

MAT-FIZ: Variációszámítás, Tenzor kalkulus, Differenciálgeometria

Szemináriumi jegyzet

Szabó Zsombor

2025. szeptember 14.

Tartalomjegyzék

1. Variációszámítás	4
1.1. Bevezetés	4
1.2. Egydimenziós skalárfüggvények elsőrendű szükséges feltételei: az első variáció — C^1 elmélet	4
1.2.1. Az Euler–Lagrange-egyenlet — erős alak	4
1.2.2. Közjáték a tesztfüggvényekről és a lokális minimalitásról	8
1.2.3. A Du Bois-Reymond-egyenlet — erős alak	10
1.2.4. A minimalizálók létezéséről	10
1.2.5. Egy megmaradási törvény: Beltrami-azonosság	11
1.2.6. Egy regularitási eredmény	11
1.3. Egydimenziós skalárfüggvények elsőrendű szükséges feltételei: az első variáció — C^1 elmélet	12
1.3.1. Az Euler–Lagrange-egyenlet	12
1.3.2. A Du Bois-Reymond-egyenlet	14
1.3.3. Egy regularitási eredmény	15
1.4. Problémák szabad végpontokkal	16
1.5. Izoperimetrikus problémák	17
1.5.1. A Lagrange-multiplikátor szabály variációs problémákra	17
1.6. Elsőrendű szükséges feltételek általános függvényekre	18
1.6.1. Az Euler–Lagrange-egyenlet	18
1.6.2. Természetes peremfeltételek	19
1.6.3. Belső variációk és az energia-impulzus tenzor	19

1.6.4.	Izoperimetrikus problémák	20
1.6.5.	Holonóm kényszerek	20
1.7.	Másodrendű szükséges feltételek	20
1.7.1.	A második variáció nem-negativitása	20
1.7.2.	A Legendre–Hadamard szükséges feltétel	21
1.7.3.	A Weierstrass-féle szükséges feltétel	21
1.8.	Null-Lagrange függvények	22
1.9.	Elégséges feltételek	22
1.9.1.	A második variáció koercitivitása	23
1.9.2.	Jacobi-féle konjugált pontok	23
1.9.3.	Konjugált pontok és a gyenge lokális minimalitás szükséges feltétele	24
1.9.4.	Weierstrass-féle térelmélet (erős minimumhoz)	24
1.10.	Fontosabb Példák	25
1.10.1.	A legrövidebb út	25
1.10.2.	A brachisztokron-probléma	25
1.10.3.	Minimálfelületek	27
1.10.4.	Forgási minimálfelület	27
1.10.5.	Fénysugár útja változó törésmutatójú közegben	28
1.10.6.	Izoperimetrikus egyenlőtlenség	29
1.10.7.	Képhelyreállítás	29
1.10.8.	Képszegmentáció	30
1.10.9.	Variációs autoenkóder (VAE)	31
2.	Tenzor kalkulus	33
2.1.	Vektorterek és bázisok	33
2.2.	Duális vektorterek és duális bázisok	34
2.3.	A Kronecker-tenzor	36
2.4.	Skálárszorzatok	37
2.5.	Reciprok bázisok	40
2.6.	Bázisok, Duális Bázisok, Reciprok Bázisok	41
2.7.	Tenzorok	41
2.7.1.	Tenzorok általánosságban	41
2.7.2.	Szimmetrikus tenzorok	42
2.7.3.	Vektorterek orientált térfogattal	45

2.8. Tenzorok skalárszorzos téren	45
2.8.1. „Pseudotenzorok”	47
2.8.2. Kontrakciók	48
2.9. A Hodge-féle csillag operátor	48
3. Differenciálgeometria	49
3.1. Euklideszi tér: Descartes- és görbevonali koordináták	49
3.2. Differenciálható sokaságok	49
3.3. Érintővektorok	50
3.4. Érintő- és kotangens nyaláb	50
3.5. Külső derivált	51
3.6. Affin konnexió	51
3.7. Lie-derivált	52
3.8. Torzió	52
3.9. Levi-Civita-konnexió	52
3.10. Geodetikusok	53
3.11. Görbület	53
3.12. Előretolás és visszahúzás (Push-Forward and Pull-Back)	54

1. Variációszámítás

1.1. Bevezetés

A variációszámítás a matematika egy olyan területe, amely a funkcionálok (azaz olyan valós értékű függvények, amelyek bemenetei maguk is függvények) minimalizálásával (vagy maximalizálásával) foglalkozik. A variációszámítás széles körben alkalmazható a fizikában, a mérnöki tudományokban, az alkalmazott és elméleti matematikában, és szorosan kapcsolódik a parciális differenciálegyenletekhez (PDE-k). Például egy klasszikus probléma a variációszámításban a két pont közötti legrövidebb út megtalálása. A távolság fogalma nem feltétlenül euklideszi, vagy az út egy felületre korlátozódhat, ebben az esetben a legrövidebb utat geodetikus vonalnak nevezik. A fizikában a Hamilton-elv kimondja, hogy egy fizikai rendszer pályái a hatásfunkcionál kritikus pontjai. A kritikus pontok lehetnek a hatásfunkcionál minimumai, maximumai vagy nyeregpontjai. A gépi látásban egy kép értelmes régiókra való szegmentálásának problémáját gyakran egy funkcionál minimalizálási problémájaként fogalmazzák meg az összes lehetséges szegmentációra – ez egy természetes probléma a variációszámításban. Hasonlóképpen, a képfeldolgozásban a leromlott vagy zajos képek helyreállításának problémáját nagyon sikeresen fogalmazták meg a variációszámítás problémájaként. A PDE-k a funkcionálok minimalizálóihoz vonatkozó szükséges feltételekként kerülnek elő. Emlékezzünk vissza a többváltozós analízisből, hogy ha egy $f : \mathbb{R}^d \rightarrow \mathbb{R}$ függvénynek $\mathbf{x} \in \mathbb{R}^d$ pontban minimuma van, akkor $\nabla f(\mathbf{x}) = 0$. A $\nabla f(\mathbf{x}) = 0$ szükséges feltétel felhasználható a lehetséges minimalizáló \mathbf{x} pontok megoldására. A variációszámításban, ha egy $f : \mathbb{R}^d \rightarrow \mathbb{R}$ függvény egy $I(f)$ funkcionál minimalizálóját, akkor a $\delta I(f) = 0$ szükséges feltétel egy PDE-nek bizonyul, amelyet Euler–Lagrange-egyenletnek neveznek. Az Euler–Lagrange-egyenlet tanulmányozása lehetővé teszi számunkra a minimalizálók explicit kiszámítását és tulajdonságaik vizsgálatát. Emiatt gazdag kölcsönhatás van a variációszámítás és a PDE-k elmélete között. Az ebben a fejezetben szereplő ötletek a Γ -konvergenciához kapcsolódnak, amely a funkcionálok konvergenciájának egy olyan fogalma, amely biztosítja, hogy a minimalizálók minimalizálókhoz konvergáljanak.

1.2. Egydimenziós skalárfüggvények elsőrendű szükséges feltételei: az első variáció — C^1 elmélet

Most a többváltozós analízisből jól ismert szélsőérték módszert a variációs integrálok esetére szeretnénk kiterjeszteni, azaz olyan $I : C^1([a, b]) \rightarrow \mathbb{R}$ funkcionálokra, melyek alakja:

$$I(f) := \int_a^b L(x, f(x), f'(x)) dx$$

azzal a céllal, hogy levezessünk néhány (elsőrendű) szükséges feltételt, melyet a minimalizálóknak ki kell elégíteniük.

1.2.1. Az Euler–Lagrange-egyenlet — erős alak

Néhány jelölés rögzítésével kezdünk.

1.1. Definíció. Az $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ függvényt **Lagrange függvénynek** nevezzük.

Rögzítsünk tehát egy $f \in C^1([a, b])$ függvényt és vegyünk egy $\varphi \in C^1([a, b])$ irányt. Euler ötlete az volt, hogy tekintsük az I funkcionál f_0 -nál vett iránymenti deriváltját a φ irány mentén, azaz tekintsük a $\Phi : (-\varepsilon_0, \varepsilon_0) \rightarrow \mathbb{R}$ függvényt, melynek definíciója:

$$\Phi(\varepsilon) := I(f + \varepsilon\varphi) \quad (1)$$

és deriváljuk azt. Ehhez szükségünk van a következő technikai eredményre.

1.2. Lemma. Legyen $g : [a, b] \times [c, d] \rightarrow \mathbb{R}$ egy folytonos függvény, melynek létezik a második változó szerinti parciális deriváltja, és az folytonos. Defináljuk a $G : [c, d] \rightarrow \mathbb{R}$ függvényt az alábbiak szerint:

$$G(t) = \int_a^b g(x, t) dx$$

Ekkor G egy C^1 osztályú függvény, és

$$G'(t) := \int_a^b \frac{\partial g}{\partial t}(x, t) dx$$

Bizonyítás. Rögzítsünk egy $t \in [c, d]$ értéket, és elegendően kicsi h -ra tekintsük a differenciányadost:

$$\frac{G(t+h) - G(t)}{h} = \int_a^b \frac{g(x, t+h) - g(x, t)}{h} dx = \int_a^b \frac{\partial g}{\partial t}(x, t + \theta_h) dx$$

ahol az utolsó lépésben a Lagrange-középértéktételt használtuk, és $\theta_h \in (0, 1)$ függ x , t és h értékétől. Mivel $\frac{\partial g}{\partial t}$ folytonos, egyenletesen folytonos a $[c, d]$ intervallumon. Így, rögzített $\varepsilon > 0$ -hoz találhatunk olyan $\delta > 0$ értéket, hogy ha $|h| < \delta$, akkor

$$\left| \frac{\partial g}{\partial t}(x, t + \theta_h) - \frac{\partial g}{\partial t}(x, t) \right| < \varepsilon$$

minden $x \in [a, b]$ -re. Ekkor

$$\begin{aligned} \left| \frac{G(t+h) - G(t)}{h} - \int_a^b \frac{\partial g}{\partial t}(x, t) dx \right| &= \left| \int_a^b \left(\frac{\partial g}{\partial t}(x, t + \theta_h) - \frac{\partial g}{\partial t}(x, t) \right) dx \right| \\ &\leq \int_a^b \left| \frac{\partial g}{\partial t}(x, t + \theta_h) - \frac{\partial g}{\partial t}(x, t) \right| dx \leq \varepsilon(b-a) \end{aligned}$$

Mivel ε tetszőleges, a bizonyítás kész.

Tegyük fel tehát, hogy a Lagrange függvény C^1 osztályú. Valójában csak arra van szükségünk, hogy L -nek a p (sebesség) és ξ (érték) változók szerinti parciális deriváltjai folytonosak legyenek. A fenti eredmény azt mondja, hogy a Φ függvény differenciálható $\varepsilon = 0$ -ban, és

$$\Phi'(0) = \int_a^b [L_p(x, f(x), f'(x))\varphi'(x) + L_\xi(x, f(x), f'(x))\varphi(x)] dx$$

1.3. Definíció. Bevezetjük a $\delta I : C^1([a, b]) \times C^1([a, b]) \rightarrow \mathbb{R}$ operátort a következőképpen:

$$\delta I(f; \varphi) := \Phi'(0)$$

ahol Φ a (1) szerint definiált, feltéve, hogy a jobb oldal létezik. A $\delta I(f; \varphi)$ mennyiséget az I funkcionál **első variációjának** nevezzük f -nál a φ irány mentén.

1.4. Megjegyzés. Anélkül, hogy feltételeznénk L C^1 osztályú voltát, nem tudjuk, hogy Φ deriváltja $\varepsilon = 0$ -ban létezik-e, és ha igen, hogyan írható fel.

Most azokra az $f \in C^1([a, b])$ függvényekre összpontosítunk, amelyek az I minimumai valamely $\mathcal{A} \subset C^1([a, b])$ megengedett függvényosztályon. Látni fogjuk, hogy a levont következtetések valóban függenek a megengedett osztály tulajdonságaitól.

A következőkben feltesszük, hogy a megengedett osztály:

$$\mathcal{A} := \{w \in C^1([a, b]) : w(a) = \alpha, w(b) = \beta\}$$

néhány rögzített $\alpha, \beta \in \mathbb{R}$ értékre. Ebben az esetben a megengedett variációk azok, amelyek a peremértékeket rögzítve hagyják. Emiatt csak olyan φ függvényeket veszünk figyelembe, amelyekre:

$$C_0^1([a, b]) := \{w \in C^1([a, b]) : w(a) = w(b) = 0\}$$

Tudjuk, hogy ha az operátor jól definiált, akkor

$$\delta I(f; \varphi) = 0$$

kell, hogy teljesüljön minden $\varphi \in C_0^1([a, b])$ esetén. A 3.2. Lemma alapján ez átfogalmazható:

$$\int_a^b [L_p(x, f(x), f'(x))\varphi'(x) + L_\xi(x, f(x), f'(x))\varphi(x)] dx = 0 \quad (2)$$

minden $\varphi \in C_0^1([a, b])$ esetén. Ezt az egyenletet az I **gyenge Euler–Lagrange-egyenletének** nevezzük.

1.5. Definíció. Egy $f \in C^1([a, b])$ függvényt, amely kielégíti a (2) egyenletet minden $\varphi \in C_0^1([a, b])$ esetén, az I **gyenge extrémálisának** nevezzük.

A cél egy "szebb" egyenletet kapni, amelyet az I minimalizálóinak az \mathcal{A} halmazon ki kell elégíteniük. Emiatt feltesszük, hogy a Lagrange függvény L C^2 osztályú, és a minimumhely f szintén C^2 osztályú. Ezen hipotézisek mellett, parciális integrálással, a gyenge Euler–Lagrange-egyenletet a következőképpen írhatjuk:

$$\begin{aligned} 0 &= \int_a^b L_p(x, f, f')\varphi'(x) dx + \int_a^b L_\xi(x, f, f')\varphi(x) dx \\ &= [L_p(x, f, f')\varphi(x)]_a^b - \int_a^b \frac{d}{dx} L_p(x, f, f')\varphi(x) dx + \int_a^b L_\xi(x, f, f')\varphi(x) dx \\ &= \int_a^b \left(L_\xi(x, f, f') - \frac{d}{dx} L_p(x, f, f') \right) \varphi(x) dx \end{aligned}$$

ahol az utolsó lépésben felhasználtuk, hogy $\varphi(a) = \varphi(b) = 0$.

Szükségünk van egy további lépésre, hogy a fenti feltételből egy szép egyenletet kapjunk. A következő technikai eredmény megmutatja, hogyan.

1.6. Lemma. (A variációszámítás fundamentális lemmája) Tegyük fel, hogy van egy $g \in C^0([a, b])$ függvényünk, amelyre

$$\int_a^b g(x)\varphi(x) dx = 0$$

teljesül minden $\varphi \in C_0^1([a, b])$ esetén. Ekkor $g \equiv 0$ az $[a, b]$ intervallumon.

Bizonyítás. Tegyük fel indirekt, hogy létezik egy $x_0 \in (a, b)$ pont, ahol $g(x_0) \neq 0$. Az általánosság megszorítása nélkül feltehetjük, hogy $g(x_0) > 0$. Mivel g folytonos, létezik egy $\delta > 0$, hogy $g(x) > \frac{g(x_0)}{2} > 0$ minden $x \in (x_0 - \delta, x_0 + \delta) \subset [a, b]$ esetén. Az ötlet az, hogy konstruálunk egy $\varphi \in C_0^1([a, b])$ függvényt, amely pozitív az $(x_0 - \delta, x_0 + \delta)$ intervallumon, és nulla máshol. Egy pillanatra tételezzük fel egy ilyen φ létezését. Akkor azt kapnánk, hogy

$$0 = \int_a^b g(x)\varphi(x) dx = \int_{x_0-\delta}^{x_0+\delta} g(x)\varphi(x) dx > \frac{g(x_0)}{2} \int_{x_0-\delta}^{x_0+\delta} \varphi(x) dx > 0$$

Mivel ez ellentmondás, arra következtetünk, hogy $g \equiv 0$ az $[a, b]$ intervallumon. Most konstruáljuk meg a megfelelő φ függvényt. Legyen:

$$\varphi(x) := \begin{cases} (x - (x_0 - \delta))^2(x - (x_0 + \delta))^2 & \text{ha } x \in (x_0 - \delta, x_0 + \delta) \\ 0 & \text{egyébként} \end{cases}$$

Könnyen belátható, hogy ez a függvény rendelkezik a kívánt tulajdonságokkal.

Így a következő eredményt kaptuk:

1.7. Tétel. (Euler–Lagrange-egyenlet — erős alak) Legyen $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ egy C^2 osztályú függvény. Tegyük fel, hogy az $I : C^1([a, b]) \rightarrow \mathbb{R}$ funkcionál,

$$I(u) := \int_a^b L(x, f(x), f'(x)) dx$$

egy $f \in C^2([a, b]) \cap \mathcal{A}$ minimummal rendelkezik. Ekkor a következő egyenlet teljesül:

$$L_\xi(x, f(x), f'(x)) = \frac{d}{dx} L_p(x, f(x), f'(x)) \quad (3)$$

minden $x \in [a, b]$ esetén.

1.8. Definíció. A (3) egyenletet az I (**erős**) Euler–Lagrange-egyenletének nevezzük. Egy $f \in C^2([a, b])$ függvényt, amely kielégíti ezt az egyenletet, az I (**erős**) **extremálisának** nevezzük.

1.9. Megjegyzés. A fenti tétel nem állít semmiféle létezési eredményt. Ez csupán egy szükséges feltétel, amelyet az I minimalizálóinak az \mathcal{A} halmazon ki kell elégíteniük. Továbbá az a tény, hogy egy f minimum C^2 osztályú, olyasmí, amit a priori feltételezünk, és általában nem garantált.

1.10. Megjegyzés. Megjegyezzük, hogy mivel L és f feltételezhetően C^2 osztályúak, a (3) jobb oldalát a következőképpen írhatjuk:

$$L_{xp}(x, f, f') + L_{\xi p}(x, f, f')f'(x) + L_{pp}(x, f, f')f''(x)$$

1.2.2. Kőzjáték a tesztfüggvényekről és a lokális minimalitásról

Mielőtt folytatnánk, szeretnénk néhány szót ejteni két fontos kérdéstről: a megengedett variációk terének megválasztásáról és a lokális minimalizálók fogalmáról.

Eddig a $C_0^1([a, b])$ teret használtuk a megengedett variációk, vagyis a tesztfüggvények terének az \mathcal{A} osztályon való minimalizálók számára. Ezt a teret ad hoc választottuk ki az adott helyzetre. Tegyük fel, hogy egy hasonló szükséges feltételt szeretnénk levezetni az

$$I(f) := \int_a^b L(x, f(x), f'(x), f''(x)) dx$$

típusú $F : C^2([a, b]) \rightarrow \mathbb{R}$ funkcionálok minimalizálóiira a $\mathcal{B} := \{w \in C^2([a, b]) : w(a) = \alpha_0, w'(a) = \alpha_1, w(b) = \beta_0, w'(b) = \beta_1\}$ osztályon. Ebben az esetben a tesztfüggvények tere a $C_0^2([a, b])$ lesz. Hasonlóképpen, ha a Lagrange függvény az f k -adik deriváltjától függ, akkor a tesztfüggvények tere $C_0^k([a, b])$ lesz, és így tovább. Ezért szokás a tesztfüggvények standard terének a

$$C_c^\infty([a, b]) := \{w \in C^\infty([a, b]) : \text{supp}(w) \subseteq (a, b)\}$$

teret venni, azaz azon C^∞ függvények terét, amelyek tartója (azaz annak a halmaznak a lezártja, ahol a függvény nem nulla) kompaktan tartalmazott az (a, b) intervallumban. Ennek a térnek a választását indokolni kell. Először is, vegyük észre, hogy $C_c^\infty([a, b]) \subset C_0^k([a, b])$ minden $k \in \mathbb{N}$ esetén. Így ez a tér bármilyen rendű deriváltaktól függő Lagrange függvényekhez használható. Továbbá kiderül, hogy a variációszámítás fundamentális lemmája még ebben a kisebb tesztfüggvényterben is érvényes. Vagyis fennáll a következő:

1.11. Lemma. (Fundamentális lemma — második változat) Tegyük fel, hogy van egy $g \in C^0([a, b])$ függvényünk, amelyre

$$\int_a^b g(x)\varphi(x) dx = 0$$

minden $\varphi \in C_c^\infty([a, b])$ esetén. Ekkor $g \equiv 0$ az $[a, b]$ intervallumon.

1.12. Megjegyzés. Valójában egy általánosabb állítás is igaz: vehetjük, hogy $g \in L^1(a, b)$, és ugyanarra a következtetésre jutunk!

A bizonyítás ötlete ugyanaz, mint a 3.6. Lemmáé. Csak a megfelelő, $C_c^\infty([a, b])$ -beli tesztfüggvényt kell megkonstruálni, ugyanazokkal a tulajdonságokkal, mint amit a 3.6. Lemmában konstruáltunk. Ez azt jelenti, hogy még ha csak azt tudjuk is, hogy az iránymenti derivált ezen kisebb tér irányjaiban nulla, ez elegendő ahhoz, hogy megkapjuk a $\delta I(f; \varphi) = 0$ szükséges feltétel differenciális alakját.

Most a lokális minimalizálók fogalmát tárgyaljuk. Mint tudjuk, végtelen dimenzióban nem minden norma ekvivalens. Ez azt jelenti, hogy a lokális fogalma attól a normától függ, amelyet

a terünkben választunk. Koncentráljunk most a $C^1([a, b])$ térre. A hozzá tartozó természetes norma az úgynevezett C^1 -norma, $\|\cdot\|_{C^1}$, amelyet a következőképpen adunk meg:

$$\|f\|_{C^1} := \max_{x \in [a, b]} |f(x)| + \max_{x \in [a, b]} |f'(x)| =: \|f\|_{C^0} + \|f'\|_{C^0}$$

De ezen a téren a C^0 -normát is tekinthetjük (alapvetően nem törődünk a deriválttal!):

$$\|f\|_{C^0} := \max_{x \in [a, b]} |f(x)|$$

Nyilvánvalóan $\|f\|_{C^0} \leq \|f\|_{C^1}$, de a két norma nem ekvivalens, amint azt az $f_n(x) := \frac{1}{n} \sin(nx)$ függvénysorozat mutatja. Tehát a $\|\cdot\|_{C^1}$ norma erősebb, mint a $\|\cdot\|_{C^0}$ norma. Így a lokális minimalitásnak két fogalma van:

1.13. Definíció. Egy $f \in \mathcal{A}$ függvényt, amelyre

$$I(f) \leq I(v)$$

teljesül minden $v \in \mathcal{A}$ esetén, amelyre $\|f - v\|_{C^1} < \varepsilon$ valamely $\varepsilon > 0$ -ra, az I **gyenge lokális minimalizálójának** nevezzük. Ha az egyenlőség csak akkor áll fenn, ha $v = f$, akkor azt mondjuk, hogy f az I **szigorú gyenge lokális minimalizálója**.

1.14. Definíció. Egy $f \in \mathcal{A}$ függvényt, amelyre

$$I(f) \leq I(v)$$

teljesül minden $v \in \mathcal{A}$ esetén, amelyre $\|f - v\|_{C^0} < \varepsilon$ valamely $\varepsilon > 0$ -ra, az I **erős lokális minimalizálójának** nevezzük. Ha az egyenlőség csak akkor áll fenn, ha $v = f$, akkor azt mondjuk, hogy f az I **szigorú erős lokális minimalizálója**.

Nyilvánvaló, hogy egy erős lokális minimalizáló egyben gyenge lokális minimalizáló is. Az ellenkezője nem igaz, amint azt a következő példa mutatja.

1.15. Példa. (Bolza) Tekintsük a funkcionált

$$I(f) := \int_0^1 ((f'(x))^2 - (f'(x))^4) dx$$

a $C_0^1([0, 1])$ halmazon definiálva. Bizonyítsuk be, hogy $f \equiv 0$ egy szigorú gyenge lokális minimalizáló, de nem erős. Az ötlet a következő: a $g(p) := p^2(1 - p^2)$ Lagrange függvénynek izolált lokális minimuma van $p = 0$ -ban. Ezért $f \equiv 0$ az I szigorú gyenge lokális minimalizálója. Másrészt lehetséges olyan $(f_n)_n$ függvénysorozatot konstruálni, hogy $f_n \rightarrow 0$ egyenletesen a $[0, 1]$ intervallumon, és $I(f_n) \rightarrow -\infty$. Ehhez az ötlet az, hogy az f_n -ek deriváltját "felrobbantjuk". Ez bizonyítja, hogy $f \equiv 0$ nem erős lokális minimalizálója az I -nek.

Tehát minden alkalommal, amikor lokális minimalizálókat vizsgálunk, meg kell határoznunk, hogy melyik metrikát (vagy topológiát) vesszük figyelembe.

1.16. Megjegyzés. Nyilvánvaló, hogy a 3.7. Tétel érvényes az I gyenge lokális minimalizálóira is, és így az I erős lokális minimalizálóira is.

1.2.3. A Du Bois-Reymond-egyenlet — erős alak

Most egy másik elsőrendű szükséges feltételt szeretnénk levezetni az I lokális minimalizálóra. Egyelőre az eredményt anélkül közöljük, hogy megmagyaráznánk a levezetés mögött rejlő ötletet.

1.17. Tétel. (A Du Bois-Reymond-egyenlet — erős alak) Tekintsünk egy $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ C^2 osztályú függvényt. Tegyük fel, hogy az $I : C^1([a, b]) \rightarrow \mathbb{R}$ funkcionál

$$I(f) := \int_a^b L(x, f(x), f'(x)) dx$$

egy $f \in C^2([a, b]) \cap \mathcal{A}$ gyenge lokális minimummal rendelkezik. Ekkor a következő egyenlet teljesül:

$$\frac{d}{dx} (L(x, f, f') - f'(x)L_p(x, f, f')) = L_x(x, f, f') \quad (4)$$

minden $x \in [a, b]$ esetén.

Bizonyítás. Közvetlen számítással kapjuk, hogy

$$\begin{aligned} \frac{d}{dx} (L - f' L_p) &= L_x + L_\xi d' + L_p f'' - f'' L_p - f' \frac{d}{dx} L_p \\ &= L_x + f' \left(L_\xi - \frac{d}{dx} L_p \right) \end{aligned}$$

Mivel f kielégíti az Euler–Lagrange-egyenletet, a zárójelben lévő kifejezés nulla, ami bizonyítja az állítást.

1.18. Megjegyzés. A (4) egyenletet az Euler–Lagrange-egyenlet második alakjának is nevezzük.

1.2.4. A minimalizálók létezéséről

Az összes eddigi szükséges feltételben adottnak vettük egy minimalizáló létezését. Itt meg akarjuk mutatni, hogy még egy nagyon egyszerű esetben is előfordulhat, hogy a minimalizálók létezése nem teljesül.

A bemutatandó példa az úgynevezett Euler-paradoxon. Tekintsük az $L(p) := (1-p^2)^2$ függvényt. Definiáljuk a funkcionált:

$$I(f) := \int_0^1 L(f'(x)) dx$$

minden $f \in C_0^1([0, 1])$ esetén. Ekkor $I(f) \geq 0$, de nincs olyan függvény, amelyre $I(f) = 0$ lenne. Valóban, egy ilyen f függvénynek csak $f' = \pm 1$ lehetne, és ki kellene elégítenie az $f(0) = f(1) = 0$ feltételt. De ez nem egyeztethető össze az $f \in C^1$ követelménnyel. Ha a megengedett függvények osztályát kiterjesztjük a szakaszonként C^1 függvényekre, akkor létezik minimalizáló.

1.2.5. Egy megmaradási törvény: Beltrami-azonosság

Tegyük fel, hogy a Lagrange függvény expliciten nem függ az x változótól, azaz $L = L(\xi, p)$. Ebben az esetben a Du Bois-Reymond-egyenlet (4) a következőre egyszerűsödik:

$$\frac{d}{dx} (L(f, f') - f' L_p(f, f')) = 0$$

ami azt jelenti, hogy létezik egy $c \in \mathbb{R}$ konstans, amelyre

$$L(f(x), f'(x)) - f'(x) L_p(f(x), f'(x)) = c$$

minden $x \in [a, b]$ esetén. Ezt az egyenletet **Beltrami-azonosságnak** vagy az energia-megmaradás törvényének nevezik.

1.2.6. Egy regularitási eredmény

Most tegyük fel magunknak a következő kérdést: tegyük fel, hogy van egy (lokális) $f \in C^1([a, b])$ minimalizálója az I -nek. Lehetséges-e a gyenge Euler–Lagrange-egyenletből, anélkül, hogy expliciten megoldanánk, arra következtetni, hogy f valójában simább? A következő eredmény választ ad erre.

1.19. Tétel. Legyen $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ egy C^2 osztályú Lagrange függvény. Legyen $f \in C^1([a, b])$ a gyenge Euler–Lagrange-egyenlet egy megoldása. Tegyük fel, hogy

$$L_{pp}(x, f(x), f'(x)) \neq 0$$

minden $x \in [a, b]$ esetén. Ekkor $f \in C^2([a, b])$.

Bizonyítás. Tudjuk, hogy létezik $c \in \mathbb{R}$ konstans, amelyre a következő egyenlet teljesül:

$$L_p(x, f(x), f'(x)) = g(x) := c + \int_a^x L_\xi(t, f(t), f'(t)) dt$$

minden $x \in [a, b]$ esetén. A g függvény C^1 osztályú. Definiáljuk a $G(x, p) := L_p(x, f(x), p) - g(x)$ függvényt, amely C^1 osztályú. Tudjuk, hogy $G(x, f'(x)) = 0$ minden $x \in [a, b]$ esetén. Továbbá

$$\frac{\partial G}{\partial p}(x, f'(x)) = L_{pp}(x, f(x), f'(x)) \neq 0$$

minden $x \in [a, b]$ esetén. Az implicitfüggvény-tétel alkalmazásával azt kapjuk, hogy $f'(x) = h(x)$ valamely C^1 osztályú h függvényre. Tehát $f \in C^2([a, b])$.

1.20. Megjegyzés. Lehetséges a fenti tételt a következőképpen kiterjeszteni: tegyük fel, hogy a Lagrange függvény L C^k osztályú és továbbra is kielégíti a nem-elfajulási feltételt. Ekkor a gyenge Euler–Lagrange-egyenlet bármely C^1 megoldása valójában C^k osztályú.

A következő példa mutatja, hogy a nem-elfajulási feltétel valóban szükséges egy ilyen regularitási eredmény eléréséhez.

1.21. Példa. Legyen $L \in C^2(\mathbb{R})$ egy konvex függvény, amelyre $L(p) = |p|$ a $[-1, 1]$ intervallumon. Ekkor az $\int_0^1 L(f'(x)) dx$ funkcionál minimalizálói az $f(0) = 0, f(1) = 1$ feltételek mellett szingularitásokat mutathatnak.

1.3. Egydimenziós skalárfüggvények elsőrendű szükséges feltételei: az első variáció — C^1 elmélet

Az előző fejezetben két fontos elsőrendű szükséges feltételt (az Euler–Lagrange-egyenletet és a Du Bois-Reymond-egyenletet) vezettünk le, feltételezve, hogy a minimalizáló C^2 osztályú. A következő példa mutatja, hogy általában ez egy olyan feltételezés, amelyet nem tehetünk meg a priori.

1.22. Példa. Tekintsük a következő funkcionált:

$$I(f) := \int_{-1}^1 (f'(x))^2 (2x - f'(x))^2 dx$$

és vizsgáljuk a minimalizálás problémáját a $v \in C^1([-1, 1])$ függvények körében, amelyekre $v(-1) = 0$ és $v(1) = 1$. Könnyen belátható, hogy a funkcionált egyértelműen minimalizálja a következő függvény:

$$f(x) := \begin{cases} 0 & x \in [-1, 0] \\ x^2 & x \in [0, 1] \end{cases}$$

Ez a függvény $C^1([-1, 1])$, de nem eleme a $C^2([-1, 1])$ térnek.

Tehát hasznos lenne olyan elsőrendű szükséges feltételeket találni, amelyek igazak a csak $C^1([a, b])$ osztályba tartozó (lokális) minimalizálókra is.

1.3.1. Az Euler–Lagrange-egyenlet

Visszatekintve arra, amit az Euler–Lagrange-egyenlet levezetése során tettünk, észrevevesszük, hogy az alapvető lépés, ahol az f minimalizáló (és az L Lagrange függvény) további regularitása igazán számít, az a parciális integrálás. Valóban, csupán feltételezve, hogy L C^1 osztályú és a minimalizáló f C^1 osztályú, tudjuk, hogy:

$$\int_a^b [L_p(x, f, f') \varphi'(x) + L_\xi(x, f, f') \varphi(x)] dx = 0 \quad (5)$$

minden $\varphi \in C_c^\infty((a, b))$ esetén. Ebből a feltételből szeretnénk egy differenciálegyenletet kapni.

Tehát, tegyük ezt lépésről lépésre: tegyük fel, hogy van két folytonos $g, h : [a, b] \rightarrow \mathbb{R}$ függvényünk, amelyekre a következő igaz:

$$\int_a^b [g(x) \varphi'(x) + h(x) \varphi(x)] dx = 0 \quad (6)$$

minden $\varphi \in C_c^\infty((a, b))$ esetén. Szeretnénk valamilyen kapcsolatot levezetni g és h között. Mivel nem tudunk parciálisan integrálni, az első tagot kell kezelni. A technikai eredmény, ami segít nekünk, a következő:

1.23. Lemma. (Du Bois-Reymond-lemma) Legyen $g : [a, b] \rightarrow \mathbb{R}$ egy folytonos függvény, amelyre

$$\int_a^b g(x) \psi'(x) dx = 0$$

minden $\psi \in C_c^\infty((a, b))$ esetén. Ekkor létezik egy $c \in \mathbb{R}$ konstans, úgy, hogy $g(x) = c$ az $[a, b]$ intervallumon.

1.24. Megjegyzés. Vegyük észre, hogy ha $\psi \in C_c^\infty((a, b))$, akkor $\psi' \in C_c^\infty((a, b))$. Így a fenti eredmény azt mondja, hogy ha $\int_a^b g(x) v(x) dx = 0$ csak olyan függvényekre igaz, amelyek deriváltak, akkor arra következtethetünk, hogy g konstans, de nem következtethetünk arra, hogy a konstans 0, ahogy azt a variációszámítás fundamentális lemmájában tehetjük.

Bizonyítás. (A 3.20. Lemma bizonyítása) Először is szeretnénk megérteni (jellemezni) a $C_c^\infty((a, b))$ azon részhalmazát, amely a $C_c^\infty((a, b))$ függvényeinek deriváltjaiból áll. Legyen $v \in C_c^\infty((a, b))$; ekkor $v = \psi'$ valamely $\psi \in C_c^\infty((a, b))$ -re, és

$$\int_a^b v(x) dx = \psi(b) - \psi(a) = 0 - 0 = 0$$

Azt állítjuk, hogy ez a tulajdonság jellemzi a deriváltakat. Pontosabban, ha van egy $v \in C_c^\infty((a, b))$ függvényünk, amelyre

$$\int_a^b v(x) dx = 0,$$

akkor létezik $\psi \in C_c^\infty((a, b))$ úgy, hogy $v = \psi'$. Valóban, a $\psi(x) := \int_a^x v(t) dt$ definícióval $\psi \in C_c^\infty((a, b))$ és $\psi' = v$.

Az ötlet most az, hogy a variációszámítás fundamentális lemmáját használjuk fel ennek az eredménynek a bizonyítására. Vegyünk egy $\varphi \in C_c^\infty((a, b))$ függvényt. Annak érdekében, hogy ezt a függvényt a problémánkban tesztfüggvényként használhassuk, át kell alakítanunk egy deriválttá, azaz át kell alakítanunk egy olyan függvénné, amelynek az $[a, b]$ feletti integrálja nulla. A legegyszerűbb módja ennek a következő függvény megfontolása:

$$\tilde{\varphi}(x) := \varphi(x) - \eta(x) \int_a^b \varphi(t) dt$$

ahol $\eta \in C_0^\infty([a, b])$ egy olyan rögzített függvény, amelyre $\int_a^b \eta(t) dt = 1$. Most ellenőrizzük, hogy jól jártunk-e el: $\tilde{\varphi} \in C_c^\infty((a, b))$ és $\int_a^b \tilde{\varphi}(x) dx = 0$. Tehát ezt a függvényt használhatjuk tesztfüggvényként a problémánkban. Ekkor

$$\begin{aligned} 0 &= \int_a^b g(x) \tilde{\varphi}(x) dx \\ &= \int_a^b g(x) \left(\varphi(x) - \eta(x) \int_a^b \varphi(t) dt \right) dx \\ &= \int_a^b g(x) \varphi(x) dx - \left(\int_a^b \varphi(t) dt \right) \left(\int_a^b g(x) \eta(x) dx \right) \\ &= \int_a^b \left(g(x) - \int_a^b g(t) \eta(t) dt \right) \varphi(x) dx \end{aligned}$$

Mivel ez minden $\varphi \in C_c^\infty((a, b))$ esetén igaz, a variációszámítás fundamentális lemmáját használva arra következtetünk, hogy

$$g(x) - \int_a^b g(t)\eta(t) dt = 0.$$

Ez azt mondja, hogy g egy konstans.

Az előző eredmény segít nekünk a (6) kifejezés kezelésében.

1.25. Következmény. Legyenek $g, h : [a, b] \rightarrow \mathbb{R}$ folytonos függvények. Tegyük fel, hogy

$$\int_a^b [g(x)\varphi'(x) + h(x)\varphi(x)] dx = 0$$

minden $\varphi \in C_c^\infty((a, b))$ esetén. Ekkor a g függvény $C^1([a, b])$ osztályú és $g'(x) = h(x)$.

A fenti eredményt a mi esetünkre alkalmazva a következő szükséges feltételt kapjuk az f és u természetes feltételezései mellett.

1.26. Tétel. Legyen $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ egy C^1 osztályú függvény. Tegyük fel, hogy az $I : C^1([a, b]) \rightarrow \mathbb{R}$ funkcionál

$$I(f) := \int_a^b L(x, f(x), f'(x)) dx$$

egy $f \in C^1([a, b]) \cap \mathcal{A}$ minimummal rendelkezik. Ekkor a $x \mapsto L_p(x, f(x), f'(x))$ függvény C^1 osztályú, és létezik egy konstans, úgy, hogy a következő egyenlet teljesül:

$$L_p(x, f(x), f'(x)) = c + \int_a^x L_\xi(t, f(t), f'(t)) dt$$

minden $x \in [a, b]$ esetén.

1.27. Megjegyzés. Általában a fenti egyenletet a következő formában írják:

$$\frac{d}{dx} L_p(x, f(x), f'(x)) = L_\xi(x, f(x), f'(x))$$

Mindazonáltal mi inkább az integrális formában írjuk, hogy emlékeztessük magunkat arra, hogy a bal oldalt (általában) nem lehet a láncszabállyal kifejtetni, mivel csak azt tételezzük fel, hogy L és f C^1 osztályúak.

1.3.2. A Du Bois-Reymond-egyenlet

Mivel szerencsénk volt, és sikerült visszanyerni az Euler–Lagrange-egyenletet (gyengébb formában!) csupán az L és f természetes hipotéziseinek feltételezésével, most azt szeretnénk megérteni, hogy vajon a Du Bois-Reymond-egyenletet is visszanyerhetjük-e ugyanezen gyenge feltételezések mellett.

A tétel bizonyításának ötlete a következő: eddig egy f függvény olyan variációit vizsgáltuk,

amelyek külső variációknak tekinthetők. De mivel függvényekkel dolgozunk, kihasználhatjuk azt a tényt is, hogy a független változót is variálhatjuk.

(A szöveg itt egy komplex levezetést mutat be a független változó variálásával, diffeomorfizmusok segítségével, ami végül a következő tételhez vezet.)

1.28. Tétel. Legyen $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ egy C^1 osztályú függvény. Tegyük fel, hogy az $I : C^1([a, b]) \rightarrow \mathbb{R}$ funkcionál

$$I(f) := \int_a^b L(x, f(x), f'(x)) dx$$

egy $f \in C^1([a, b]) \cap \mathcal{A}$ minimummal rendelkezik. Ekkor létezik egy $c \in \mathbb{R}$ konstans, úgy, hogy a következő egyenlet teljesül:

$$L(x, f, f') - f' L_p(x, f, f') = c + \int_a^x L_x(t, f(t), f'(t)) dt \quad (7)$$

minden $x \in [a, b]$ esetén.

1.29. Megjegyzés. Vegyük észre, hogy az Euler–Lagrange-egyenlet és a Du Bois-Reymond-egyenlet általában különböző egyenletek, mivel az L különböző deriváltjai szerepelnek bennük.

1.3.3. Egy regularitási eredmény

Most tegyük fel magunknak a következő kérdést: tegyük fel, hogy van egy (lokális) $f \in C^1([a, b])$ minimalizálója I -nek. Lehetséges-e a gyenge Euler–Lagrange-egyenletből, anélkül, hogy expliciten megoldanánk, arra következtetni, hogy f valójában simább? A következő eredmény ad választ.

1.30. Tétel. Legyen $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ egy C^2 osztályú Lagrange függvény. Legyen $f \in C^1([a, b])$ a gyenge Euler–Lagrange-egyenlet egy megoldása. Tegyük fel, hogy

$$L_{pp}(x, f(x), f'(x)) \neq 0$$

minden $x \in [a, b]$ esetén. Ekkor $f \in C^2([a, b])$.

Bizonyítás. Tudjuk, hogy létezik $c \in \mathbb{R}$ úgy, hogy a következő egyenlet teljesül:

$$L_p(x, f(x), f'(x)) = g(x) := c + \int_a^x L_\xi(t, f(t), f'(t)) dt$$

minden $x \in [a, b]$ esetén. A g függvény C^1 osztályú. Definiáljuk a $G(x, p) := L_p(x, f(x), p) - g(x)$ függvényt, amely C^1 osztályú. Tudjuk, hogy $G(x, u'(x)) = 0$. Továbbá

$$\frac{\partial G}{\partial p}(x, f'(x)) = L_{pp}(x, f(x), f'(x)) \neq 0$$

minden $x \in [a, b]$ esetén. Az implicitfüggvény-tétel alkalmazásával azt kapjuk, hogy $f'(x) = h(x)$ valamely C^1 osztályú h függvényre. Tehát $f \in C^2([a, b])$.

1.31. Megjegyzés. Lehetséges a fenti tételt a következőképpen kiterjeszteni: tegyük fel, hogy az L Lagrange függvény C^k osztályú, és továbbra is kielégíti a nem-elfajulási feltételt. Ekkor a gyenge Euler–Lagrange-egyenlet bármely C^1 megoldása valójában C^k osztályú lesz.

A következő példa mutatja, hogy a nem-elfajulási feltétel valóban szükséges egy ilyen regularitási eredmény eléréséhez.

1.32. Példa. Legyen $L \in C^2(\mathbb{R})$ egy konvex függvény, amelyre $L(p) = |p|$ a $[-1, 1]$ intervallumon, és L injektív $\mathbb{R} \setminus [-1, 1]$ -en. Ekkor az $\int_0^1 L(f'(x)) dx$ funkcionál minimalizálói az $f(0) = 0, f(1) = 1$ feltételekkel szingularitásokat mutathatnak.

1.33. Megjegyzés. A feltétel, amit az előző tételben megköveteltünk, a $p \mapsto f(x, f(x), p)$ függvény konvexitásával kapcsolatos a $p = f'(x)$ pontban.

1.4. Problémák szabad végpontokkal

Ebben a szakaszban azokat az eseteket vizsgáljuk, amikor a funkcionált a $C^1([a, b])$ osztályon minimalizáljuk, azaz nincsenek peremfeltételek a végpontokra. Legyen $f \in C^1([a, b])$ egy lokális minimalizáló. Vegyünk egy tetszőleges $\varphi \in C^\infty([a, b])$ függvényt. Mivel $f + \varepsilon\varphi$ is a megengedett halmazba tartozik, az első variációnak el kell tűnnie:

$$0 = \int_a^b [L_p(x, f, f')\varphi' + L_\xi(x, f, f')\varphi] dx.$$

Tegyük fel egyelőre, hogy L és f C^2 osztályúak, így parciálisan integrálhatunk:

$$0 = \int_a^b \left(L_\xi - \frac{d}{dx} L_p \right) \varphi(x) dx + [L_p(x, f, f')\varphi(x)]_a^b.$$

Ez az egyenlőség minden $\varphi \in C^\infty([a, b])$ esetén fennáll. Mivel a kompakt tartójú függvényekre $(C_c^\infty((a, b)))$ is igaz, a korábbiakból tudjuk, hogy az integrál alatti kifejezésnek nullának kell lennie, tehát az Euler–Lagrange-egyenlet továbbra is érvényes:

$$L_\xi(x, f, f') = \frac{d}{dx} L_p(x, f, f').$$

Így a peremtagoknak is el kell tűnniük:

$$L_p(b, f(b), f'(b))\varphi(b) - L_p(a, f(a), f'(a))\varphi(a) = 0$$

minden $\varphi \in C^\infty([a, b])$ esetén. Mivel $\varphi(a)$ és $\varphi(b)$ értékeit tetszőlegesen, egymástól függetlenül megválaszthatjuk, a következő lemma adja a következtetést.

1.34. Lemma. Ha $g(b)\varphi(b) - g(a)\varphi(a) = 0$ minden $\varphi \in C^\infty([a, b])$ esetén, akkor $g(a) = 0$ és $g(b) = 0$.

1.35. Tétel. Legyen L egy C^2 osztályú Lagrange függvény, és $f \in C^2([a, b])$ a funkcionál lokális minimalizálója a $C^1([a, b])$ osztályon. Ekkor f kielégíti az Euler–Lagrange-

egyenletet az $[a, b]$ intervallumon, valamint a következő **természetes peremfeltételeket**:

$$L_p(a, f(a), f'(a)) = 0 \quad \text{és} \quad L_p(b, f(b), f'(b)) = 0. \quad (8)$$

1.36. Megjegyzés. Ha csak az egyik végpont szabad (pl. $f(a) = \alpha$ rögzített, de $f(b)$ szabad), akkor a természetes peremfeltétel csak a szabad végpontban, azaz b -ben érvényesül.

1.5. Izoperimetrikus problémák

Most olyan minimalizálási problémákat vizsgálunk, amelyek egy integrális kényszert is tartalmaznak: minimalizáljuk

$$I(f) := \int_a^b L(x, f, f') dx$$

funkcionált az

$$\mathcal{A} := \{f \in C^1([a, b]) : f(a) = \alpha, f(b) = \beta, \text{ és } G(f) = c\}$$

osztályon, ahol $c \in \mathbb{R}$ és

$$G(f) := \int_a^b g(x, f, f') dx.$$

1.5.1. A Lagrange-multiplikátor szabály variációs problémákra

Az ötlet hasonló a véges dimenziós Lagrange-multiplikátoros módszerhez. Ott egy $I(\mathbf{x})$ függvényt minimalizálunk a $G(\mathbf{x}) = 0$ feltétel mellett, és a szükséges feltétel az, hogy $\nabla I(\mathbf{x}) = \lambda \nabla G(\mathbf{x})$ valamely $\lambda \in \mathbb{R}$ konstansra.

1.37. Tétel. (Lagrange-multiplikátor szabály) Legyenek L és g C^1 osztályú függvények. Legyen $f \in \mathcal{A}$ az I funkcionál lokális minimumhelye a $G(f) = c$ kényszer mellett. Tegyük fel, hogy f nem "degenerált" pontja a kényszernek (azaz létezik ψ_0 , amire $\delta G(f)[\psi_0] \neq 0$). Ekkor létezik egy $\lambda \in \mathbb{R}$ konstans (a Lagrange-multiplikátor), amelyre f kielégíti a

$$\delta I(f)[\varphi] + \lambda \delta G(f)[\varphi] = 0$$

egyenletet minden megengedett φ variációra. Ez ekvivalens azzal, hogy f kielégíti a $H = I + \lambda G$ funkcionálhoz tartozó Euler–Lagrange-egyenletet, melynek a Lagrange függvénye $h = L + \lambda g$.

Bizonyítás. (A bizonyítás vázlata) Vegyünk egy φ variációt, amely általában nem tartja a $G(f) = c$ kényszert. Az ötlet az, hogy ezt a variációt egy másik, ψ_0 irányú (amelyre $\delta G(f)[\psi_0] \neq 0$) kis perturbációval korrigáljuk úgy, hogy az új $f + \varepsilon \varphi + s(\varepsilon) \psi_0$ függvény már eleget tegyen a kényszernek. Az implicitfüggvény-tétel garantálja egy ilyen $s(\varepsilon)$ korrekciós függvény létezését. Mivel a korrigált függvényen az I funkcionálnak minimuma van $\varepsilon = 0$ -ban, a deriválnak el kell tűnnie, ami a tétel állításához vezet.

1.38. Megjegyzés. A módszer a gyakorlatban úgy működik, hogy megoldjuk a $h = L + \lambda g$ Lagrange függvényhez tartozó Euler–Lagrange-egyenletet. A megoldás, $f(x; \lambda, c_1, c_2)$, általában függ a λ multiplikátortól és két integrálási állandótól. Ezt a három konstans a

peremfeltételekből és a $G(f) = c$ kényszer-egyenletből határozzuk meg.

1.6. Elsőrendű szükséges feltételek általános függvényekre

Ebben a fejezetben az elsőrendű szükséges feltételekhez vezető gondolatokat általánosítjuk $u : \Omega \rightarrow \mathbb{R}^M$ függvények esetére, ahol $\Omega \subset \mathbb{R}^N$ egy C^1 osztályú peremmel rendelkező nyílt halmaz. Az ilyen függvényekhez

$$L : \Omega \times \mathbb{R}^M \times \mathbb{R}^{M \times N} \rightarrow \mathbb{R}$$

típusú Lagrange függvényeket kell tekintenünk. A változókat a következőképpen jelöljük: $(x, \xi, p) = (x_1, \dots, x_N; \xi^1, \dots, \xi^M; p_1^1, \dots, p_N^M)$. A deriváltakra a $p_\alpha^i = \frac{\partial f^i}{\partial x_\alpha}$ jelölést használjuk. Az egyszerűség kedvéért az ismétlődő indexekre az Einstein-féle szummázási konvenciót alkalmazzuk (görög indexek 1-től N-ig, latin indexek 1-től M-ig futnak).

1.6.1. Az Euler–Lagrange-egyenlet

Az egyváltozós esethez hasonlóan, egy f gyenge lokális minimalizálót egy $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ teszt-függvénnyel perturbálunk. Az $I(f + \varepsilon\varphi)$ funkcionál $\varepsilon = 0$ pontbeli első deriváltjának el kell tűnnie:

$$0 = \delta I(f)[\varphi] = \int_{\Omega} \left(\frac{\partial L}{\partial \xi^i} \varphi^i + \frac{\partial L}{\partial p_\alpha^i} \frac{\partial \varphi^i}{\partial x_\alpha} \right) dx.$$

A második tagra parciális integrálást (Gauss-Ostrogradszki tételt) alkalmazva, és kihasználva, hogy φ a peremen nulla, kapjuk:

$$0 = \int_{\Omega} \left(\frac{\partial L}{\partial \xi^i} - \frac{\partial}{\partial x_\alpha} \left(\frac{\partial L}{\partial p_\alpha^i} \right) \right) \varphi^i(x) dx.$$

Mivel ez minden φ^i tesztfüggvényre igaz, a variációszámítás (többdimenziós) fundamentális lemmája szerint az integrálandó kifejezésnek el kell tűnnie.

1.39. Lemma. (A variációszámítás fundamentális lemmája) Legyen $g \in C^0(\Omega)$ olyan, hogy $\int_{\Omega} g(x)\varphi(x) dx = 0$ minden $\varphi \in C_c^\infty(\Omega)$ esetén. Ekkor $g \equiv 0$ az Ω -n.

1.40. Definíció. Az L Lagrange függvényhez tartozó **Euler-operátor** (E_L):

$$E_L(f)_i := \frac{\partial L}{\partial \xi^i} - \sum_{\alpha=1}^N \frac{\partial}{\partial x_\alpha} \left(\frac{\partial L}{\partial p_\alpha^i} \right) \quad (i = 1, \dots, M).$$

Ezt tömörebben is írhatjuk: $E_L(f) = L_\xi - \operatorname{div}(L_p) = 0$.

1.41. Tétel. Legyen $\Omega \subset \mathbb{R}^N$ egy korlátos, nyílt halmaz C^1 peremmel. Tegyük fel, hogy az L Lagrange függvény C^2 osztályú. Ha $f \in C^2(\Omega; \mathbb{R}^M)$ egy gyenge lokális minimalizálója az I funkcionálnak, akkor kielégíti az $E_L(f) = 0$ Euler–Lagrange-egyenletrendszer Ω -ban.

1.42. Példa. (Példák) • **Dirichlet-funkcionál:** $I(f) = \frac{1}{2} \int_{\Omega} |\nabla f|^2 dx$. Az Euler–Lagrange-egyenlet a **Laplace-egyenlet:** $\Delta f = 0$. Mivel a Lagrange függvény szigorúan konvex, minden megoldás egyedi minimalizáló.

- **Poisson-egyenlet:** Ha a funkcionál $I(f) = \int_{\Omega} (\frac{1}{2}|\nabla f|^2 + hf)dx$, az egyenlet $\Delta f = h$.
- **Minimálfelület-egyenlet:** Az $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ függvény grafikonjának felszínét adó $A(f) = \int_{\Omega} \sqrt{1 + |\nabla f|^2}dx$ funkcionál Euler–Lagrange-egyenlete:

$$\operatorname{div} \left(\frac{\nabla f}{\sqrt{1 + |\nabla f|^2}} \right) = 0.$$

- **Előírt közepes görbületű felület:** Az $I(f) = \int_{\Omega} (\sqrt{1 + |\nabla f|^2} + Hu) dx$ funkcionál egyenlete:

$$\operatorname{div} \left(\frac{\nabla f}{\sqrt{1 + |\nabla f|^2}} \right) = H.$$

1.6.2. Természetes peremfeltételek

Ha a minimalizálási feladatban a peremértékek nincsenek rögzítve (azaz a $C^1(\overline{\Omega}; \mathbb{R}^M)$ téren minimalizálunk), a parciális integrálás egy peremintegrált is eredményez:

$$\int_{\Omega} (E_L(f)_i \varphi^i) dx + \int_{\partial\Omega} \left(\frac{\partial L}{\partial p_{\alpha}^i} \nu_{\alpha} \right) \varphi^i dS = 0.$$

Mivel az Euler–Lagrange-egyenlet az Ω belsejében továbbra is érvényes, a peremintegrálnak el kell tűnnie minden φ variációra. A variációszámítás peremre vonatkozó fundamentális lemmája szerint ez csak akkor lehetséges, ha a φ^i együtthatója nulla.

1.43. Tétel. Egy lokális minimalizálónak nemcsak az $E_L(f) = 0$ egyenletet kell kielégítenie Ω -ban, hanem a peremen a következő **természetes peremfeltételt** is:

$$\sum_{\alpha=1}^N \frac{\partial L}{\partial p_{\alpha}^i} \nu_{\alpha} = 0 \quad \text{az } \partial\Omega\text{-n minden } i = 1, \dots, M \text{ esetén,}$$

ahol ν a peremre állított külső normálvektor.

1.44. Példa. • A Dirichlet-funkcionál esetében ez a homogén **Neumann-peremfeltételt** adja: $\frac{\partial f}{\partial \nu} = 0$.

- A minimálfelület-funkcionálnál a természetes peremfeltétel geometriai jelentése az, hogy a felület merőlegesen metszi a peremet.

1.6.3. Belső variációk és az energia-impulzus tenzor

A független x változó variálásával ("belső variációk") a Du Bois–Reymond-egyenlet általánosítását kapjuk. Az ötlet az, hogy az Ω tartományt egy η vektormező mentén "deformáljuk", és a funkcionál stacionaritását vizsgáljuk. A levezetés az **energia-impulzus tenzor** fogalmához vezet.

1.45. Definíció. Az **energia-impulzus tenzor** egy $T_{\alpha\beta}$ mátrix, melynek elemei:

$$T_{\alpha\beta} := \frac{\partial L}{\partial p_\alpha^i} \frac{\partial f^i}{\partial x_\beta} - L \delta_{\alpha\beta}.$$

1.46. Tétel. Ha f egy lokális minimalizáló, akkor teljesül rá a $\operatorname{div}(T) + L_x = 0$ egyenlet, ami egy megmaradási törvényt fejez ki. (Itt L_x az L explicit x -függéséből származó parciális derivált.)

1.6.4. Izoperimetrikus problémák

Egy $G(f) = c$ integrális kényszer esetén a megoldás egy λ **Lagrange-multiplikátor** konstanssal módosított Euler-Lagrange egyenletet elégít ki, amely a $H = L + \lambda G_{\text{lagrangian}}$ funkcionálhoz tartozik.

1.6.5. Holonóm kényszerek

Ha a kényszer egy $G(\mathbf{x}, f(\mathbf{x})) = 0$ egyenlettel adott (azaz a megoldásnak egy sokaságon kell lennie), a szükséges feltétel az, hogy az Euler-operátor merőleges legyen a kényzersokaság érintőjára. Ez $\lambda_j(\mathbf{x})$ **függvény értékű Lagrange-multiplikátorok** megjelenéséhez vezet:

$$E_L(f)_i = \sum_{j=1}^k \lambda_j(\mathbf{x}) \frac{\partial G_j}{\partial \xi^i}.$$

1.7. Másodrendű szükséges feltételek

Eddig csak az első variáció eltűnésével kapcsolatos szükséges feltételeket vizsgáltuk. Most magasabb rendű szükséges feltételeket vizsgálunk. Három ilyen fogunk látni: a második variáció nem-negativitását, a Legendre–Hadamard-feltételt gyenge lokális minimalizálókra, és a Weierstrass-feltételt erős lokális minimalizálókra. Az első teljes általánosságban bizonyítjuk, míg a másik kettő esetében a görbék, azaz $f : [a, b] \rightarrow \mathbb{R}^M$ függvények esetére szakosodunk. Ennek oka, hogy ebben az esetben a számítások egyszerűbbek, így jobban megragadható a fő gondolat. Minden bizonyítás, technikai részletektől eltekintve, adaptálható az általánosabb esetre.

1.7.1. A második variáció nem-negativitása

Ez az első feltétel nem meglepő. A véges dimenziós esetben egy lokális minimumhely egyben kritikus és stabil pont is, azaz $\nabla g(\mathbf{x}) = 0$ és $D^2g(\mathbf{x}) \geq 0$. Azt állítjuk, hogy ugyanez igaz a variációs integrálokra is.

1.47. Tétel. Legyen L egy C^2 osztályú Lagrange függvény, és tegyük fel, hogy $f \in C^1(\Omega; \mathbb{R}^M)$ egy gyenge lokális minimalizáló. Ekkor

$$\delta^2 I(f)[\varphi] \geq 0$$

minden $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ esetén, ahol

$$\delta^2 I(f)[\varphi] := \int_{\Omega} \left(\varphi^T L_{\xi\xi}(x, f, Df) \varphi + 2\varphi^T L_{\xi p}(x, f, Df) D\varphi + (D\varphi)^T L_{pp}(x, f, Df) D\varphi \right) dx.$$

Bizonyítás. Vegyünk egy $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ függvényt, és tekintsük a $\Phi(\varepsilon) := I(f + \varepsilon\varphi)$ függvényt. Tudjuk, hogy $\Phi'(0) = 0$ és $\Phi''(0) \geq 0$. Az állítás az utóbbi feltételből következik.

1.7.2. A Legendre–Hadamard szükséges feltétel

Az előző integrális feltételből szeretnénk egy pontonkénti feltételt levezetni. Kiderül, hogy a második variációban szereplő három tag közül a legfontosabb az L_{pp} -t tartalmazó.

1.48. Tétel. (Legendre–Hadamard-feltétel) Legyen L egy C^2 Lagrange függvény és tegyük fel, hogy u egy gyenge lokális minimalizáló. Ekkor az $M \times M$ -es $L_{pp}(x, f(x), Df(x))$ mátrix pozitív szemidefinit, azaz

$$\eta^T L_{pp}(x, f(x), Df(x)) \eta \geq 0$$

minden $x \in [a, b]$ és $\eta \in \mathbb{R}^M$ esetén.

Bizonyítás. (*A bizonyítás vázlata*) A bizonyítás egy speciális, "tüskés" tesztfüggvény-sorozat konstruálásán alapul. Ezek a függvények egy x_0 pont körül egyre keskenyebb és meredekebb tartományon vesznek fel nem nulla értéket. Behelyettesítve őket a $\delta^2 I(f)[\varphi_k] \geq 0$ feltételbe és elvégezve a $k \rightarrow \infty$ határátmenetet, a deriváltakat tartalmazó tagok dominálnak, és megkapjuk a pontonkénti feltételt.

1.7.3. A Weierstrass-féle szükséges feltétel

Ha feltesszük, hogy f egy **erős** lokális minimalizáló, egy globálisabb jellegű szükséges feltételt kaphatunk.

1.49. Definíció. (Weierstrass-féle többlet függvény) Az \mathcal{E} **többlet függvény** a következőképpen definiált:

$$\mathcal{E}(x, \xi, p, q) := L(x, \xi, q) - L(x, \xi, p) - (q - p)^T L_p(x, \xi, p).$$

Geometriailag az \mathcal{E} függvény azt méri, hogy a $p \mapsto L(x, \xi, p)$ függvény mennyivel van a p_0 pontbeli érintője felett. A konvex függvényekre $\mathcal{E} \geq 0$.

1.50. Tétel. (Weierstrass-feltétel) Ha f egy erős lokális minimalizáló, akkor

$$\mathcal{E}(x, f(x), Df(x), q) \geq 0$$

minden $x \in [a, b]$ és minden $q \in \mathbb{R}^{M \times N}$ esetén.

Bizonyítás. (*A bizonyítás vázlata*) A bizonyítás egy "fűrészfog" alakú variáció konstruálásán alapul. Egy kis intervallumon eltérítjük a megoldást egy lineáris függvénnyel, majd visszavezetjük az eredeti görbéhez. Mivel erős minimumról van szó, a deriváltak tetszőlegesen nagyok lehetnek, így ez a variáció megengedett. A funkcionál változásának nulladrendűnek kell lennie, amiből a határátmenet után az $\mathcal{E} \geq 0$ feltétel adódik.

1.8. Null-Lagrange függvények

Ebben a részben olyan Lagrange függvényeket vizsgálunk, amelyekre $E_L(f) = 0$ minden f függvényre. Ezek azért érdekesek, mert az Euler–Lagrange-egyenlet nem ad semmilyen információt a minimumhelyekről.

1.51. Definíció. Egy L Lagrange függvény **null-lagrange függvénynek** nevezünk, ha $L_L(f) = 0$ minden sima f függvényre.

1.52. Állítás. A következő állítások ekvivalensek:

- (i) L egy null-lagrange függvény.
- (ii) Az $I(f)$ funkcionál értéke csak a peremértékektől függ.
- (iii) L egy teljes divergencia, azaz létezik egy Φ vektormező, amelyre $L = \text{div}(\Phi)$.

Bizonyítás. (*A bizonyítás vázlata*) (i) \Leftrightarrow (ii) az Euler-operátor definíciójából és egy útvonal-integrálos trükkből következik. (iii) \Rightarrow (i) a Stokes-tétel direkt következménye, mivel egy teljes divergencia integrálja a peremen kiértékelt potenciállal egyenlő.

Az egydimenziós skalár esetben ($N = M = 1$) a null-lagrange függvény szerkezete különösen egyszerű:

1.53. Állítás. Egy $L(x, \xi, p)$ Lagrange függvény pontosan akkor null-lagrange függvény, ha létezik egy $S(x, \xi)$ függvény, amelyre

$$L(x, \xi, p) = \frac{\partial S}{\partial x}(x, \xi) + p \frac{\partial S}{\partial \xi}(x, \xi) = \frac{d}{dx} S(x, f(x)).$$

Ez azt jelenti, hogy az $I(f)$ funkcionál egyszerűen $I(f) = \int_a^b \frac{d}{dx} S(x, f(x)) dx = S(b, f(b)) - S(a, f(a))$, ami nyilvánvalóan csak a peremértékektől függ.

1.9. Elégséges feltételek

Ebben a fejezetben a következő kérdésre keressük a választ: milyen feltételeket kell hozzáadnunk az előző fejezetekben levezetett első- és másodrendű szükséges feltételekhez annak érdekében, hogy gyenge vagy erős lokális minimalitást biztosítsunk?

1

¹Vegyük észre, hogy az összes bemutatott feltétel nem globális minimalitási tulajdonságokkal foglalkozik, hanem csak lokálisakkal. A globális minimalitási eredmények eléréséhez bonyolultabb érvelések szükségesek.

1.9.1. A második variáció koercitivitása

Az I második variációjának vizsgálatával kezdünk. Tudjuk, hogy a véges dimenziós esetben, ha van egy $g \in C^2(\mathbb{R}^N)$ függvényünk és egy $x_0 \in \mathbb{R}^N$ pontunk, amelyre

$$\nabla g(x_0) = 0 \quad \text{és} \quad D^2g(x_0) > 0,$$

akkor x_0 a g függvény izolált lokális minimalizálója. Így feltételezhetnénk, hogy ugyanez igaz a variációs számítás funkcionáljaira is. A válasz azonban nemleges, amint azt a következő, Bolza által bemutatott példa mutatja. A probléma az, hogy egy végtelen dimenziós térben vagyunk. A véges dimenziós esetben, ha egy folytonos függvény pozitív az egységgyömbön, akkor az infimuma is pozitív a kompaktság miatt. Végtelen dimenzióban az egységgyömb nem kompakt, így ez az érvelés megbukik. Előfordulhat, hogy létezik egy olyan függvénysorozat, amelyen a második variáció nullához tart, anélkül, hogy a sorozat normában konvergálna. A helyes feltétel a **koercitivitás**.

1.54. Tétel. Legyen L egy C^2 osztályú Lagrange függvény, és legyen $f \in C^2((a, b))$ egy kritikus pont, amelyre

$$\inf_{\|\varphi\|_{H^1}=1} \delta^2 I(f)[\varphi] = c > 0,$$

azaz létezik $c > 0$ konstans, hogy

$$\delta^2 I(f)[\varphi] \geq c \int_a^b ((\varphi'(x))^2 + (\varphi(x))^2) dx = c \|\varphi\|_{H^1}^2$$

minden $\varphi \in C_c^\infty((a, b))$ tesztfüggvényre. Ekkor f egy gyenge lokális minimalizáló. Pontosabban, f egy izolált gyenge lokális minimalizáló.

1.55. Megjegyzés. A fenti érvelés nem alkalmazható erős lokális minimalitás elérésére, mivel valóban szükségünk van a függvény deriváltjának kontrolljára az L másodrendű deriváltjainak becsléséhez.

1.9.2. Jacobi-féle konjugált pontok

Most azt vizsgáljuk, mikor lehetséges a második variáció alulról történő becslése, ahogy azt a tétel hipotézise megköveteli. Ebben a szakaszban csak $f \in C^2((x_1, x_2))$ függvényekkel foglalkozunk. Rögzítsünk egy f kritikus pontot és definiáljuk:

$$a(x) := L_{pp}(x, f, f'), \quad b(x) := L_{p\xi}(x, f, f'), \quad c(x) := L_{\xi\xi}(x, f, f').$$

A Legendre–Hadamard-feltételből tudjuk, hogy $a(x) \geq 0$. Tegyük fel az erős Legendre-feltételt, azaz $a(x) > 0$. Jacobi és Legendre ötlete az volt, hogy a járulékos Lagrange függvényt (a második variáció integranduszát) teljes négyzetté egészítsék ki egy null-lagrange függvény hozzáadásával. Ez egy w segédfüggvény bevezetéséhez és egy Riccati-típusú egyenlethez vezet. Egy okos helyettesítéssel ($w = -a \frac{v'}{v} - b$) ez az egyenlet egy másodrendű lineáris differenciálegyenletre redukálható:

Ez azonban nem meglepő, hiszen, ahogy a véges dimenziós esetben is, ezek a technikák a 'második deriváltak' tulajdonságain alapulnak, amelyek definíciójuk szerint lokális természetűek.

$$(a(x)v'(x))' + (c(x) - b'(x))v(x) = 0. \quad (9)$$

Ezt az egyenletet **Jacobi-egyenletnek** nevezzük. Egy pozitív megoldását **Jacobi-mezőnek**. Ha létezik Jacobi-mező (azaz egy $v > 0$ megoldás) az $[x_1, x_2]$ intervallumon, akkor a második variáció alulról becsülhető, és a kritikus pont egy gyenge lokális minimalizáló.

1.56. Tétel. Legyen L egy C^2 osztályú Lagrange függvény és $f \in C^2((x_1, x_2))$ egy kritikus pont. Tegyük fel, hogy létezik egy Jacobi-mező (amely f -hoz tartozik) az $[x_1, x_2]$ intervallumon. Ekkor f egy szigorú gyenge lokális minimalizáló.

1.9.3. Konjugált pontok és a gyenge lokális minimalitás szükséges feltétele

A Jacobi-egyenlet megoldásainak zérushelyei központi szerepet játszanak.

1.57. Definíció. Egy x_0 ponthoz **konjugált pontoknak** nevezzük a Jacobi-egyenlet azon megoldásának izolált zérushelyeit, amely $v(x_0) = 0$ és $v'(x_0) \neq 0$ feltételekkel indul.

A Sturm-féle oszcillációs tétel segítségével megmutatható, hogy egy lineárisan független megoldáspár zérushelyei "váltogatják" egymást. Ez a konjugált pontok elméletének alapja.

1.58. Tétel. Legyen f egy kritikus pont, és tegyük fel, hogy az erős Legendre-feltétel ($a(x) > 0$) teljesül. Ekkor:

- (i) ha az $(x_1, x_2]$ intervallumban nincsenek x_1 -hez konjugált pontok, akkor f egy izolált gyenge lokális minimalizáló.
- (ii) ha létezik az (x_1, x_2) nyílt intervallumban x_1 -hez konjugált pont, akkor f nem gyenge lokális minimalizáló.
- (iii) ha az első konjugált pont éppen x_2 , akkor a helyzet határeset, bármi előfordulhat.

1.9.4. Weierstrass-féle térelmélet (erős minimumhoz)

Eddig a gyenge lokális minimalitáshoz adtunk elégséges feltételeket. Ebben a szakaszban az erős lokális minimalitást biztosító feltételeket vizsgáljuk.

1.59. Definíció. Egy extrémálisokból álló $f(x, \alpha)$ sereg egy **extrémálisok mezejét** alkotja, ha a görbesereg egyszeresen fedi le a sík egy tartományát.

Az elmélet lényege, hogy egy f_0 extrémalist beágyazunk egy ilyen mezőbe. Ez lehetővé teszi a **Hilbert-féle invariáns integrál** és a **Weierstrass-féle többlet (excess) függvény** definiálását.

$$\mathcal{E}(x, \xi, p, q) = L(x, \xi, q) - L(x, \xi, p) - (q - p)L_p(x, \xi, p)$$

A Weierstrass-féle szükséges feltétel szerint egy erős lokális minimumnál $\mathcal{E} \geq 0$ kell, hogy legyen.

1.60. Tétel. (Elégséges feltétel erős minimumra) Legyen $f \in C^2([a, b])$ egy extrémális. Tegyük fel, hogy:

1. f beágyazható egy extrémálisok mezejébe.
2. Az erős Legendre-feltétel, $L_{pp} > 0$, teljesül.
3. A Weierstrass-féle feltétel, $\mathcal{E} > 0$, teljesül a vizsgált pontok környezetében.

Ekkor f egy izolált erős lokális minimalizáló.

Megmutatható, hogy az extrémálisok mezejébe való beágyazhatóság feltétele szorosan kapcsolódik a konjugált pontok hiányához. Így a konjugált pontok hiánya a gyenge minimum, míg az (erősített) Weierstrass-feltétellel kiegészítve az erős minimum elégséges feltételét adja.

1.10. Fontosabb Példák

Folytatjuk a példák sorát az Euler–Lagrange-egyenletek kiszámításával és megoldásával az 1.1. szakaszból származó példákra.

1.10.1. A legrövidebb út

Emlékezzünk vissza, hogy a legrövidebb út problémájánál a következő funkcionált szeretnénk minimalizálni:

$$I(f) = \int_0^a g(x, f(x)) \sqrt{1 + f'(x)^2} dx,$$

az $f(0) = 0$ és $f(a) = b$ feltételek mellett. Itt $d = 1$ és a Lagrange függvény:

$$L(x, z, p) = g(x, z) \sqrt{1 + p^2}.$$

Ezért $L_z(x, z, p) = g_z(x, z) \sqrt{1 + p^2}$ és $L_p(x, z, p) = g(x, z)(1 + p^2)^{-\frac{1}{2}}p$. Az Euler–Lagrange-egyenlet:

$$g_z(x, f(x)) \sqrt{1 + f'(x)^2} - \frac{d}{dx} \left(g(x, f(x)) (1 + f'(x)^2)^{-\frac{1}{2}} f'(x) \right) = 0.$$

Ezt általában nehéz megoldani. Abban a speciális esetben, ha $g(x, z) = 1$, akkor $g_z = 0$, és az egyenlet a következőre egyszerűsödik:

$$\frac{d}{dx} \left(\frac{f'(x)}{\sqrt{1 + f'(x)^2}} \right) = 0.$$

A deriváltat elvégezve $f''(x) = 0$ adódik, tehát a megoldás egy egyenes! Ez megerősíti azt az intuíciónkat, hogy két pont között a legrövidebb út az egyenes.

1.10.2. A brachisztochron-probléma

A brachisztochron-problémánál a legrövidebb lecsúszási időt keressük, ami a következő funkcionállal írható le:

$$T[y(x)] = \int_A^B dt = \int_A^B \frac{ds}{v}.$$

Homogén gravitációs térben, ha a test az origóból indul ($y(0) = 0$), az energiamegmaradás törvénye szerint $v = \sqrt{-2gy}$ (a lefelé mutató y tengely miatt). Az $ds = \sqrt{1 + y'^2} dx$ ívhosszelemmel a funkcionál:

$$T[y] = \frac{1}{\sqrt{2g}} \int_0^a \sqrt{\frac{1 + y'(x)^2}{-y(x)}} dx,$$

az $y(0) = 0$ és $y(a) = b$ ($b < 0$) feltételek mellett. Itt a Lagrange függvény

$$L(x, y, y') = \sqrt{\frac{1 + y'^2}{-y}}.$$

Vegyük észre, hogy L nem függ expliciten x -től. Ezért használhatjuk az Euler–Lagrange-egyenlet alternatív, energiamegmaradást kifejező alakját (Beltrami-azonosság: $y' \frac{\partial L}{\partial y'} - L = \text{konstans}$), amely a következőhöz vezet:

$$y(x)(1 + y'(x)^2) = -C_0, \quad (10)$$

ahol C_0 egy pozitív konstans. Ebből több kvalitatív tulajdonság is levezethető:

- A pálya függőlegesen indul, $\lim_{x \rightarrow 0^+} y'(x) = -\infty$.
- A függvény konvex ($y''(x) > 0$).
- A pálya szimmetrikus a legmélyebb pontjára.

Az (10) egyenlet egy elsőrendű, szétválasztható differenciálegyenlet y' -re:

$$\frac{dy}{dx} = -\sqrt{\frac{-C_0 - y}{y}}.$$

Az integrál elvégzéséhez célszerű egy új θ paramétert bevezetni a következő helyettesítéssel:

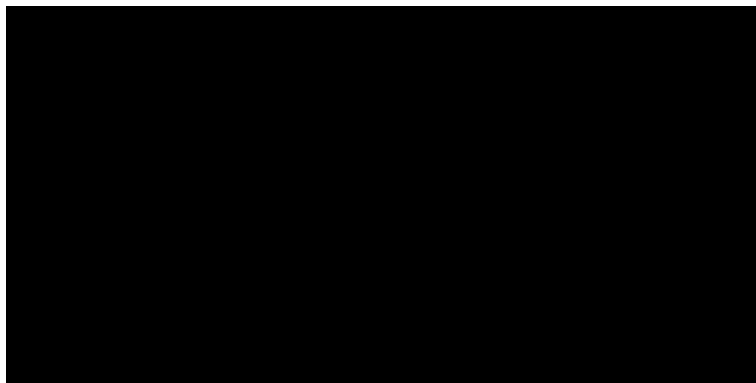
$$y(\theta) = -\frac{C_0}{2}(1 - \cos \theta).$$

Ezt visszahelyettesítve és integrálva x -re, a következő paraméteres alakot kapjuk:

$$\begin{aligned} x(\theta) &= \frac{C_0}{2}(\theta - \sin \theta) \\ y(\theta) &= -\frac{C_0}{2}(1 - \cos \theta) \end{aligned}$$

A C_0 konstans úgy kell megválasztani, hogy a görbe áthaladjon a (a, b) végponton.

A megoldás egy **ciklois** görbe, amelyet egy egyenesen legördülő kerék peremének egy pontja ír le.



1. ábra. Brachisztochron-görbék családja. A megoldás egy lefelé fordított ciklois ív.

Érdekesség, hogy a ciklois egyben **tautochron** görbe is: a súrlódásmentesen lecsúszó testnek ugyanannyi időbe telik elérni a legalsó pontot, bárhonnan is indítjuk a pályán.

1.10.3. Minimálfelületek

A minimálfelület-probléma célja az

$$I(f) = \int_U \sqrt{1 + |\nabla f|^2} dx$$

funkcionál minimalizálása, az $f = g$ peremfeltétel mellett a ∂U peremen. Itt $L_z = 0$, így az Euler–Lagrange-egyenlet a következő:

$$\operatorname{div} \left(\frac{\nabla f}{\sqrt{1 + |\nabla f|^2}} \right) = 0 \quad U\text{-ban.} \quad (11)$$

Ezt **minimálfelület-egyenletnek** nevezik. Kifejtve egy bonyolult, nemlineáris másodrendű parciális differenciálegyenletet kapunk:

$$(1 + |\nabla f|^2) \Delta f - \sum_{i,j=1}^d f_{x_i x_j} f_{x_i} f_{x_j} = 0.$$

1.61. Példafeladat. Mutassuk meg, hogy az $f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b$ sík megoldja a minimálfelület-egyenletet.

1.62. Példafeladat. Mutassuk meg, hogy $n = 2$ esetén a Scherk-felület, $f(x_1, x_2) = \log \left(\frac{\cos(x_1)}{\cos(x_2)} \right)$, megoldja a minimálfelület-egyenletet a $(-\pi/2, \pi/2) \times (-\pi/2, \pi/2)$ négyzeten.

1.10.4. Forgási minimálfelület

Mi a minimális felület alakja két, egymástól $2L$ távolságra lévő, r sugarú gyűrű között? A szimmetria miatt feltételezhetjük, hogy a felület egy $f : [-L, L] \rightarrow \mathbb{R}$ függvény x -tengely körüli

megforgatásával jön létre. A felszín:

$$I(f) = 2\pi \int_{-L}^L f(x) \sqrt{1 + f'(x)^2} dx.$$

A Lagrange függvény nem függ expliciten x -től, így a Beltrami-azonosságot használva a következő egyenlethez jutunk:

$$cf(x) = \sqrt{1 + f'(x)^2}.$$

Az egyenlet megoldása egy **lánCGörbe** (katenáris):

$$f(x) = \frac{1}{c} \cosh(cx).$$

Az ebből származó forgásfelület a **katenoid**. Érdekes, hogy nem mindig létezik megoldás: ha a gyűrűk túl messze vannak egymástól a sugarukhoz képest (r/L arány túl kicsi), a szappanbuborék elpattan, és nem jön létre stabil minimálfelület. Ha létezik megoldás, akkor általában kettő is van; ezek közül a "kevésbé domború" adja a valódi minimumot.

1.10.5. Fénysugár útja változó törésmutatójú közegben

Milyen pályán halad a fénysugár, ha a törésmutató $n(y) = n_0 \frac{d}{y}$? A Fermat-elv szerint a fénysugár azt az utat követi, amelyen az optikai úthossz minimális:

$$S = \int n(\mathbf{r}) ds.$$

Az $ds = \sqrt{1 + y'^2} dx$ ívhosszelemmel a minimalizálandó funkcionál:

$$S[y] = \int_{x_0}^{x_1} n_0 \frac{d}{y} \sqrt{1 + y'(x)^2} dx.$$

A Lagrange függvény $L(y, y') = n_0 \frac{d}{y} \sqrt{1 + y'^2}$ nem függ expliciten x -től, így ismét a Beltrami-azonosságot használhatjuk:

$$y' \frac{\partial L}{\partial y'} - L = E \quad (\text{konstans}).$$

Behelyettesítés és egyszerűsítés után a következőhöz jutunk:

$$-\frac{n_0 d}{y \sqrt{1 + y'^2}} = E.$$

Átrendezve y' -re:

$$y'(x) = \frac{dy}{dx} = \pm \sqrt{\frac{n_0^2 d^2}{E^2 y^2} - 1}.$$

Ez egy szétválasztható differenciálegyenlet. A változókat szétválasztva és integrálva:

$$\int dx = \pm \int \frac{y}{\sqrt{\frac{n_0^2 d^2}{E^2} - y^2}} dy.$$

Az integrál elvégzése után kapjuk:

$$x + c = \mp \sqrt{\frac{n_0^2 d^2}{E^2} - y^2}.$$

Mindkét oldalt négyzetre emelve és átrendezve az eredmény:

$$(x + c)^2 + y^2 = \frac{n_0^2 d^2}{E^2}.$$

Tehát a fénysugár pályája egy olyan körív, amelynek a középpontja az x -tengelyen helyezkedik el.

1.10.6. Izoperimetrikus egyenlőtlenség

Melyik az a síkbeli zárt görbe, amely adott ℓ kerület mellett a lehető legnagyobb A területet zárja körbe? A sejtés (és a helyes válasz) a kör. Ezt az állítást az **izoperimetrikus egyenlőtlenség** fogalmazza meg:

$$4\pi A \leq \ell^2,$$

ahol egyenlőség pontosan akkor áll fenn, ha a görbe egy kör. A bizonyításhoz a Lagrange-multiplikátoros módszert alkalmazzuk funkcionálokra. Maximalizáljuk az $A(x, y)$ területfunkcionált az $\ell(x, y) = \ell_0$ kerület-kényszer mellett. Az Euler–Lagrange-egyenletek megoldása valóban egy kört ad.

1.10.7. Képhelyreállítás

A zajos képek "megtisztítására" szolgáló egyik sikeres módszer a **teljes variáció (Total Variation, TV)** regularizált minimalizálás. A feladat az

$$I(u) = \int_U \left(\frac{1}{2}(f - u)^2 + \lambda |\nabla u| \right) dx$$

funkcionál minimalizálása az összes $u : U \rightarrow \mathbb{R}$ függvényre, ahol $U = (0, 1)^2$ a képtartomány. Az f függvény az eredeti zajos kép, a minimalizáló u pedig a zajtalanított kép. A Lagrange függvény itt:

$$L(x, z, p) = \frac{1}{2}(f(x) - z)^2 + \lambda |p|.$$

Ez a Lagrange függvény a $|p| = |\nabla u|$ tag miatt nem differenciálható $p = 0$ -ban. Ez kisebb problémákat okoz a numerikus szimulációk során, ezért gyakori, hogy a TV funkcionál egy differenciálható approximációját használják. Egy népszerű választás:

$$I_\varepsilon(u) = \int_U \left(\frac{1}{2}(f - u)^2 + \lambda \sqrt{|\nabla u|^2 + \varepsilon^2} \right) dx,$$

ahol $\varepsilon > 0$ egy kicsi paraméter. Ha $\varepsilon = 0$, visszkapjuk az eredeti teljes variáció funkcionált. Ebben az esetben a Lagrange függvény:

$$L_\varepsilon(x, z, p) = \frac{1}{2}(f(x) - z)^2 + \lambda \sqrt{|p|^2 + \varepsilon^2},$$

amely már z -ben és p -ben is differenciálható. Bizonyítható, hogy az I_ε minimalizálói konvergálnak az I minimalizálóihoz, amint $\varepsilon \rightarrow 0$, de a bizonyítás nagyon technikai. Az ötlet tehát az, hogy rögzítünk egy kis $\varepsilon > 0$ értéket, és minimalizáljuk az I_ε funkcionált. Az Euler–Lagrange-egyenlet kiszámításához vegyük észre, hogy:

$$L_{\varepsilon,z}(x, z, p) = z - f(x) \quad \text{és} \quad \nabla_p L_\varepsilon(x, z, p) = \frac{\lambda p}{\sqrt{|p|^2 + \varepsilon^2}}.$$

Ezért az Euler–Lagrange-egyenlet a következő:

$$u - \lambda \operatorname{div} \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \right) = f \quad U\text{-ban},$$

homogén Neumann-peremfeltételekkel: $\frac{\partial u}{\partial \nu} = 0$ a ∂U peremen. Ezt az egyenletet szinte soha nem lehet analitikusan megoldani, így numerikus approximációkat kell alkalmaznunk.

1.10.8. Képszegmentáció

A képszegmentáció célja egy kép felosztása értelmes régiókra, például egy objektum elválasztása a háttértől. A Chan–Vese modell ezt a feladatot a következő funkcionál minimalizálásaként fogalmazza meg:

$$I(u, a, b) = \int_U (H(u)(f - a)^2 + (1 - H(u))(f - b)^2 + \lambda \delta(u) |\nabla u|) dx,$$

ahol a minimalizálás az $u : U \rightarrow \mathbb{R}$ "szintfüggvényre" és az a, b valós számokra történik. Itt f a szegmentálandó kép, a és b a két régió (pl. objektum és háttér) átlagos intenzitása, u pedig egy olyan függvény, amelynek nulla-szinthalmaza ($u(x) = 0$) adja a régiók közötti határt. A $H(u)$ a Heaviside-függvény, a $\delta(u)$ pedig a Dirac-delta függvény, amely a határvonal hosszát "méri".

A Lagrange függvény:

$$L(x, z, p) = H(z)(f(x) - a)^2 + (1 - H(z))(f(x) - b)^2 + \lambda \delta(z) |p|$$

a Heaviside- és delta-függvények miatt nem folytonos, ami numerikus problémákat okoz. A gyakorlatban ezért sima approximációkat használunk. Egy $\varepsilon > 0$ paraméterrel definiáljuk a sima Heaviside-függvényt:

$$H_\varepsilon(x) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \left(\frac{x}{\varepsilon} \right) \right).$$

Ennek deriváltja a sima delta-függvény approximációja:

$$\delta_\varepsilon(x) := H'_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + x^2}.$$

Ezzel a simított funkcionál:

$$I_\varepsilon(u) = \int_U (H_\varepsilon(u)(f - a)^2 + (1 - H_\varepsilon(u))(f - b)^2 + \lambda \delta_\varepsilon(u) |\nabla u|) dx.$$

Ennek a Lagrange függvénynek már levezethetjük az Euler–Lagrange-egyenletét u -ra (rögzített a, b mellett). A számítások elvégzése után a következő PDE-t kapjuk:

$$\delta_\varepsilon(u) [(f - a)^2 - (f - b)^2] - \lambda \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = 0 \quad U\text{-ban},$$

homogén Neumann-peremfeltételek mellett ($\frac{\partial u}{\partial \nu} = 0$).

Mivel a minimalizálás u, a és b szerint történik, egy **alternáló minimalizálási algoritmust** alkalmazunk.

1. **a és b frissítése:** Rögzített u mellett az I_ε funkcionál minimuma a -ra és b -re analitikusan számolható. Az optimális értékek a két régió súlyozott átlagai:

$$a = \frac{\int_U H_\varepsilon(u) f \, dx}{\int_U H_\varepsilon(u) \, dx}, \quad b = \frac{\int_U (1 - H_\varepsilon(u)) f \, dx}{\int_U (1 - H_\varepsilon(u)) \, dx}.$$

2. **u frissítése:** Rögzített a és b mellett végzünk egy kis lépést az I_ε funkcionál gradiensének ellentétes irányába. Ez a **gradiens ereszkedés** egy parciális differenciálegyenlet megoldását jelenti:

$$\frac{\partial u}{\partial t} + \delta_\varepsilon(u) [(f - a)^2 - (f - b)^2] - \lambda \operatorname{div} \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \right) = 0.$$

(Itt a $|\nabla u|$ tagot is simítjuk a nevezőben, hogy elkerüljük a nullával való osztást.)

Ezt a két lépést iteratívan ismételjük, amíg a folyamat konvergál. Az algoritmus eredményeként az u függvény nulla-szinthalmaza kirajzolja a szegmentált kép határait.

1.10.9. Variációs autoenkóder (VAE)

A variációs autoenkóder (VAE) egy generatív, mélytanuló modell. A modell két részből áll: egy **enkóderből** ($q_\phi(z|x)$) és egy **dekóderből** ($p_\theta(y|z)$). Az enkóder egy x bemeneti képet egy alacsony dimenziós latens tér valószínűségi eloszlására képezi le (ezt z -vel jelöljük). A dekóder egy z latens vektorból rekonstruálja a megfelelő y szegmentációs maszkot.

A VAE célja az evidencia alsó korlátjának (Evidence Lower Bound, ELBO) maximalizálása, ami ekvivalens a következő negatív ELBO funkcionál minimalizálásával:

$$L(\theta, \phi; x, y) = \mathbb{E}_{q_\phi(z|x)} [-\log p_\theta(y|z)] + D_{KL}(q_\phi(z|x) || p(z)),$$

ahol a minimalizálás az enkóder (ϕ) és a dekóder (θ) paramétereire történik.

- Az első tag a **rekonstrukciós hiba**. Azt méri, hogy a dekóder által generált maszk mennyire illeszkedik a valós y maszkra. Kétsíkú (objektum/háttér) szegmentációnál ezt tipikusan pixel-szintű bináris keresztentrópiával valószínűsítjük meg. Ez a tag analóg a Chan-Vese modell adatilleszkedési tagjaival.
- A második tag a **Kullback–Leibler-divergencia** (D_{KL}). Ez egy regularizációs tag, amely biztosítja, hogy az enkóder által tanult latens eloszlás ($q_\phi(z|x)$) közel maradjon egy egyszerű, előre definiált prior eloszláshoz ($p(z)$), ami tipikusan egy standard normális eloszlás ($\mathcal{N}(0, I)$). Ez a tag analóg a Chan-Vese modell regularizációs tagjával, mivel mindkettő a megoldás "simaságát" szorgalmazza.

A fő numerikus probléma az, hogy a latens vektort a $q_\phi(z|x)$ eloszlásból való mintavételezéssel kapjuk ($z \sim q_\phi(z|x)$), ami nem teszi lehetővé a gradiens vissza-terjesztését. Ennek feloldására

a **reparametrizációs trükköt** alkalmazzuk. Az enkóder nem közvetlenül a z vektort, hanem annak eloszlásának paramétereit, egy átlagvektort (μ) és egy kovarianciamátrixot (σ^2) tanulja meg. A mintavételezés ezután így történik:

$$z = \mu + \sigma \odot \epsilon, \quad \text{ahol} \quad \epsilon \sim \mathcal{N}(0, I).$$

Így a véletlenszerűség "külsővé" válik, és a gradiens szabadon áramolhat vissza μ -n és σ -n keresztül.

Az optimalizálás egy **gradiens ereszkedési algoritmussal** történik.

1. **Előre terjesztés (Forward Pass):** Egy (x, y) kép-maszk párt adunk a modellnek.
 - (a) Az enkóder az x képből kiszámítja a μ és σ paramétereket.
 - (b) A reparametrizációs trükkel mintát veszünk a látens térből: $z = \mu + \sigma \odot \epsilon$.
 - (c) A dekóder a z latens vektorból legenerálja a \hat{y} inferált maszkot.
2. **Hiba számítása:** Kiszámítjuk a teljes L hibát a rekonstrukciós hiba és a KL-divergencia összegeként.
3. **Vissza-terjesztés (Backward Pass):** A teljes hiba gradiensét visszaterjesztjük a hálózaton, és frissítjük az enkóder (ϕ) és a dekóder (θ) súlyait.

Ezt a folyamatot iteratívan ismétljük egy nagy adathalmazon. A betanított modell ezután képes új képekhez is szegmentációs maszkokat generálni.

2. Tenzor kalkulus

2.1. Vektorterek és bázisok

Legyen V egy n -dimenziós vektortér \mathbb{R} felett, ellátva egy $\{e_i\}$ bázissal.

2.1. Jelölés. Amennyiben másként nincs jelezve, az azonos alsó és felső indexpárokra az Einstein-féle összegzési konvenció érvényes. Tehát, ha X_j^i egy tetszőleges számhalmaz, akkor $X_i^i := \sum_{i=1}^n X_i^i$.

Azt a műveletet, amikor egy felső és egy alsó indexet egyenlővé teszünk – és ezzel hallgatólagosan feltételezzük az összegzést (v.s.) –, **indexkontrakciónak** vagy **nyomképzésnek** nevezzük.

2.2. Megjegyzés. A kontrahált indexek **néma indexek**, amelyek tetszőlegesen átcímkezhetőek, ellentétben a **szabad indexekkel**. Így például $X_i^i = X_j^j$.

2.3. Megjegyzés. A V tér minden $\{e_i\}$ bázisához és minden $x \in V$ elemhez létezik egy egyértelmű $\{x^i\}$ valós számhalmaz úgy, hogy $x = x^i e_i$ (lásd 1. fejezet).

Az x^i valós számokat az x vektor $\{e_i\}$ vektorbázisra vonatkozó **kontravariáns komponenseinek** nevezzük.

2.4. Definíció. A **Kronecker-szimbólum**, vagy **Kronecker-delta**, definíciója a következő:

$$\delta_j^i = \begin{cases} 1 & \text{ha } i = j \\ 0 & \text{ha } i \neq j \end{cases}.$$

2.5. Tétel. Legyen $x = x^i e_i \in V$ és $f_j = A_j^i e_i$ egy báziscsere definíciója. Ha $x^i e_i = y^i f_i$, akkor

$$x^i = A_j^i y^j \quad \text{vagy, ezzel ekvivalensen,} \quad y^i = B_j^i x^j,$$

ahol definíció szerint $A_k^i B_j^k = \delta_j^i$.

Ügyeljünk az A_j^i komponensek megjelenésére az e_i bázisvektorok, illetve az x^i kontravariáns komponensek transzformációjában.

Adott egy sima sokaság – melyet később definiálunk – egy $\{x^i\}$ koordináta-rendszerrel. A lineáris parciális deriválás operátorainak $\{e_i \equiv \partial_i \equiv \partial/\partial x^i\}$ halmaza bármely (implicit) referenciapontban egy vektortérbázist definiál. Egy $x^i = x^i(\bar{x}^j)$ koordináta-transzformáció során a láncszabály a megfelelő bázisokat a következőképpen kapcsolja össze:

$$\bar{\partial}_j = A_j^i \partial_i,$$

ahol $f_j \equiv \bar{\partial}_j \equiv \partial/\partial \bar{x}^j$, és

$$A_j^i = \frac{\partial x^i}{\partial \bar{x}^j}.$$

2.6. Jelölés. Sima sokaságon gyakran találkozunk egy vektor koordináta-bázis szerinti felbontásával, azaz $v = v^i \partial_i$. (A referenciapont itt implicit.)

A $\{\partial_i\}$ halmazt **koordináta (vektor) bázisnak**, vagy **holonóm (vektor) bázisnak** nevezzük. Definíció szerint ez feszíti ki a sokaság **érintőterét** a referenciapontban.

- A bázisvektorokat alsó indexekkel jelöljük.
- A kontravariáns vektorkomponenseket felső indexekkel jelöljük.

2.7. Megjegyzés. A felső és alsó indexek megkülönböztetése csak akkor releváns, ha az indexek eleve arra korlátozódnak, hogy az $\{1, \dots, n\}$ tartományból vegyenek fel értékeket, ahol $n = \dim V$.

2.2. Duális vektorterek és duális bázisok

Legyen V egy n -dimenziós vektortér \mathbb{R} felett.

2.8. Definíció. A V téren értelmezett **lineáris funkcionál** egy lineáris leképezés V -ből \mathbb{R} -be. Más szavakkal, $\hat{f} : V \rightarrow \mathbb{R}$ egy lineáris funkcionál, ha minden $x, y \in V$, $\lambda, \mu \in \mathbb{R}$ esetén,

$$\hat{f}(\lambda x + \mu y) = \lambda \hat{f}(x) + \mu \hat{f}(y).$$

2.9. Megjegyzés. A V^* halmaz, amely V összes lineáris funkcionálját tartalmazza, a szokásos módon vektorterré tehető, nevezetesen: legyen $\hat{f}, \hat{g} \in V^*$ és $\lambda, \mu \in \mathbb{R}$, a $\lambda \hat{f} + \mu \hat{g} \in V^*$ lineáris kombinációt a

$$(\lambda \hat{f} + \mu \hat{g})(x) = \lambda \hat{f}(x) + \mu \hat{g}(x)$$

azonosítás révén definiáljuk minden $x \in V$ esetén (lásd 1. fejezet).

2.10. Definíció. A **duális tér** a V téren értelmezett lineáris funkcionálok vektortere: $V^* = \mathcal{L}(V, \mathbb{R})$.

Legyen V egy n -dimenziós vektortér \mathbb{R} felett, ellátva egy $\{e_i\}$ bázissal, és legyen V^* a duális tere.

2.11. Megjegyzés. A V minden $\{e_i\}$ bázisa indukál egy $\{\hat{e}^i\}$ bázist V^* -ban a következőképpen. Ha $x = x^i e_i$, akkor $\hat{e}^i(x) = x^i$.

A V^* tér $\{\hat{e}^i\}$ duális bázisát egyértelműen meghatározza a definiáló tulajdonsága:

$$\hat{e}^i(e_j) = \delta_j^i.$$

Bizonyítás. (A 2.1. Eredmény bizonyítása.) Legyen $\hat{f} \in V^*$.

- Bármely $x = x^i e_i \in V$ esetén $\hat{f}(x) = \hat{f}(x^i e_i) = \hat{f}(e_i) x^i = \hat{f}(e_i) \hat{e}^i(x) = (\hat{f}(e_i) \hat{e}^i)(x) = (f_i \hat{e}^i)(x)$, ahol $f_i = \hat{f}(e_i)$. Ez azt mutatja, hogy a $\{\hat{e}^i\}$ halmaz kifeszíti a V^* teret.
- Tegyük fel, hogy $\alpha_i \hat{e}^i = 0 \in V^*$. Ekkor $\alpha_i = \alpha_j \delta_j^i = \alpha_j \hat{e}^j(e_i) = (\alpha_j \hat{e}^j)(e_i) = 0 \in \mathbb{R}$ minden $i = 1, \dots, n$ esetén. Ez azt mutatja, hogy a $\{\hat{e}^i\}$ halmaz lineárisan független.

Ezen eredmény fényében a továbbiakban hallgatólagosan feltételezzük, hogy egy adott V vektortér ($\{e_i\}$ bázissal) magában foglalja a duális térnek (V^*) létezését (a megfelelő $\{\hat{e}^i\}$ duális bázissal).

2.12. Megjegyzés. A duális tér, V^* , dimenziója megegyezik V dimenziójával.

Egy $\hat{f} \in V^*$ lineáris funkcionálra alternatívaként **kovektorként** vagy **1-formaként** hivatkozunk. A V vektortér V^* duális terét **kovektor térnek** vagy az **1-formák terének** is nevezik.

2.13. Megjegyzés. A V^* tér minden $\{\hat{e}^i\}$ kovektor bázisához és minden $\hat{x} \in V^*$ elemhez létezik egy egyértelmű $\{x_i\}$ valós számhalmaz úgy, hogy $\hat{x} = x_i \hat{e}^i$.

2.14. Definíció. Az x_i valós számokat az \hat{x} kovektor $\{\hat{e}^i\}$ kovektor bázisra vonatkozó **kovariáns komponenseinek** nevezzük.

Ezt a szakaszt a 7. oldalon található 2.1. Tétel duális analógiájával zárjuk:

2.15. Tétel. Az $f_j = A_j^i e_i$ báziscsere a duális vektorbázis következő cseréjét indukálja:

$$\hat{f}^j = B_i^j \hat{e}^i,$$

ahol $A_j^k B_k^i = \delta_j^i$. Következésképpen, ha $\hat{x} = x_i \hat{e}^i = y_i \hat{f}^i \in V^*$, akkor

$$x_i = B_i^j y_j \quad \text{vagy, ezzel ekvivalensen,} \quad y_i = A_i^j x_j.$$

Bizonyítás. (A 2.2. Tétel bizonyítása.) • Adott $f_j = A_j^\ell e_\ell$ és $\hat{f}^i = B_k^i \hat{e}^k$ esetén kapjuk, hogy $\delta_j^i = \hat{f}^i(f_j) = B_k^i \hat{e}^k(A_j^\ell e_\ell) = B_k^i A_j^\ell \hat{e}^k(e_\ell) = B_k^i A_j^\ell \delta_\ell^k = B_k^i A_j^k$, amiből következik, hogy $B = A^{-1}$.

- Így ha $x_i \hat{e}^i = y_i \hat{f}^i$ -t definiálunk, akkor $y_j \hat{f}^j = y_j B_i^j \hat{e}^i$, amiből $x_i = B_i^j y_j$.
- Inverzióval megkapjuk az utolsó azonosságot.

- A 7. oldalon lévő 2.1. Tétel definiálja az úgynevezett **vektor transzformációs törvényt**.
- A 9. oldalon lévő 2.2. Tétel definiálja az úgynevezett **kovektor transzformációs törvényt**.

2.16. Megjegyzés. Adott egy sima sokaság $\{x^i\}$ koordináta-rendszerrel, és egy $x^i = x^i(\bar{x}^j)$ koordináta-transzformáció. A láncszabály (bármely referenciapontban) kimondja, hogy

$$d\bar{x}^j = B_i^j dx^i \quad \text{ahol} \quad B_i^j = \frac{\partial \bar{x}^j}{\partial x^i}.$$

Az $\hat{e}^i \equiv dx^i$ és $\hat{f}^j \equiv d\bar{x}^j$ azonosítással ez formailag azonos a 2.2. Tételben szereplő duális vektorbázis transzformációval. Különösen, a 2.1. Eredményt felidézve,

$$dx^i(\partial_j) = \delta_j^i.$$

2.17. Jelölés. Sima sokaságon gyakran találkozunk egy kovektor koordináta-bázis szerinti felbontásával, azaz $\hat{v} = v_i dx^i$. (A referenciapont itt implicit.)

A $\{dx^i\}$ halmazt **koordináta (duális vektor) bázisnak**, vagy **holonóm (duális vektor) bázisnak** nevezzük. Definíció szerint ez feszíti ki a sokaság **kotangens terét** a referenciapontban.

- A bázis kovektorokat felső indexekkel jelöljük.
- A kovariáns komponenseket alsó indexekkel jelöljük.

2.3. A Kronecker-tenzor

Legyen V egy n -dimenziós vektortér \mathbb{R} felett.

2.18. Jelölés. Az $\hat{f} \in V^*$ jelölésére a standard függvényjelölés helyett gyakran a $\langle \hat{f}, \cdot \rangle \in V^*$ bracket-formalizmust alkalmazzák. A függvényérték kiszámítása egy $x \in V$ vektor argumentum üres helyre való beillesztésével történik: $\langle \hat{f}, x \rangle = \hat{f}(x) \in \mathbb{R}$. Különösen, duális bázisokra

$$\langle \hat{e}^i, e_j \rangle = \hat{e}^i(e_j) = \delta_j^i$$

(lásd a 9. oldalon lévő 2.1. Eredményt).

2.19. Definíció. A $\langle \cdot, \cdot \rangle : V^* \times V \rightarrow \mathbb{R} : (\hat{f}, x) \mapsto \langle \hat{f}, x \rangle$ bilineáris leképezést **Kronecker-tenzornak** nevezzük.

2.20. Megjegyzés. A tenzor általános fogalmát a 2.8. szakaszban magyarázzuk el.

Ne keverjük össze a Kronecker-tenzort a skalárszorzáttal. A Kronecker-tenzorral ellentétben a skalárszorzat két vektor argumentumot igényel ugyanabból a vektortérből.

2.21. Megjegyzés. A Kronecker-tenzor bilineáris:

- $\forall \hat{f}, \hat{g} \in V^*, x \in V, \lambda, \mu \in \mathbb{R} : \langle \lambda \hat{f} + \mu \hat{g}, x \rangle = \lambda \langle \hat{f}, x \rangle + \mu \langle \hat{g}, x \rangle,$
- $\forall \hat{f} \in V^*, x, y \in V, \lambda, \mu \in \mathbb{R} : \langle \hat{f}, \lambda x + \mu y \rangle = \lambda \langle \hat{f}, x \rangle + \mu \langle \hat{f}, y \rangle.$
- A 9. oldalon lévő 2.1. Eredmény azt mutatja, hogy a Kronecker-tenzor komponenseit a Kronecker-szimbólum adja meg.
- Ha $\hat{f} = f_i \hat{e}^i \in V^*$ és $x = x^i e_i \in V$, akkor a Kronecker-tenzor bilinearitása maga után vonja, hogy $\langle \hat{f}, x \rangle = f_i x^j \langle \hat{e}^i, e_j \rangle = f_i x^j \delta_j^i = f_i x^i$. Így a Kronecker-tenzor és az indexkontrakció szorosan összefüggnek.

2.22. Megjegyzés. • A Kronecker-tenzor komponensei függetlenek a bázisválasztástól. (Ellenőrizze!)

- Adott V esetén a 9. oldalon tett feltételezésből következik, hogy a Kronecker-tenzor létezik.
- Mivel V^* maga is egy vektortér, tekinthetjük annak duálisát $(V^*)^* = V^{**}$. Bizonyítás nélkül kijelentjük, hogy ha V (és így V^*) véges dimenziós, akkor V^{**} izomorf V -vel,

azaz azonosítható V -vel. Ez nem általánosítható végtelen dimenziós vektorterekre.

2.23. Jelölés. A formálisan ekvivalens \hat{f} és $\langle \hat{f}, \cdot \rangle$ jelölések használata egy V^* elemre azt a hajlandóságunkat tükrözi, hogy azt vagy egy atomi egységként, vagy egy lineáris leképezésként tekintjük, meghatározatlan argumentummal. Az előző megjegyzés igazolja hasonló alternatív jelölések használatát egy V elemre, nevezetesen vagy x -ként, vagy mint $\langle \cdot, x \rangle \in V^{**} \sim V$.

2.4. Skalárszorzatok

Legyen V egy n -dimenziós vektortér \mathbb{R} felett.

2.24. Definíció. Egy valós vektortéren értelmezett **skalárszorzat** egy $(\cdot|\cdot) : V \times V \rightarrow \mathbb{R}$ alakú, nem-degenerált, pozitív definit, szimmetrikus bilineáris leképezés, amely kielégíti a következőket:

- $\forall x, y \in V : (x|y) = (y|x),$
- $\forall x, y, z \in V, \lambda, \mu \in \mathbb{R} : (x|\lambda y + \mu z) = \lambda(x|y) + \mu(x|z),$
- $\forall x \in V : (x|x) \geq 0$ és $(x|x) = 0$ akkor és csak akkor, ha $x = 0$.

Ez a 2.5. Definíció lesz az alapértelmezett. Sima sokaságok kontextusában ezt **Riemann-féle skalárszorzatnak** vagy, különösen a fizikában, ‘**Riemann-metrikának**’ nevezik. Esetenként találkozhatunk komplex vektortéren értelmezett skalárszorzattal. Az alábbiakban $\bar{z} = a - bi \in \mathbb{C}$ a $z = a + bi \in \mathbb{C}$ komplex konjugáltját jelöli.

2.25. Definíció. Egy komplex vektortéren értelmezett **skalárszorzat** egy $(\cdot|\cdot) : V \times V \rightarrow \mathbb{C}$ alakú, nem-degenerált, pozitív definit, Hermite-féle szeszkvilineáris leképezés, amely kielégíti a következőket:

- $\forall x, y \in V : (x|y) = \overline{(y|x)},$
- $\forall x, y, z \in V, \lambda, \mu \in \mathbb{C} : (x|\lambda y + \mu z) = \lambda(x|y) + \mu(x|z),$
- $\forall x \in V : (x|x) \geq 0$ és $(x|x) = 0$ akkor és csak akkor, ha $x = 0$.

2.26. Megjegyzés. • Vegyük észre, hogy a komplex esetben $(\lambda x + \mu y|z) = \bar{\lambda}(x|z) + \bar{\mu}(y|z)$, innen a ‘szeszkvilineáris’ jelző.

- A “skalárszorzat” kifejezést különböző kontextusokban használják olyan bilineáris vagy szeszkvilineáris formák jelölésére, amelyek különböző axiómarendszereket elégítenek ki. Példák erre a speciális relativitáselmélet kontextusában a Lorentz-féle skalárszorzat (v.i.) és a klasszikus mechanika Hamilton-féle formalizmusában a szimplektikus skalárszorzat.
- A Kronecker-tenzorral ellentétben a skalárszorzat létezése soha nem magától értetődő, hanem mindig explicit módon ki kell kötni.

2.27. Definíció. Egy valós V vektortéren értelmezett **pszeudo-Riemann-féle skalárszorzat** egy $(\cdot|\cdot) : V \times V \rightarrow \mathbb{R}$ alakú, nem-degenerált, szimmetrikus bilineáris leképezés, amely kielégíti a következőket:

- $\forall x, y \in V : (x|y) = (y|x),$
- $\forall x, y, z \in V, \lambda, \mu \in \mathbb{R} : (x|\lambda y + \mu z) = \lambda(x|y) + \mu(x|z),$
- $\forall x \in V, \text{ ahol } x \neq 0, \exists y \in V \text{ úgy, hogy } (x|y) \neq 0.$

A **Lorentz-féle skalárszorzat** egy pszeudo-Riemann-féle skalárszorzat $(+, -, \dots, -)$ vagy $(-, +, \dots, +)$ szignatúrával. Vegyük észre, hogy a 2.5. Definíció a 2.7. Definíció speciális esete, nevezetesen $(+, \dots, +)$ szignatúrával. A Lorentz-féle skalárszorzat szignatúrája pontosan egy $+$ vagy egy $-$ előjelet tartalmaz. A szignatúra fogalmát az alábbi 2.8. Definíció magyarázza el.

2.28. Definíció. Egy V vektorteret, amely $(\cdot|\cdot)$ skalárszorzattal van ellátva, **skalárszorzos térnek** is nevezünk.

Legyen V egy n -dimenziós skalárszorzos tér \mathbb{R} felett, $\{e_i\}$ bázissal ellátva.

2.29. Definíció. A **Gram-mátrix** G az a mátrix, amelynek komponensei $g_{ij} = (e_i|e_j)$. A **szignatúrája** (s_1, \dots, s_n) az a (bázisfüggetlen) n -es, amely a Gram-mátrix (valós) sajátértékeinek $s_i = \pm$ előjeleiből áll, $i = 1, \dots, n$, a sajátvektorok adott sorrendjéhez.

2.30. Megjegyzés. Nyilvánvaló, hogy G szimmetrikus. A definíció másik következménye, hogy G invertálható a pszeudo-Riemann-esetben. Mert tegyük fel, hogy G szinguláris lenne, akkor létezne $x = x^i e_i \in V, x \neq 0$ úgy, hogy $x^i g_{ij} = 0$. Ebből azt kapnánk, hogy minden $y = y^j e_j \in V$ esetén $0 = x^i g_{ij} y^j = x^i (e_i|e_j) y^j = (x^i e_i|y^j e_j) = (x|y)$. Ez ellentmond a 2.7. Definíció nem-degeneráltsági axiómájának.

2.31. Definíció. Az inverz Gram-mátrix G^{-1} komponenseit g^{ij} -vel jelöljük, azaz

$$g^{ik} g_{kj} = \delta_j^i.$$

Egy skalárszorzat, ha létezik, kapcsolatba hozható a Kronecker-tenzorral, amint azt alább látni fogjuk.

Legyen V egy n -dimenziós skalárszorzos tér \mathbb{R} felett.

2.32. Tétel. Létezik egy bijektív lineáris leképezés $G : V \rightarrow V^*, G^{-1} : V^* \rightarrow V$ inverzzel, úgy, hogy

- $\forall x, y \in V : (x|y) = \langle G(x), y \rangle,$
- $\forall \hat{x} \in V^*, y \in V : (G^{-1}(\hat{x})|y) = \langle \hat{x}, y \rangle.$

A G és G^{-1} mátrixreprezentációit a Gram-mátrix G és annak inverze G^{-1} adja meg, (lásd a 12. oldali 2.8. Definíciót).

Bizonyítás. (A 2.3. Tétel bizonyítása.) • Rögzített $x \in V$ esetén definiáljuk $\hat{x} \in V^*$ -t mint az $\hat{x} : V \rightarrow \mathbb{R} : y \mapsto (x|y)$ lineáris leképezést. Ezt követően definiáljuk a $G : V \rightarrow V^* : x \mapsto G(x) \equiv \hat{x}$ lineáris leképezést.

- A második részhez be kell bizonyítanunk G invertálhatóságát. Tegyük fel, hogy $x \in V, x \neq 0$ olyan, hogy $G(x) = 0$. Ekkor $0 = \langle G(x), y \rangle = (x|y)$ minden $y \in V$ esetén, ami ellentmond a skalárszorzat nem-degeneráltsági tulajdonságának. Tehát G egy-az-egyben^a. Mivel $\dim V = n < \infty$, ezért bijektív is, tehát invertálható.
- A G mátrixrepresentációjának kiszámításához bontsuk fel $x = x^i e_i, y = y^j e_j$ és $G(x) = G(x^i e_i) = x^i G_{ij} \hat{e}^j$ alakban, ahol a mátrixkomponenseket a $G(e_i) = G_{ij} \hat{e}^j$ útján definiáljuk.
 - A konstrukció (fenti első pont), a skalárszorzat bilinearitása és a Gram-mátrix definíciója révén $\langle G(x), y \rangle = (x|y) = x^i y^j (e_i|e_j) = g_{ij} x^i y^j$.
 - A dualitás definíciója szerint $\langle G(x), y \rangle = G_{ij} x^i y^k \langle \hat{e}^j, e_k \rangle = G_{ij} x^i y^k \delta_k^j = G_{ij} x^i y^j$.

Mivel $x^i, y^j \in \mathbb{R}$ tetszőlegesek, következik, hogy $G_{ij} = g_{ij}$ (lásd 2.8. Definíció).

- Következésképpen az inverz $G^{-1} : V^* \rightarrow V$ leképezésnek egy G^{-1} mátrixrepresentációja van, amely megfelel az inverz Gram-mátrixnak: $(G^{-1})^{ij} = g^{ij}$.

^aAz “egy-az-egyben” és az “injektív” kifejezések szinonimák. Egy $f : D \rightarrow C$ függvény injektív, ha különböző argumentumokhoz különböző értékeket rendel. Ha a függvény C képhalmazát a tényleges $f(D)$ értékkészletére cseréljük, akkor szürjektív is, tehát bijektív.

2.33. Megjegyzés. • Vegyük észre a szinonim megfogalmazásokat: $G(x) \equiv \hat{x} \equiv \langle G(x), \cdot \rangle \equiv (x|\cdot)$. Idézzük fel a jelölésről szóló megjegyzést a 11. oldalon. Az alábbi 2.9. Definíció egy további szinonim formát vezet be.

- A Riesz-féle reprezentációs tétel, amely kimondja, hogy minden $\hat{x} \in V^*$ elemhez létezik egy és csak egy $x \in V$ úgy, hogy $\langle \hat{x}, y \rangle = (x|y)$ minden $y \in V$ esetén, a 2.3. Tétel közvetlen következménye, nevezetesen $x = G^{-1}(\hat{x})$.

2.34. Jelölés. • Félkövér szimbólumokat tartunk fenn az intrinzik mennyiségeknek, mint például vektorok, kovektorok és (más) lineáris leképezések, hogy megkülönböztessük őket a referenciabázishoz viszonyított komponenseiktől.

2.35. Definíció. A $\sharp : V \rightarrow V^*$ és $\flat : V^* \rightarrow V$ **konverziós operátorokat** a G , illetve G^{-1} rövidítéseként vezetjük be, nevezetesen, ha $v \in V, \hat{v} \in V^*$, akkor

$$\sharp v := G(v) \in V^* \quad \text{illetve} \quad \flat \hat{v} := G^{-1}(\hat{v}) \in V.$$

A G -t általában **metrikának**, a G^{-1} -t pedig **duális metrikának** nevezik. A $\flat : V^* \rightarrow V$ és $\sharp : V \rightarrow V^*$ operátorokat **sharp**-nak, illetve **flat**-nek ejtik.

Legyen $v = v^i e_i \in V$ és $\hat{v} = v_i \hat{e}^i \in V^*$, ekkor

- $\flat \hat{v} = g_{ij} v^i \hat{e}^j = v_j \hat{e}^j = \hat{v} \in V^*$, ahol $v_j = g_{ij} v^i$;

- $[\hat{v} = g^{ij}v_i e_j = v^j e_j = v \in V]$, ahol $v^j = g^{ij}v_i$.

Bizonyítás. (A 2.2. Eredmény bizonyítása.) Ez közvetlenül következik a $]$ és $[$ definíciójából, felhasználva a következőket:

- $]e_i = G(e_i) = g_{ij}\hat{e}^j \in V^*$, illetve
- $[\hat{e}^i = G^{-1}(\hat{e}^i) = g^{ij}e_j \in V$.

2.36. Jelölés. Amennyiben másként nem jelezzük, a következő konvenciót követjük. Ha $v^i \in \mathbb{R}$, akkor $v_j := g_{ij}v^i$. Fordítva, ha $v_i \in \mathbb{R}$, akkor $v^j := g^{ij}v_i$.

Az előzőekben $G \in \mathcal{L}(V, V^*)$ és $G^{-1} \in \mathcal{L}(V^*, V)$, a $G : V \rightarrow V^*$ és $G^{-1} : V^* \rightarrow V$ prototípusokkal. Azonban gyakran a metrikát és duálisát $G : V \times V \rightarrow \mathbb{R}$, illetve $G^{-1} : V^* \times V^* \rightarrow \mathbb{R}$ típusú bilineáris leképezésként értelmezik. A kontextusból egyértelműnek kell lennie, hogy melyik prototípusra utalunk. Erre a kétértelműsége a 2.8. szakaszban térünk vissza.

Egy skalárszorzatos térben a vektorok “álrúhas kovektoroknak” tekinthetők, és fordítva.

2.5. Reciprok bázisok

Legyen V egy n -dimenziós skalárszorzatos tér \mathbb{R} felett, $\{e_i\}$ bázissal ellátva.

2.37. Definíció. A V tér $\{e_i\}$ bázisa indukál egy új, $\{[\hat{e}^i]\}$ bázist V -ben, melyet $[\hat{e}^i] = g^{ij}e_j$ ad meg, az úgynevezett **reciprok bázist**.

2.38. Definíció. A V^* tér $\{\hat{e}^i\}$ bázisa indukál egy új, $\{e_i\}$ bázist V^* -ban, melyet $]e_i = g_{ij}\hat{e}^j$ ad meg, az úgynevezett **reciprok duális bázist**.

- $[\hat{e}^i]$ helyett néha e^i -t írnak az e_i reciprokára.
- Hasonlóképpen, e_i helyett néha \hat{e}_i -t írnak az \hat{e}^i reciprokára.
- Mivel konvencionálisan a V elemeit alsó indexekkel, a V^* elemeit pedig felső indexekkel címkézzük, ragaszkodni fogunk a $]$ és $[$ operátorokat alkalmazó jelöléshez.

2.39. Megjegyzés. • A reciprok bázist egyértelműen meghatározza a skalárszorzat.

- A reciprok bázis Gram-mátrixa G^{-1} , komponensei

$$([\hat{e}^i][\hat{e}^j]) = g^{ik}g^{j\ell}(e_k|e_\ell) = g^{ik}g^{j\ell}g_{k\ell} = g^{ij}.$$

- Továbbá

$$([\hat{e}^i]|e_j) = g^{ik}(e_k|e_j) = g^{ik}g_{kj} = \delta_j^i.$$

2.40. Definíció. Azt mondjuk, hogy a $[\hat{e}^i]$ és e_j vektorok **merőlegese**k, ha $i \neq j$.

2.41. Megjegyzés. A V tér minden $\{\hat{e}^i\}$ reciprok bázisához és minden $x \in V$ elemhez létezik egy egyértelmű $\{x_i\}$ valós számhalmaz úgy, hogy $x = x_i \hat{e}^i$.

2.42. Definíció. Az x_i valós számokat az x vektor $\{e_i\}$ bázisra vonatkozó **kovariáns komponenseinek** nevezzük.

2.6. Bázisok, Duális Bázisok, Reciprok Bázisok

2.43. Jelölés. Nem valószínű, hogy zavar keletkezik, ha azonos szimbólumokat, $\hat{x}_i \equiv x_i$, használunk a duális és a reciprok vektorkomponensek jelölésére. Az alapul szolgáló bázist ($\{\hat{e}^i\}$, illetve $\{e_i\}$) és így a helyes geometriai értelmezést a kontextusból kell kikövetkeztetni. Alapértelmezés szerint az x_i -t duális vektorkomponensekként fogjuk értelmezni, a továbbiakban elhagyva a $\hat{}$ szimbólumot, azaz $\hat{x} = x_i \hat{e}^i \in V^*$.

2.7. Tenzorok

2.7.1. Tenzorok általánosságban

2.44. Definíció. A **tenzor** egy $T : \underbrace{V^* \times \cdots \times V^*}_p \times \underbrace{V \times \cdots \times V}_q \rightarrow \mathbb{R}$ típusú multilineáris leképezés, ahol $p, q \in \mathbb{N}_0$. Az ilyen típusú tenzorokat együttesen a következőképpen jelöljük: $T \in \underbrace{V \otimes \cdots \otimes V}_p \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_q := \mathcal{T}_q^p(V)$.

Bizonyítás. (*Eredmény bizonyítása.*) Az első eset egy korábban kifejtett konvenció. A második eset a 8. oldalon található 2.2. Definíció alapján igaz, felidézve a 9. oldalon bevezetett terminológiát. A harmadik eset a $V^{**} = V$ izomorfizmus révén igaz, felidézve a 11. oldalon tett megjegyzést.

2.45. Definíció. Két \mathbb{R} -értékű függvény, $f : X \rightarrow \mathbb{R}$ és $g : Y \rightarrow \mathbb{R}$, **külső szorzata** ($f \otimes g : X \times Y \rightarrow \mathbb{R}$) a következőképpen definiálható: $(f \otimes g)(x, y) = f(x)g(y)$ minden $x \in X, y \in Y$ esetén.

2.46. Definíció. Mivel a vektorok és kovektorok \mathbb{R} -értékű leképezések, a következő definíció természetes. A kevert $\binom{p}{q}$ -tenzor, $e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \in \mathcal{T}_q^p(V)$, minden $(i_1, \dots, i_p, j_1, \dots, j_q)$ index $(p+q)$ -esre a következőképpen van definiálva:

$$(e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q})(\hat{x}_1, \dots, \hat{x}_p, y_1, \dots, y_q) = \langle \hat{x}_1, e_{i_1} \rangle \cdots \langle \hat{x}_p, e_{i_p} \rangle \langle \hat{e}^{j_1}, y_1 \rangle \cdots \langle \hat{e}^{j_q}, y_q \rangle,$$

minden $\hat{x}_1, \dots, \hat{x}_p \in V^*$ és $y_1, \dots, y_q \in V$ esetén.

Bizonyítás. (A 2.4. Eredmény bizonyítása.) Adott:

$$\begin{aligned} T(\hat{e}^{i_1}, \dots, \hat{e}^{i_p}, e_{j_1}, \dots, e_{j_q}) &= t_{l_1 \dots l_q}^{k_1 \dots k_p} (e_{k_1} \otimes \dots \otimes \hat{e}^{l_q})(\hat{e}^{i_1}, \dots, e_{j_q}) \\ &= t_{l_1 \dots l_q}^{k_1 \dots k_p} \langle \hat{e}^{i_1}, e_{k_1} \rangle \dots \langle \hat{e}^{l_q}, e_{j_q} \rangle \\ &= t_{l_1 \dots l_q}^{k_1 \dots k_p} \delta_{k_1}^{i_1} \dots \delta_{k_p}^{i_p} \delta_{j_1}^{l_1} \dots \delta_{j_q}^{l_q} = t_{j_1 \dots j_q}^{i_1 \dots i_p}. \end{aligned}$$

A multilinearitás miatt általános kovektor és vektor argumentumokra, $\hat{\omega}_1, \dots, \hat{\omega}_p \in V^*$ és $v_1, \dots, v_q \in V$, érvényes, hogy $T(\hat{\omega}_1, \dots, v_q) = t_{j_1 \dots j_q}^{i_1 \dots i_p} \omega_{1, i_1} \dots \omega_{p, i_p} v_1^{j_1} \dots v_q^{j_q}$.

2.47. Jelölés. Kivéve, ha ez zavart okozhatna, a továbbiakban nem teszünk jelölésbeli különbséget az olyan különböző prototípusok között, mint a g és a G . Mindig G -t fogunk írni, akár kovariáns 2-tenzorként (g a példában, amit a $G(v, w)$ írásmód sugall), akár vektor-kovektor átalakítóként (\lrcorner -operátor, amit a $G(v)$ írásmód sugall) értendő. A \lrcorner és \lrcorner -operátorok kényelmes rövidítéseket biztosítanak a (duális) metrikára, ha hangsúlyozni kívánjuk annak vektor-kovektor (illetve kovektor-vektor) átalakítóként való működését.

2.48. Tétel. Ha $S \in \mathcal{T}_q^p(V)$, $T \in \mathcal{T}_s^r(V)$, ahol $S = s_{j_1 \dots j_q}^{i_1 \dots i_p} e_{i_1} \otimes \dots \otimes \hat{e}^{j_q}$ és $T = t_{j_1 \dots j_s}^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes \hat{e}^{j_s}$, akkor $S \otimes T \in \mathcal{T}_{q+s}^{p+r}(V)$, ahol $S \otimes T = s_{j_1 \dots j_q}^{i_1 \dots i_p} t_{j_{q+1} \dots j_{q+s}}^{i_{p+1} \dots i_{p+r}} e_{i_1} \otimes \dots \otimes e_{i_{p+r}} \otimes \hat{e}^{j_1} \otimes \dots \otimes \hat{e}^{j_{q+s}}$.

Bizonyítás. (Tétel bizonyítása.) Alkalmazzuk a 19. oldalon található 2.13. Definíciót, figyelembe véve a multiplikatív együttthatók megjelenését: $((\lambda f) \otimes (\mu g))(x, y) = ((\lambda f)(x))((\mu g)(y)) = \lambda \mu f(x)g(y)$. A további részletek az olvasóra vannak bízva (figyelembe véve az előző figyelmeztetést).

2.7.2. Szimmetrikus tenzorok

2.49. Megjegyzés. Idézzük fel a 8. oldalon tett korábbi megjegyzést az indexek elhelyezésének konvenciójáról. Különösen az alábbi 2.15. és 2.16. Definíciókban az indexek egy eleve korlátlan számú argumentum címkézésére szolgálnak, ezért a felső/alsó index konvenció nem alkalmazható.

2.50. Jelölés. $\pi = \{\pi(\alpha_1), \dots, \pi(\alpha_p)\}$ a p darab rendezett szimbólum, $\{\alpha_1, \dots, \alpha_p\}$, egy permutációját jelöli. $\pi(\alpha_i)$, $i = 1, \dots, p$, a π i -edik elemének rövidített jelölése.

2.51. Definíció. Legyen $\pi = \{\pi(1), \dots, \pi(p)\}$ a $\{1, \dots, p\}$ címkék egy tetszőleges permutációja, akkor $T \in \mathcal{T}_p^0(V)$ **szimmetrikus kovariáns tenzor**, ha minden $v_1, \dots, v_p \in V$ esetén $T(v_{\pi(1)}, \dots, v_{\pi(p)}) = T(v_1, \dots, v_p)$. Hasonlóképpen, $T \in \mathcal{T}_0^p(V)$ **szimmetrikus kontravariáns tenzor**, ha minden $\hat{v}_1, \dots, \hat{v}_p \in V^*$ esetén $T(\hat{v}_{\pi(1)}, \dots, \hat{v}_{\pi(p)}) = T(\hat{v}_1, \dots, \hat{v}_p)$.

2.52. Jelölés. A szimmetrikus kovariáns p -tenzorok vektorterét $\mathcal{W}^p(V) \subset \mathcal{T}_p^0(V)$ jelöli. A szimmetrikus kontravariáns p -tenzorok vektorterét $\mathcal{W}_p(V) \subset \mathcal{T}_0^p(V)$ jelöli. Alternatíva-

ként néha $\mathcal{W}^p(V) = \underbrace{V^* \otimes_S \cdots \otimes_S V^*}_p$ és $\mathcal{W}_p(V) = \underbrace{V \otimes_S \cdots \otimes_S V}_p$ írásmódot használnak.

2.53. Definíció. Jelölje $\text{sgn}(\pi)$ a π permutáció előjelét, azaz $\text{sgn}(\pi) = 1$, ha páros, és $\text{sgn}(\pi) = -1$, ha páratlan. A $T \in \mathcal{T}_p^0(V)$ tenzort ****antiszimmetrikus kovariáns tenzornak**** nevezzük, ha minden $v_1, \dots, v_p \in V$ és a $\{1, \dots, p\}$ címkék bármely $\pi = \{\pi(1), \dots, \pi(p)\}$ permutációja esetén $T(v_{\pi(1)}, \dots, v_{\pi(p)}) = \text{sgn}(\pi)T(v_1, \dots, v_p)$. Hasonlóképpen, a $T \in \mathcal{T}_0^p(V)$ tenzort ****antiszimmetrikus kontravariáns tenzornak**** nevezzük, ha minden $\hat{v}_1, \dots, \hat{v}_p \in V^*$ esetén $T(\hat{v}_{\pi(1)}, \dots, \hat{v}_{\pi(p)}) = \text{sgn}(\pi)T(\hat{v}_1, \dots, \hat{v}_p)$.

2.54. Jelölés. Az antiszimmetrikus kovariáns p-tenzorok vektorterét $\Lambda^p(V) \subset \mathcal{T}_p^0(V)$ jelöli. Az antiszimmetrikus kontravariáns p-tenzorok vektorterét $\Lambda_p(V) \subset \mathcal{T}_0^p(V)$ jelöli. Alternatívaként néha $\Lambda^p(V) = \underbrace{V^* \otimes_A \cdots \otimes_A V^*}_q$ és $\Lambda_p(V) = \underbrace{V \otimes_A \cdots \otimes_A V}_p$ írásmódot használnak.

2.8.3 Szimmetriát és antiszimmetriát megőrző szorzási operátorok

2.55. Definíció. Minden $\hat{v}_1, \dots, \hat{v}_k \in V^*$ esetén az antiszimmetrikus kovariáns k-tenzort, $\hat{v}_1 \wedge \cdots \wedge \hat{v}_k \in \Lambda^k(V)$, a következőképpen definiáljuk: $(\hat{v}_1 \wedge \cdots \wedge \hat{v}_k)(x_1, \dots, x_k) = \det\langle \hat{v}_i, x_j \rangle$ minden $x_j \in V$ esetén.

2.56. Definíció. A **perm** operátor egy négyzetes mátrixhoz rendel egy számot a \det operátorhoz hasonló módon, de minden tagot $+$ előjellel vesz figyelembe:

$$\text{perm} A = \sum_{j_1, \dots, j_n=1}^n |[j_1, \dots, j_n]| A_{1j_1} \cdots A_{nj_n}$$

2.57. Megjegyzés. A fenti kifejezésekben a teljesen antiszimmetrikus szimbólumok formális szerepe az összegzés releváns tagokra való korlátozása. Az abszolút érték elhagyja a páratlan paritású tagok mínusz előjelét.

2.58. Definíció. Minden $\hat{v}_1, \dots, \hat{v}_k \in V^*$ k-as esetén a szimmetrikus kovariáns k-tenzort, $\hat{v}_1 \vee \cdots \vee \hat{v}_k \in \mathcal{W}^k(V)$, a következőképpen definiáljuk: $(\hat{v}_1 \vee \cdots \vee \hat{v}_k)(x_1, \dots, x_k) = \text{perm} \langle \hat{v}_i, x_j \rangle$.

2.59. Tétel. Vö. a fenti észrevételekkel. Fennáll, hogy

$$\dim \Lambda^p(V) = \binom{n}{p} \quad \text{és} \quad \dim \mathcal{W}^p(V) = \binom{n+p-1}{p}.$$

Bizonyítás. (A 2.5. Tétel bizonyítása.) • Az első azonosságot illetően, egy $\hat{e}_{i_1} \wedge \cdots \wedge \hat{e}_{i_p}$ alakú szorzatban a címkék szigorúan növekvő sorrendje pontosan annyi különböző módon valósítható meg, ahányféleképpen p elemet ki lehet választani n különböző objektumból.

- A második azonosságot illetően, az a módok száma, ahogyan egy $\hat{e}_{i_1} \vee \dots \vee \hat{e}_{i_p}$ alakú szorzatban a címkék nem csökkenő sorrendje megvalósítható, a következő érvelésből vezethető le. Bármely megengedett címkesorrend szimbolikusan ábrázolható p darab pont (\bullet) sorozatával, amelyeket $n - 1$ darab kereszt (\times) egyenlő címkéjű részsorozatokra oszt, ahol mindegyik kereszt egy egységnyi címkeátmenetet jelöl. A lehetséges permutációk száma $\binom{n+p-1}{p}$.

2.60. Megjegyzés. Vegyük észre a kvalitatív különbségeket $\Lambda^p(V)$ és $\mathcal{W}^p(V)$ között a dimenzió p -vel való skálázódását illetően. Az előbbi esetben $\dim \Lambda^p(V) = 0$, amint $p > n = \dim V$, azaz nem léteznek nem triviális antiszimmetrikus tenzorok, amelyek rangja meghaladja a vektortér dimenzióját. Ezzel szemben bármilyen rangú nem triviális szimmetrikus tenzorok léteznek.

2.61. Definíció. Az $S : \mathcal{T}_p^0(V) \rightarrow \mathcal{W}^p(V)$ lineáris operátort, melyet a

$$(S(T))(v_1, \dots, v_p) = \frac{1}{p!} \sum_{\pi} T(v_{\pi(1)}, \dots, v_{\pi(p)})$$

kifejezés ad meg, ahol az összegzés a $\{1, \dots, p\}$ indexhalmaz összes π permutációjára kiterjed, **szimmetrizáló leképezésnek** nevezzük.

Bizonyítás. (*Eredmény bizonyítása.*) Legyen $T = t_{i_1 \dots i_p} \hat{e}^{i_1} \otimes \dots \otimes \hat{e}^{i_p}$ és $U = u_{i_{p+1} \dots i_{p+q}} \hat{e}^{i_{p+1}} \otimes \dots \otimes \hat{e}^{i_{p+q}}$, szimmetrikus $t_{i_1 \dots i_p}$ és $u_{i_{p+1} \dots i_{p+q}}$ holorokkal. Ekkor $T \vee U = t_{i_1 \dots i_p} u_{i_{p+1} \dots i_{p+q}} (\hat{e}^{i_1} \otimes \dots) \vee (\hat{e}^{i_{p+1}} \otimes \dots) = \frac{(p+q)!}{p!q!} S(T \otimes U)$.

2.62. Definíció. Az $A : \mathcal{T}_p^0(V) \rightarrow \Lambda^p(V)$ lineáris operátort, melyet a

$$(A(T))(v_1, \dots, v_p) = \frac{1}{p!} \sum_{\pi} \text{sgn}(\pi) T(v_{\pi(1)}, \dots, v_{\pi(p)})$$

kifejezés ad meg, **antiszimmetrizáló leképezésnek** nevezzük.

2.63. Jelölés. Az Einstein-féle összegzési konvenciót, (lásd 7. oldal), általában az antiszimmetrikus tenzorok kontextusában adaptálják. Általában, ha $T \in \Lambda^p(V)$, akkor a következő ekvivalens felbontások bármelyikét használhatjuk:

- $T = t_{i_1 \dots i_p} \hat{e}^{i_1} \otimes \dots \otimes \hat{e}^{i_p} = \sum_{i_1, \dots, i_p} t_{i_1 \dots i_p} \hat{e}^{i_1} \otimes \dots \otimes \hat{e}^{i_p}$,
- $T = t_{|i_1 \dots i_p|} \hat{e}^{i_1} \wedge \dots \wedge \hat{e}^{i_p} = \sum_{i_1 < \dots < i_p} t_{|i_1 \dots i_p|} \hat{e}^{i_1} \wedge \dots \wedge \hat{e}^{i_p}$.

2.64. Megjegyzés. • Mind az S , mind az A idempotens: $S \circ S = S$, illetve $A \circ A = A$. Tehát definíció szerint projekciók $\mathcal{T}_p^0(V)$ -ről $\mathcal{W}^p(V)$ -re, illetve $\Lambda^p(V)$ -re.

- Vegyük észre, hogy a (anti)szimmetria tulajdonságát nem definiáltuk valódi vegyes tenzorra.
- Egy (mondjuk kovariáns) 2-tenzor mindig felbontható egy szimmetrikus és egy antiszimmetrikus 2-tenzor összegére.

- A metrikus tenzor G szimmetrikus, és annak inverze, a duális metrikus tenzor G^{-1} is az.

2.7.3. Vektorterek orientált térfogattal

2.65. Definíció. Legyen (V, μ) a V vektortér, melyet a $\mu \in \Lambda^n(V)$ n -formával láttunk el, és legyen $\{e_i\}$ a V egy adott bázisa.

- (V, μ) -t **orientált térfogattal rendelkező vektortérnek** nevezzük.
- (V, μ) **pozitív orientációjú**, ha $\mu(e_1, \dots, e_n) > 0$.
- (V, μ) **negatív orientációjú**, ha $\mu(e_1, \dots, e_n) < 0$.

2.66. Jelölés. (V, μ) helyett egyszerűen V -t fogunk írni. Az $\{e_i\}$ bázis, amelyre $\mu(e_1, \dots, e_n) = 1$ teljesül, a kontextusból egyértelmű kell, hogy legyen.

2.67. Definíció. Legyen $a_1, \dots, a_k \in V$, $k \in \{0, \dots, n\}$, ekkor a $(\mu \rfloor a_1 \rfloor \dots \rfloor a_k) \in \Lambda^{n-k}(V)$ $(n-k)$ -formát a következőképpen definiáljuk:

$$(\mu \rfloor a_1 \rfloor \dots \rfloor a_k)(x_{k+1}, \dots, x_n) = \mu(a_1, \dots, a_k, x_{k+1}, \dots, x_n).$$

2.68. Megjegyzés. • Egy V vektortéren értelmezett orientált térfogat $\mu \in \Lambda^n(V)$ nem igényel skalárszorzatot.

- Egy orientált térfogat $\mu \in \Lambda^n(V)$ a duális tér V^* létezése révén létezik.
- Ha $\mu(e_1, \dots, e_n) = 1$, akkor $\mu = \hat{e}^1 \wedge \dots \wedge \hat{e}^n$ (magával μ -vel ellentétben a tényezők nem egyediek).

2.8. Tenzorok skalárszorzatot téren

2.69. Definíció. Idézzük fel a 14. oldalon lévő 2.2. Eredményt és a 18. oldalon lévő 2.12. Definíciót. Bármely $p, q \in \mathbb{N}_0$ esetén a $T \in \mathcal{T}_q^p(V)$ tenzor **kovariáns reprezentációját**, $\rfloor T \in \mathcal{T}_{p+q}^0(V)$, a következőképpen definiáljuk:

$$\rfloor T(v_1, \dots, v_{p+q}) = T(\rfloor v_1, \dots, \rfloor v_p, v_{p+1}, \dots, v_{p+q}) \quad \text{minden } v_1, \dots, v_{p+q} \in V \text{ esetén.}$$

Hasonlóképpen, a $T \in \mathcal{T}_q^p(V)$ tenzor **kontravariáns reprezentációját**, $[T \in \mathcal{T}_0^{p+q}(V)$, a következőképpen definiáljuk:

$$[T(\hat{v}_1, \dots, \hat{v}_{p+q}) = T(\hat{v}_1, \dots, \hat{v}_p, [\hat{v}_{p+1}, \dots, [\hat{v}_{p+q}) \quad \text{minden } \hat{v}_1, \dots, \hat{v}_{p+q} \in V^* \text{ esetén.}$$

Általánosságban egy $T \in \mathcal{T}_q^p(V)$ tenzor **vegyes reprezentációját** úgy kapjuk, hogy a tenzor definíciós tartományában lévő V és V^* (duális) vektorterek egy tetszőleges részhalmozatát a metrika által indukált $\{\rfloor, [\}$ -mechanizmussal hasonló módon „átkapcsoljuk”.

2.70. Jelölés.

- A jelöléssel való visszaélés révén általában ugyanazt a szimbólumot, mondjuk T -t, használjuk a fent tárgyalt (kovariáns, kontravariáns vagy vegyes) prototípusok bármelyikének jelölésére.
- Ugyanezen elv alapján a $\mathcal{T}_k(V) = \bigcup_{p+q=k} \mathcal{T}_q^p(V)$ tér együttesen utal az összes k totális rangú tenzorra.
- Ha egy adott $T \in \mathcal{T}_q^p(V)$ prototípusra gondolunk, annak a kontextusból egyértelműnek kell lennie.
- A potenciálisan zavaró $t_{j_1 \dots j_s}^{i_1 \dots i_r}$ jelölés helyett egy vegyes holorra gyakran $t_{j_1 \dots j_s}^{i_1 \dots i_r}$ -t írunk, azaz olyat, amelyben az indexek sorrendje egyértelmű.

2.8.6 Tenzor transzformációk

2.8.6.1 „Abszolút tenzorok”

Bizonyítás. (A 2.8. Eredmény bizonyítása.) Tegyük fel, hogy $\hat{f}^i = C_j^i \hat{e}^j$. Ekkor a dualitás definíciója szerint

$$\delta_j^i := \langle \hat{f}^i, f_j \rangle = \langle C_k^i \hat{e}^k, A_j^l e_l \rangle^* = C_k^i A_j^l \langle \hat{e}^k, e_l \rangle := C_k^i A_j^l \delta_l^k = C_k^i A_j^k,$$

amiből $C = A^{-1} = B$. A $*$ -gal jelölt lépésben a Kronecker-tenzor bilinearitását használtuk fel. Következésképpen, $e_i = B_i^l f_l$ és $\hat{e}^j = A_k^j \hat{f}^k$. Ezeket behelyettesítve a T bázisfelbontásába, megkapjuk a megfelelő holorokra vonatkozó állítást.

2.71. Megjegyzés. A tenzor transzformációs törvény homogén természete azt jelenti, hogy egy $t_{j_1 \dots j_q}^{i_1 \dots i_p} = 0$ alakú holor-egyenlet bármely bázisra vonatkozóan fennáll, ha egy adott bázisra fennáll.

2.8.6.2 „Relatív tenzorok”

- 2.72. Megjegyzés.**
- A 39. oldalon található 2.1. Lemmában a $\mu_{i_1 \dots i_n} = [i_1, \dots, i_n]$ holor az abszolút tenzor transzformációs törvény szerint transzformálódik, (lásd a 37. oldalon lévő 2.8. Eredményt), ami azt jelenti, hogy $\bar{\mu}_{i_1 \dots i_n} = \det(A) \mu_{i_1 \dots i_n} \neq \mu_{i_1 \dots i_n}$.
 - A 2.9. Eredményben a $\mu_{i_1 \dots i_n} = [i_1, \dots, i_n]$ holor a relatív tenzor transzformációs törvény szerint transzformálódik, ami a $\mu_{i_1 \dots i_n} = [i_1, \dots, i_n]$ holort invariánssá teszi abban az értelemben, hogy $\bar{\mu}_{i_1 \dots i_n} = \mu_{i_1 \dots i_n} = [i_1, \dots, i_n]$ a bázistól függetlenül.
 - A relatív tenzor transzformációs törvény elfogadása lehetővé teszi, hogy a $\mu_{i_1 \dots i_n}$ -t a permutációs szimbólum $[i_1, \dots, i_n]$ 'kovariáns álneveként' értelmezzük, amely definíció szerint nem függ a bázistól.

2.73. Jelölés. A következő konvenciót alkalmazzuk a továbbiakban az Einstein-féle összegzési konvenció következetes használatának támogatására és a $\mu_{i_1 \dots i_n}$ jelentésével kapcsolatos kétértelműség feloldására:

- A továbbiakban $\mu_{i_1 \dots i_n}$ -t és $\mu^{i_1 \dots i_n}$ -t a permutációs szimbólum kovariáns és kontravariáns álneveként fogadjuk el, azaz $\mu_{i_1 \dots i_n} \equiv \mu^{i_1 \dots i_n} \equiv [i_1, \dots, i_n]$.
- Következésképpen a $\mu_{i_1 \dots i_n}$ holor a relatív tenzor transzformációs törvénynek, a 2.9. Eredménynek van alávetve.

2.8.1. „Pseudotenzorok”

2.74. Jelölés. Rövidség kedvéért a Gram-mátrix, $g_{ij} = (e_i | e_j)$, determinánsát $g = \det(g_{ij})$ jelöli.

2.75. Definíció. A **Levi-Civita-tenzor** az egyedi, pozitív orientációjú, egységnyi n-forma:

$$\epsilon = \sqrt{g} \mu = \sqrt{g} \hat{e}^1 \wedge \dots \wedge \hat{e}^n = \epsilon_{|i_1 \dots i_n|} \hat{e}^{i_1} \wedge \dots \wedge \hat{e}^{i_n}.$$

2.76. Tétel. A Levi-Civita-tenzor kontravariáns reprezentációja, $[\epsilon \in \Lambda_n(V)]$, (lásd a 35. oldali 2.24. Definíciót), a következő: $[\epsilon = \epsilon^{i_1 \dots i_n} e_{i_1} \wedge \dots \wedge e_{i_n}]$, a holor reprezentációja pedig

$$\epsilon^{i_1 \dots i_n} = \frac{1}{\sqrt{g}} [i_1, \dots, i_n].$$

Bizonyítás. (Tétel bizonyítása.) Elegendő a holor reprezentáció érvényességét indexemeléssel igazolni:

$$\epsilon^{i_1 \dots i_n} = g^{i_1 j_1} \dots g^{i_n j_n} \epsilon_{j_1 \dots j_n} = g^{i_1 j_1} \dots g^{i_n j_n} \sqrt{g} [j_1, \dots, j_n] = \frac{1}{\sqrt{g}} [i_1, \dots, i_n].$$

Az utolsó lépésben felhasználtuk azt a tényt, hogy $g^{i_1 j_1} \dots g^{i_n j_n} [j_1, \dots, j_n] = \det(g^{ij}) [i_1, \dots, i_n]$, valamint azt, hogy $\det(g^{ij}) = 1 / \det(g_{ij})$.

Bizonyítás. (Tétel bizonyítása.) A g_{ij} transzformációs tulajdonságai révén (lásd 2.8. Eredmény) kapjuk, hogy $\bar{g}_{ij} = A_i^k A_j^l g_{kl}$, amiből $\sqrt{\bar{g}} = |\det(A)| \sqrt{g}$. A pseudotenzor transzformációs törvény definíciója szerint, felhasználva egy korábbi megfigyelést, $\bar{\epsilon}_{i_1 \dots i_n} := \text{sgn}(\det A) A_{i_1}^{j_1} \dots A_{i_n}^{j_n} \epsilon_{j_1 \dots j_n} = \text{sgn}(\det A) \det(A) \sqrt{g} \mu_{i_1 \dots i_n} = |\det A| \sqrt{g} \mu_{i_1 \dots i_n} = \sqrt{\bar{g}} \mu_{i_1 \dots i_n}$.

2.77. Megjegyzés. • A Levi-Civita-tenzort **pseudotenzornak** nevezik, mert a forma-invariáns holorja (!) a 2.10. Eredmény pseudotenzor transzformációs törvénye szerint transzformálódik, nem pedig a 2.8. Eredmény abszolút tenzor transzformációs törvénye szerint.

- A \sqrt{g} tényező révén elért formai invarianciát nem szabad összekeverni a numerikus invarianciával.
- Nem pozitív definit skalárszorzat esetén (lásd 2.7. Definíció) enyhe módosításokra lehet szükség, tipikusan a \sqrt{g} tényező helyettesítése $\sqrt{|g|}$ -vel, ha g negatív.

2.78. Definíció. Az általánosított Kronecker- vagy permutációs szimbólum definíciója:

$$\begin{pmatrix} \delta_{j_1}^{i_1} & \cdots & \delta_{j_k}^{i_1} \\ \vdots & \ddots & \vdots \\ \delta_{j_1}^{i_k} & \cdots & \delta_{j_k}^{i_k} \end{pmatrix} = \begin{cases} +1 & \text{ha } (i_1, \dots, i_k) \text{ a } (j_1, \dots, j_k) \text{ páros permutációja,} \\ -1 & \text{ha } (i_1, \dots, i_k) \text{ a } (j_1, \dots, j_k) \text{ páratlan permutációja,} \\ 0 & \text{egyébként.} \end{cases} \quad (12)$$

2.8.2. Kontrakciók

2.79. Definíció. Legyen $T \in \mathcal{T}_1^1(V)$. A T **nyoma**, $\mathbb{T} \setminus T \in \mathbb{R}$, a következőképpen van definiálva:

$$\mathbb{T} \setminus T = T(\hat{e}^i, e_i).$$

2.80. Megjegyzés. • A $\mathbb{T} \setminus$ -operátor bármely $T \in \mathcal{T}_q^p(V)$ vegyes tenzorra alkalmazható, feltéve, hogy $p, q \geq 1$, nevezetesen $p - 1$ kovektor és $q - 1$ vektor argumentum tetszőleges rögzítésével.

- „Befagyasztott” argumentumok nélkül a $\mathbb{T} \setminus T$ -t vegyes tenzorként, $\mathbb{T} \setminus T \in \mathcal{T}_{q-1}^{p-1}(V)$, értelmezhetjük.
- A $\mathbb{T} \setminus$ -operátor a kovariáns és a kontravariáns rangot is eggyel csökkenti.

2.9. A Hodge-féle csillag operátor

2.81. Definíció. Legyenek $a_1, \dots, a_k \in V$. A **Hodge-féle csillag operátor** $*$: $\Lambda^k(V) \rightarrow \Lambda^{n-k}(V)$ a következőképpen van definiálva:

$$\begin{cases} *1 = \epsilon & \text{ha } k = 0, \\ *([a_1 \wedge \cdots \wedge a_k]) = \epsilon[a_1] \cdots [a_k] & \text{ha } k = 1, \dots, n, \end{cases}$$

lineáris kiterjesztéssel.

Ezt a definíciót átfogalmazhatnánk úgy, hogy a bal oldalon a $[a_1, \dots, a_k] \in V^*$ elemeket $\hat{a}_1, \dots, \hat{a}_k \in V^*$ -ra cseréljük, a jobb oldalon pedig az $a_1, \dots, a_k \in V$ elemeket $\hat{a}_1, \dots, \hat{a}_k \in V$ -re. Hasonló definíció adható a kontravariáns megfelelőre, $* : \Lambda_k(V) \rightarrow \Lambda_{n-k}(V)$, amely esetben az $a_1, \dots, a_k \in V$ vektor argumentumok]-konverziója elhagyható.

2.82. Megjegyzés. A Hodge-féle csillag operátor izomorfizmust hoz létre $\Lambda^k(V)$ és $\Lambda^{n-k}(V)$ között, idézzük fel a 26. oldalon található dimenzió-érvelést:

$$\dim \Lambda^k(V) = \binom{n}{k} = \binom{n}{n-k} = \dim \Lambda^{n-k}(V).$$

2.83. Megjegyzés. • Ha a fenti észrevételben $k = 0$ -t veszünk, az összhangban van a $*1 = \epsilon$ definícióval.

- Ha $k = n$ -et veszünk, az megmutatja, hogy $**1 = 1$. Ez az eredmény lényegében

a 2.5. Definíció alapján. A 2.7. Definíció általános esetében ellenőrizhető, hogy $**1 = \sim\partial \times g$, és hogy egy általános $A \in \Lambda^k(V)$ k-formára $**A = (-1)^{k(n-k)} \sim\partial \times gA$.

3. Differenciálgeometria

3.1. Euklideszi tér: Descartes- és görbevonali koordináták

3.1. Definíció. Egy n -dimenziós **euklideszi tér**, E , egy metrikus tér, azaz egy affin tér, amely el van látva egy $d : E \times E \rightarrow \mathbb{R}$ távolságfüggvénnyel, valamint egy $+$: $E \times V \rightarrow E$ leképezéssel, ahol V egy n -dimenziós euklideszi vektortér, úgy, hogy

- minden $x, y \in E$ esetén létezik egy egyértelmű $v \in V$ úgy, hogy $y = x + v$, és $d(x, y) = \sqrt{(v|v)}$;
- minden $x, y \in E$ és minden $v \in V$ esetén $d(x + v, y + v) = d(x, y)$;
- minden $x \in E$ és minden $u, v \in V$ esetén $(x + u) + v = x + (u + v)$.

3.2. Differenciálható sokaságok

3.2. Definíció. Egy n -dimenziós **sima differenciálható sokaság** egy Hausdorff-féle topologikus tér, M , amely el van látva sima diffeomorfizmusok egy $\{\phi_\alpha\}$ családjával, $\phi_\alpha : \Omega_\alpha \subset M \rightarrow \phi_\alpha(\Omega_\alpha) \subset \mathbb{R}^n$, a következő tulajdonságokkal:

- $\{\Omega_\alpha\}$ az M egy nyílt lefedése, azaz Ω_α nyílt és $\cup_\alpha \Omega_\alpha = M$.
- Ha $\Omega_\alpha \cap \Omega_\beta \neq \emptyset$, akkor $\phi_\alpha(\Omega_\alpha \cap \Omega_\beta), \phi_\beta(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n$ nyílt halmazok, és a $\phi_\alpha \circ \phi_\beta^{-1}|_{\phi_\beta(\Omega_\alpha \cap \Omega_\beta)}$ és (így) a $\phi_\beta \circ \phi_\alpha^{-1}|_{\phi_\alpha(\Omega_\alpha \cap \Omega_\beta)}$ leképezések diffeomorfizmusok.
- Az $\{(\Omega_\alpha, \phi_\alpha)\}$ atlasz maximális az előző két axiómára nézve.

3.3. Jelölés. A jelöléssel való visszaélés révén a (3.1–3.2) egyenleteket gyakran $y = y(x)$, illetve $x = x(y)$ alakban egyszerűsítik, a megfelelő értelmezési tartományok hallgatólagos elhagyásával.

3.4. Definíció. A $\phi_\beta \circ \phi_\alpha^{-1}$, illetve $\phi_\alpha \circ \phi_\beta^{-1}$ koordináta-transzformáció **Jacobi-mátrixa** az elsőrendű parciális deriváltak $n \times n$ -es mátrixa:

$$S_j^i(x) = \frac{\partial(\phi_\beta \circ \phi_\alpha^{-1})^i(x)}{\partial x^j} := \frac{\partial y^i}{\partial x^j} \quad \text{ill.} \quad T_j^i(y) = \frac{\partial(\phi_\alpha \circ \phi_\beta^{-1})^i(y)}{\partial y^j} := \frac{\partial x^i}{\partial y^j}.$$

3.3. Érintővektorok

3.5. Definíció. Legyen $\gamma : (-\epsilon, \epsilon) \rightarrow M : t \mapsto \gamma(t)$ egy sima paraméterezett görbe, ahol $\gamma(0) = p \in M$, és $f \in C^\infty(M)$ egy tetszőleges sima skalármező. Ekkor a γ görbe **érintővektora** a p pontban az a lineáris leképezés, amelyet a következő ad meg:

$$\dot{\gamma}(0) : C^\infty(M) \rightarrow \mathbb{R} : f \mapsto \left. \frac{d}{dt}(f \circ \gamma) \right|_{t=0}.$$

3.6. Jelölés. Idézzük fel a 3.4. Definíciót és a (3.3–3.6) egyenleteket. A $\dot{\gamma}(0)$ helyett ezt írjuk:

$$v|_p = v^i \frac{\partial}{\partial x^i} \Big|_p$$

ahol $v|_p = \dot{\gamma}(0)$ és $v^i = \dot{x}^i(0)$. Az alsó index arra emlékeztet, hogy az érintővektor a $p = \gamma(0)$ pontban van definiálva. Ez a konstrukció megismételhető bármely $x \in M$ pontra és bármely $\gamma \in C^\infty((-\epsilon, \epsilon), M)$ sima görbére.

3.7. Megjegyzés. A bázispont, $p \in M$, fontos. Általában magától értetődőnek vesszük, hogy $\langle \cdot, \cdot \rangle \in T_p^*M \otimes T_pM$, $dx^i \in T_p^*M$, $\partial_j \in T_pM$, stb., a $p \in M$ bázispont explicit említése nélkül a jelölés egyszerűsítése érdekében. Ha explicit szeretnénk lenni, a p címkeként csatolható minden lokális geometriai entitáshoz emlékeztetőül.

3.4. Érintő- és kotangens nyaláb

3.8. Definíció. Az **érintőtér** egy $x \in M$ pontban, T_xM , az $x \in M$ pontbeli érintővektorok vektortere.

3.9. Megjegyzés. Idézzük fel az előző szakaszból, hogy a definíció megadható intrinzik módon úgy, hogy a T_xM -et egy $x \in M$ pontban egymást transzverzálisan metsző görbecsalád n érintővektorának feszített tereként definiáljuk.

3.10. Definíció. Az **érintőnyaláb**, $TM = \cup_{x \in M} T_xM$, az érintőterek uniója az M összes x pontjára nézve.

3.11. Jelölés. • Idézzük fel a 3.5. Definíciót. A kotangens teret egy $x \in M$ pontban T_x^*M jelöli.

• Idézzük fel a 3.6. Definíciót. A kotangens nyalábot $T^*M = \cup_{x \in M} T_x^*M$ alakban írjuk.

3.5. Külső derivált

3.12. Definíció. Legyen $\omega^k = \omega_{|i_1 \dots i_k|} dx^{i_1} \wedge \dots \wedge dx^{i_k} = \frac{1}{k!} \omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \in \Lambda^k(TM)$ egy k -ad rangú antiszimmetrikus kovariáns tenzormező, vagy röviden k -forma, ahol $x \in M$. Az ω **külső deriváltja** az a $(k+1)$ -forma, amelyet a következő ad meg:

$$d\omega^k = \frac{1}{k!} d\omega_{i_1 \dots i_k} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} := \frac{1}{k!} \partial_i \omega_{i_1 \dots i_k} dx^i \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \in \Lambda^{k+1}(TM),$$

ahol $df := \partial_i f dx^i$ bármely kellően sima $f : M \rightarrow \mathbb{R}$ skalármező esetén.

3.13. Megjegyzés. Az $1/k!$ tényezőt azért adtuk meg, hogy az $\omega_{i_1 \dots i_k}$ -t az ω^k holorjaként értelmezhesük, mivel

$$\omega^k = \omega_{|i_1 \dots i_k|} dx^{i_1} \wedge \dots \wedge dx^{i_k} = \frac{1}{k!} \omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

ahol az utolsó lépés a holor antiszimmetriája miatti automatikus antiszimmetrizáláson alapul.

3.6. Affin konnexió

3.14. Definíció. Legyen $f \in C^\infty(M)$ egy skalárfüggvény, és $v = v^i \partial_i \in TM$. Ekkor

$$vf := \langle df, v \rangle := v^i \partial_i f.$$

Tehát, amikor egy skalármezőre hat, hallgatólagosan értendő, hogy a $v = \langle d\cdot, v \rangle$ formális azonosítást végezzük el.

3.15. Definíció. Egy **affin konnexió** egy M sokaságon egy $\nabla : TM \times TM \rightarrow TM : (v, w) \mapsto \nabla_v w$ leképezés, amelyet a w **kovariáns deriváltjának** is neveznek a v irányában, a következő tulajdonságokkal. Legyen $f, f_1, f_2 \in C^\infty(M)$, $v, v_1, v_2, w, w_1, w_2 \in TM$, és $\lambda, \lambda_1, \lambda_2 \in \mathbb{R}$, ekkor

1. $\nabla_{f_1 v_1 + f_2 v_2} w = f_1 \nabla_{v_1} w + f_2 \nabla_{v_2} w$ minden $f_1, f_2 \in C^\infty(M)$ esetén,
2. $\nabla_v (\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 \nabla_v w_1 + \lambda_2 \nabla_v w_2$, minden $\lambda_1, \lambda_2 \in \mathbb{R}$ esetén,
3. $\nabla_v (fw) = f \nabla_v w + (\nabla_v f)w$ minden $f \in C^\infty(M)$ esetén. Itt $\nabla_v f \equiv vf$ -et definiáltunk.

3.16. Megjegyzés. Ha elhagyjuk az irányvektor argumentumot, akkor ∇w -t a w **kovariáns differenciáljának** nevezzük, és $\nabla f = df$ -et azonosítunk. Így egy lokális keretvektor (azaz a lokális érintőtér bázisvektorának) kovariáns deriváltja egy vektor ugyanabban az érintőtérben. Következésképpen felírható a keretvektorok lineáris kombinációjaként.

3.17. Definíció. A **Christoffel-szimbólumokat**, Γ_{ij}^k -t, a $\nabla_{e_j} e_i = \Gamma_{ij}^k e_k$ egyenlet definiálja.

3.18. Megjegyzés. A sokaság bármely rögzített pontjában a $\bar{\Gamma}_{ij}^k = 0$ egyenletrendszer meg lehet oldani a Jacobi-mátrix ügyes megválasztásával, azaz egy megfelelő lokális koordináta-rendszerre való áttéréssel. Azonban általában ezt nem lehet globálisan elérni, hacsak a sokaság nem sík (nulla Riemann-görbületű, v.i.).

3.19. Definíció. Lásd 3.9. Definíció. Az affin konnexió kiterjeszthető tetszőleges tenzormezőkre a szorzatszabály általánosításával, nevezetesen bármely tenzormezőpárra, mondjuk $T \in \mathcal{T}_q^p(TM)$ és $S \in \mathcal{T}_s^r(TM)$, megköveteljük, hogy

$$\nabla_v(S \otimes T) = (\nabla_v S) \otimes T + S \otimes (\nabla_v T).$$

3.20. Jelölés. A tenzormezők x -függése hallgatólagosan értendő és elhagyott a jelölésből.

3.7. Lie-derivált

3.21. Definíció. A $w \in V$ **Lie-deriváltja** a $v \in V$ irányában a következőképpen van definiálva:

$$\mathcal{L}_v w = [v, w],$$

ahol az úgynevezett **Lie-zárójelet** a kommutátor definiálja:

$$[v, w]f = v(wf) - w(vf) = (v^i \partial_i w^j - w^i \partial_i v^j) \partial_j f$$

bármely sima f függvényre.

3.8. Torzió

3.22. Definíció. A jelölések a korábbiak. A **torziós tenzor** a következőképpen van megadva:

$$T(v, w) = \nabla_v w - \nabla_w v - [v, w].$$

3.23. Megjegyzés. A torziós tenzor tehát a közönséges deriváltak kommutátora (mint a Lie-deriváltban) és az affin konnexió által indukált kovariáns deriváltak kommutátora közötti különbséget tárja fel.

3.9. Levi-Civita-konnexió

3.24. Definíció. Idézzük fel a 55. oldalon lévő 3.12. Definíciót és a 55. oldalon lévő 3.13. Definíciót. Egy **metrikus konnexió** vagy **Levi-Civita-konnexió** egy (M, g) Riemann-sokaságon a következő tulajdonságokat elégíti ki:

- $z(g(v, w)) = g(\nabla_z v, w) + g(v, \nabla_z w)$ minden $v, w, z \in TM$ esetén.
- $T(v, w) = 0$ minden $v, w \in TM$ esetén.

3.25. Tétel. Idézzük fel a 3.14. Definíciót. A Levi-Civita-konnexiót egyértelműen a következő Christoffel-szimbólumok adják meg:

$$\Gamma_{ij}^k = \frac{1}{2}g^{k\ell}(\partial_i g_{\ell j} + \partial_j g_{\ell i} - \partial_\ell g_{ij}).$$

Bizonyítás. (A 3.1. Tétel bizonyítása.) A bizonyítás (i) az affin konnexió általános tulajdonságain (lásd 3.9. Definíció) és (ii) a Levi-Civita-konnexió definiáló tulajdonságain (lásd 3.14. Definíció) alapul. A részletek: Mutassuk meg, hogy egy bázisra vonatkoztatva a 3.14. Definíció első feltétele ekvivalens a $D_m g_{ij} = 0$ feltétellel, majd ciklikusan permutáljuk a három szabad indexet; vonjuk ki a két kapott kifejezést az adottból. Ezt követően a 3.13. Definíció és az azt követő megfigyelés szerint a 3.14. Definíció második feltétele ekvivalens a $\Gamma_{ij}^\ell = \Gamma_{ji}^\ell$ feltétellel. Alkalmazzuk ezt az előző egyenlőségre. A $-\frac{1}{2}g^{km}$ -mel való kontrakció mindkét oldalon, felhasználva, hogy $g^{ik}g_{kj} = \delta_j^i$, végül a kívánt eredményt adja.

3.10. Geodetikusok

3.26. Definíció. Egy $v = v^i \partial_i \in TM$ vektormező **kovariáns differenciálja**, $\nabla v \in \mathcal{T}_1^1(TM)$, a következőképpen van definiálva:

$$\nabla v = \nabla_{\partial_k} v dx^k.$$

3.27. Megjegyzés. A definíció kiterjeszthető tetszőleges $T \in \mathcal{T}_q^p(TM)$ tenzormezőkre:

$$\nabla T = DT_{j_1 \dots j_q}^{i_1 \dots i_p} dx^{j_1} \otimes \dots \otimes dx^{j_q} \otimes \partial_{i_1} \otimes \dots \otimes \partial_{i_p},$$

ahol $DT_{j_1 \dots j_q}^{i_1 \dots i_p} = D_i T_{j_1 \dots j_q}^{i_1 \dots i_p} dx^i$, lásd 3.4. Eredmény.

3.28. Definíció. A **geodetikus** egy „kovariánsan konstans érintőjű” görbe, vagy egy olyan görbe, amelyet az érintővektorának „párhuzamos eltolásával” kapunk bármely referenciapontból. Más szavakkal, ha $v(t) = \dot{x}^i(t)\partial_i$ ahol $t \in I = (t_0, t_1)$, akkor $\nabla_{v(t)}v(t) = 0$ minden $t \in I$ esetén.

3.29. Megjegyzés. • Hallgatólagosan értendő, hogy $\partial_i \in TM_{x(t)}$ a koordináta bázisvektort jelöli az $x(t)$ pontban.

- Gyakran azonosítják a $v(x(t))$ -t a $v(t)$ -vel. A kontextusból kell kiderülnie, hogy melyik prototípus alkalmazandó.

3.11. Görbület

3.30. Definíció. Adott $v, w \in TM$ esetén az (antiszimmetrikus) **görbületi operátor**, $R(v, w) : TM \rightarrow TM : u \mapsto R(v, w)u$, a következőképpen van definiálva:

$$R(v, w) = [\nabla_v, \nabla_w] - \nabla_{[v, w]}.$$

3.31. Definíció. Adott $u, v, w \in TM, \hat{z} \in T^*M$ esetén a **Riemann-féle görbületi tenzor**, $Riemann \in T^*M \otimes TM \otimes TM \otimes TM$, a következőképpen van definiálva:

$$Riemann(\hat{z}, u, v, w) = \langle \hat{z}, R(v, w)u \rangle.$$

3.32. Definíció. A **Ricci-görbületi tenzor** és a **Ricci-görbületi skalár** a következőképpen van definiálva:

$$\begin{aligned} Ric(v, w) &= R_{\mu\nu} v^\mu w^\nu, \\ R &= g^{\mu\nu} R_{\mu\nu}, \end{aligned}$$

ahol $v, w \in TM$, és a Ricci-holor definíciója $R_{\mu\nu} = R^\rho_{\mu\rho\nu}$.

3.12. Előretolás és visszahúzás (Push-Forward and Pull-Back)

3.33. Definíció. • Az $f \in C^\infty(N)$ függvény ϕ általi **visszahúzása (pull-back)**, $\phi^*f \in C^\infty(M)$, úgy van definiálva, hogy $\phi^*f = f \circ \phi$.

- A $v \in TM$ vektormező ϕ általi **előretolása (push-forward)**, $\phi_*v \in TN$, úgy van definiálva, hogy $(\phi_*v)(f) = v(\phi^*f)$.
- A $\nu \in T^*N$ kovektormező ϕ általi **visszahúzása (pull-back)**, $\phi^*\nu \in T^*M$, úgy van definiálva, hogy $\langle \phi^*\nu, v \rangle_M = \langle \nu, \phi_*v \rangle_N$.

3.34. Megjegyzés. • Vegyük észre, hogy az M és N sokaságok dimenziójának nem kell megegyeznie.