\varphi + i \sin 2\varphi), \qquad 9\text{-}3$$

és általában

$$z^n = r^n\,(\cos n\varphi + i \sin n\varphi) \quad (n = 1, 2, \ldots). \qquad 9\text{-}4$$

9-4-et Moivre-féle formulának nevezzük.

c) *Gyökvonás*

A fentiek alapján a gyökvonás is egyszerűen elvégezhető. Legyen

$$z = r\,(\cos \varphi + i \sin \varphi),$$

akkor

$$\sqrt{z} = \underset{+}{\sqrt{r}}\left(\cos\left(\frac{1}{2}\varphi\right) + i \sin\left(\frac{1}{2}\varphi\right)\right). \qquad 9\text{-}5$$

Az, hogy $\underset{+}{\sqrt{r}}$-et használunk, csak megállapodás, amivel biztosítjuk, hogy a $\sqrt{z}$ komplex szám modulusa ne legyen negatív. 9-5-ben a

$$z = r\,(\cos(\varphi + 2\pi) + i \sin(\varphi + 2\pi)) \qquad 9\text{-}6$$

komplex alakot is felhasználhattuk volna. Ekkor 9-6 segítségével

$$\sqrt{z} = \underset{+}{\sqrt{r}}\left(\cos\left(\frac{1}{2}\varphi + \pi\right) + i \sin\left(\frac{1}{2}\varphi + \pi\right)\right), \qquad 9\text{-}7$$

viszont

$$\cos\left(\frac{1}{2}\varphi+\pi\right)=-\cos\frac{1}{2}\varphi,$$

$$\sin\left(\frac{1}{2}\varphi+\pi\right)=-\sin\frac{1}{2}\varphi. \qquad 9\text{-}8$$

9-8-at beírva 9-7-be, a 9-6 és 9-7 formulák összefoglalhatók a

$$\sqrt{z}=\pm\sqrt[+]{r}\left(\cos\frac{1}{2}\varphi+i\sin\frac{1}{2}\varphi\right) \qquad 9\text{-}9$$

alakban.

Látjuk tehát, hogy komplex számok is két négyzetgyökkel rendelkeznek.

d) *Magasabb gyök*

Legyen ismét

$$z=r(\cos\varphi+i\sin\varphi), \qquad 9\text{-}10$$

akkor

$$\sqrt[n]{z}=\sqrt[n]{r}\left(\cos\frac{\varphi}{n}+i\sin\frac{\varphi}{n}\right), \qquad 9\text{-}11$$

ahol $\sqrt[n]{r}$ az r nemnegatív valós szám n-edik gyöke. 9-11 valóban z n-edik gyökét adja, hiszen a jobb oldali kifejezés n-edik hatványa a Moivre-tétel segítségével éppen z-t ad.

9-11-ben a cosinus- és sinusfüggvények periodicitását figyelembe véve, z-t a

$$z=r\left(\cos(\varphi+2\pi k)+i\sin(\varphi+2\pi k)\right) \quad (k=0,\ \pm 1,\ \pm 2,\dots) \qquad 9\text{-}12$$

alakban is felhasználhatjuk.

Így 9-11-en kívül más n-edik gyököket is találunk:

$$\sqrt[n]{z}=\sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n}+i\sin\frac{\varphi+2\pi k}{n}\right) \quad (k=0,1,2,\dots,n-1). \qquad 9\text{-}13$$

$r>0$ esetben 9-13 pontosan n különböző gyököt ad. Válasszuk φ-t a

$$0\leqq\varphi<2\pi$$

intervallumban. 9-13 jobb oldala különböző komplex számokat ad, ha k a $k=0, 1, 2, \dots, n-1$ értékeken fut végig. ($k=n$ esetén ismét az első, $k=0$-val meghatározott számot kapjuk.)

A 9-13 képlet érdekes geometriai jelentéssel bír. A $k=0, 1, 2, \dots, n-1$-hez tartozó gyökök a komplex síkban egy, az origó körül írt szabályos n-szög csúcsait határozzák meg. Ha speciálisan $r=1$ és $\varphi=0$, akkor azt kapjuk, hogy

$$z_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} \quad (k=0, 1, 2, \ldots, n-1). \qquad 9\text{-}14$$

9-14 az egységkörbe írt szabályos n-szög csúcsait adja. Látjuk, hogy komplex számokat használva, az n-edik gyök meghatározása n értékre vezet. Így az

$$x^n - a = 0 \quad (a > 0)$$

egyenlet n gyökkel rendelkezik. Emlékeztetünk a 6. pontban kimondott algebra alaptételére, amely szerint a komplex együtthatós

$$x^n + A_1 x^{n-1} + \ldots + A_n = 0 \qquad 9\text{-}15$$

egyenletnek, ha az A_i számok nem mindegyike zérus, n komplex gyöke van. (A megoldások között lehetnek többszörös gyökök is.)

Belátható, hogy ha az A_k $(k=1, 2, \ldots, n)$ együtthatók valós számok és x 9-15 megoldása, akkor x^* (az x konjugáltja) is megoldás.

A 9-15 egyenlet csak úgy teljesülhet, ha mind a komplex rész, mind a valós rész zérus. Legyen $z = r(\cos\varphi + i\sin\varphi)$ 9-15 megoldása. Ekkor

$$r^n \cos n\varphi + A_1 r^{n-1} \cos(n-1)\varphi + \ldots + A_n = 0,$$

$$i(r^n \sin n\varphi + A_1 r^{n-1} \sin(n-1)\varphi + \ldots + A_{n-1} \sin\varphi) = 0.$$

A fenti két egyenlet minden további nélkül teljesül akkor is, ha a másodikat -1-gyel beszorozzuk. Ez azonban azt jelenti, hogy az eredeti egyenletnek a

$$z^* = r(\cos\varphi - i\sin\varphi)$$

komplex konjugált is gyöke.

Ebből következik, hogy a valós együtthatós egyenletek komplex megoldásai mindig konjugáltjukkal párban lépnek fel.

A páratlan fokszámú egyenleteknek így legalább egy valós megoldása kell hogy legyen, hiszen a páratlan számú gyök nem rendezhető párokba.

Így a harmadfokú egyenletnek vagy egy valós és két komplex, vagy három valós gyöke van.

NÉV- ÉS TÁRGYMUTATÓ

TARTALOMJEGYZÉK

Tankönyvkiadó Vállalat
A kiadásért felelős: Petró András igazgató

79-1161 – Dabasi Nyomda, Budapest – Dabas
Felelős vezető: Bálint Csaba igazgató

Raktári szám: 42 235/I.

Felelős szerkesztő: Ambrus Ferenc
Műszaki vezető: Lojd Lajos
Grafikai szerkesztő: Scnedarek Péter
Műszaki szerkesztő: Szilasy Gyula
A kézirat nyomdába érkezett: 1979. április. Megjelent: 1980. február
Példányszám: 2000. Terjedelem: 25,2 (A/5) ív
Készült: krétáról fotózva, íves ofszetnyomással,
az MSZ 5601–59 és az MSZ 5602–55 szabvány szerint