MÉRTÉKELMÉLET ÉS
DINAMIKUS PROGRAMOZÁS

ÚJ SZÉCHENYI TERV

Jegyzetek és példatárak a matematika egyetemi oktatásához sorozat

Algoritmuselmélet
Algoritmusok bonyolultsága
Analitikus módszerek a pénzügyben és a közgazdaságtanban
Analízis feladatgyűjtemény I
Analízis feladatgyűjtemény II
Bevezetés az analízisbe
Complexity of Algorithms
Differential Geometry
Diszkrét matematikai feladatok
Diszkrét optimalizálás
Geometria
Igazságos elosztások
Introductory Course in Analysis
Mathematical Analysis – Exercises I
Mathematical Analysis – Problems and Exercises II
Mértékelmélet és dinamikus programozás
Numerikus funkcionálanalízis
Operációkutatás
Operációkutatási példatár
Parciális differenciálegyenletek
Példatár az analízishez
Pénzügyi matematika
Szimmetrikus struktúrák
Többváltozós adatelemzés
Variációszámítás és optimális irányítás

Magyarkuti Gyula

MÉRTÉKELMÉLET ÉS DINAMIKUS PROGRAMOZÁS

Budapesti Corvinus Egyetem

Typotex

2014

Lektorálta: Dr. Pál Jenő

ISBN 978 963 279 254 5
Készült a Typotex Kiadó (http://www.typotex.hu) gondozásában
Felelős vezető: Votisky Zsuzsa
Műszaki szerkesztő: Gerner József

Készült a TÁMOP-4.1.2-08/2/A/KMR-2009-0045 számú, „Jegyzetek és példatárak a matematika egyetemi oktatásához” című projekt keretében.

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638

MAGYARORSZÁG MEGÚJUL

A projekt az Európai Unió támogatásával, az Európai Regionális Fejlesztési Alap társfinanszírozásával valósul meg.

KULCSSZAVAK: Mérték, valószínűségi mérték, σ-algebra, Caratheodory-kiterjesztés, konvergencia tételek, kettős integrál, Radon–Nikodym-derivált, szuprémum feladat, optimális út, Bellman-egyenlet, Euler-egyenlet, sokk feltétel, Markov-operátor, sztochasztikus mátrix, Banach–Tarski-paradoxon.

ÖSSZEFOGLALÁS: Ez a Corvinus Egyetem Mértékelmélet és Dinamikus programozás kurzusainak jegyzete, amely feltételezi az undergraduális Analízis és Lineáris algebra anyag készségszintű ismeretét. A legfontosabb konvergenciatételek igazolása után a Lebesgue- és Lebesgue–Stieltjes-mértéket a Charateodory-féle kiterjesztési eljárással vezetjük be. Tárgyaljuk a σ-véges mértékek szorzatára vonatkozó Fubini-tételt, majd a Radon–Nikodym-tételnek Neumanntól származó funkcionálanalízis hátterű bizonyítását adjuk. A Dinamikus programozás részben külön tárgyaljuk a determinisztikus és a sztochasztikus esetet, de a két rész a Bellman-egyenlet megoldásáig egymással párhuzamosan fut. A Bellman-egyenlet megoldását a Banach-fixponttétel segítségével állítjuk elő. Az Euler-egyenletet csak determinisztikus esetben tárgyaljuk, míg a sztochasztikus részt a sztochasztikus mátrixokkal kapcsolatos rövid összefoglaló zárja.

Tartalomjegyzék

Ábrák jegyzéke

Előszó

A főszöveg mértékelmélet része elsősorban [Rudin (1987)] és [Browder (1996)], míg a dinamikus programozás rész szinte kizárólag [Stokey–Lucas (1989)] alapján készült. A fenti három könyv mindegyike a szakma nagy-nagy klasszikusának számít, megértésük önmagában is intellektuális élmény, és kötelező olvasmány a téma iránt további érdeklődést mutató hallgatónak. E jegyzet elérte célját, ha a fenti három könyv közül legalább kettő tanulmányozására sarkallja az érdeklődő olvasót.

Példákat, feladatokat szinte egyáltalán nem a tartalmaz a könyv. Nem azért, mert nem tartom nagyon fontosnak a megértett fogalmak gyakorlatokon keresztüli illusztrációját, hanem egyszerűen azért, mert itt csak az előadások követéséhez szükséges anyagot áll módomban összegyűjteni. Gyakorlatul a fenti alapművek gazdag példa- és feladatanyagát javaslom. Kitűzött és megoldott feladatokat a [Pál–Sali (2009)] honlapon találunk. Ezeket kötelező feladatként ajánlom a témában elmélyülni vágyó olvasónak. Igen hasznosnak tartom még [Schilling (2005)] további feladatait a mértékelmélet részhez. A dinamikus programozás rész feladataihoz [Irigoyen et al. (2002)] használata elengedhetetlen.

Néhány szó a könyv struktúrájáról. A főszöveg klasszikus definíció–tétel–bizonyítás-szerkezetben íródik. A jobb megérthetőség kedvéért igyekeztem a magyarázó szövegekben a dolgok „érzelmi” oldalát is megvilágítani. Természetes, hogy az élő kommunikációt a leírt anyag nem pótolja. A fogalmakra sok-sok interpretációt és példát kell adnunk ahhoz, hogy a gondolatok az olvasó lelkéhez férkőzzenek.

A magyarázó szöveget úgy választom el a matematikai formalizmustól, hogy a magyarázó szövegrészek mindig számozatlan bekezdésekben szerepelnek. Hogy érthető legyen, nézzünk erre egy példát.

A 7. oldalon szerepel a 0.1.7. tétel, amely persze a matematikai formalizmushoz tartozik. A tétel után annak bizonyítása következik. A tétel előtt van a 0.1.6. szakasz. E szakasz számozott, tehát olyan megjegyzések vannak itt, amelyek szintén a matematikai formalizmushoz tartoznak, átgondolásuk nélkül a felépítés lyukas marad, később hivatkozás is lehet az itteni fogalmakra. A szakaszt tehát mindenképpen a főszöveg részének tekintem. A 0.1.6. és a 0.1.7. közti rész – itt egyetlen mondat – számozott környezeteken kívül van, tehát nem tartozik a szoros matematikai formalizmus kategóriájába. Egy, a matematikai formalizmust ellenőrző számítógép, e mondat nélkül is értelmesnek találná a szöveget, voltaképpen a szakasz formálisan nézve ki is hagyható. Az ilyen részek tartalmazzák azonban a szóban forgó gondolatokhoz tartozó mögöttes érzelmeket.

Minden matematikáról író szerző számára nehézség, hogy ebből mennyit írjon le. Ha semmit nem ír le, akkor gyakorta éri a vád, hogy az írás követhetetlenül száraz, és a szerző nem őszinte. Ha sokat ír le, akkor az írás dagályossá válik, és az olvasó rászorul a kapott információ szelektálására: fontos vagy nem fontos. Kell ez nekem ahhoz, hogy megértsem mondjuk a Radon–Nikodym-tételt (5.3.5), vagy nem kell? Ha ez túl sokszor merül fel az olvasóban, az elveheti tőle azt a hitet és önbizalmat, hogy igenis képes a tananyagot megtanulni és pontosan megérteni. Ez még akkor is így van, ha mindannyian tudjuk: a valódi megértésre első olvasás után vajmi kevés esélyünk van. Éppen ez az oka annak, hogy fontosnak tartom, akár pillanatnyi érzéki csalódást is megengedve, úgy kialakítani a tananyagot, hogy birtokolhatóságának érzése biztos alapot adjon további tanulmányainkhoz.

A szerző célja tehát az, hogy egy rövid és egyszerű bevezetést nyújtson a mértékelmélet és a dinamikus programozás elméletébe. Összefoglalja az abszolút minimumot, aminek ismerete nélkül nincs esélyünk napjaink közgazdaság-tudományi dolgozatait megérteni. Nem célom, hogy kézikönyvszerű összefoglaló mű szülessen. Néhány ilyen monográfia megtalálható az irodalom-jegyzékben. Ehelyett próbáltam az anyagot úgy összeválogatni, hogy az a legszükségesebbeknél sem többet, sem kevesebbet ne tartalmazzon. Fontosnak tartom, hogy ezáltal olyan *tananyag* jöjjön létre, amely két-két félév alatt, ugyan az ilyenkor elvárható szinten, de valóban megtanulható. Semmiképpen nem szeretném a téma iránt érdeklődő olvasót már a terjedelemmel is elriasztani, ezért a magyarázó szövegeket inkább szűkebbre, mint bővebbre engedtem, és a sokszor rendkívüli módon fontos és érdekes érzelmi tartalmat inkább az előadásokra hagytam.

Köszönettel tartozom hallgatóimnak az évek alatt formálódó jegyzet hibáinak gyűjtögetéséért, és hogy a hibák ellenére néhányuk igazán kiválóan elsajátította az anyagot. Köszönöm Pál Jenő gondos lektori munkáját. Köszönet illeti a Corvinus Egyetem Matematika Tanszékének munkatársait, akik mindig nyitottak és segítőkészek voltak szakmai kérdések megvitatásában. Külön köszönet Szabó Imrének, aki nagy gondossággal olvasta a készülőben lévő verziókat.

0.1. Jelölések, elnevezések, megállapodások

Feltesszük, hogy az olvasó jól ismeri a szokásos egyetemi undergraduális analízis és lineáris algebra tananyagot, l. [Dancs–Puskás (2001)], [Dancs (1992)]. Azért, hogy a sokszor eltérő terminológiát és jelölés módot rögzítsük, felsorolok a szakaszban néhány ismert fogalmat, definíciót és állítást. A felsorolás nem teljes, pusztán azt a célt szolgálja, hogy rövid betekintést nyújtson a később már hivatkozások nélkül felhasznált fogalmakra.

0.1.1.
Az $\mathbb{R}$ szimbólum jelöli a valós számokat, $\mathbb{Q}, \mathbb{C}$ a racionális, illetve a komplex számtestet. Egy $z \in \mathbb{C}$ komplex szám esetén $\Re z, \Im z$ jelöli annak valós, illetve képzetes részét.

A nem negatív *valós számokra* az $\mathbb{R}_+$ jelölést használjuk. Feltesszük, hogy a valós számok topológiai tulajdonságait jól ismeri az olvasó. Tudjuk például, hogy minden nyílt halmaz előáll mint legfeljebb megszámlálhatóan sok nyílt intervallum egyesítése. Heine–Borel-tételként fogunk az $\mathbb{R}$-beli kompaktságra hivatkozni, arra a tényre tehát, hogy az $\mathbb{R}$ valós egyenes korlátos és zárt részhalmazai éppen azok, amelyeknek tetszőleges nyílt halmazokkal való fedéséből kiválasztható véges fedés is.

A *felsőhatár-axiómát* az $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ halmazban használjuk. Az $\mathbb{R}$-beli rendezés természetes módon terjed ki $\overline{\mathbb{R}}$-ra, azaz minden $\alpha \in \mathbb{R}$ mellett $-\infty < \alpha < +\infty$. Így minden nem üres valós halmaznak van szuprémuma, persze $\sup A = +\infty$ előfordulásával is számolnunk kell.

A valós számok szokásos összeadás és szorzás algebrai műveleteit is kiterjesztjük az $\overline{\mathbb{R}}$ halmazra. Például tetszőleges $\alpha > 0$ mellett $\alpha \cdot (+\infty) = +\infty$, értelemszerűen minden $\alpha < 0$ esetén $\alpha \cdot (+\infty) = -\infty$. Hasonlóan, az összeadás műveletet is $\overline{\mathbb{R}}$-ra kiterjesztve használjuk, azaz például $a + \infty = \infty$ minden $a \in \mathbb{R}$, vagy $a = +\infty$ mellett. Természetesen vigyáznunk kell: a bevezetett műveletekkel $\overline{\mathbb{R}}$ nem test!

Érdemes itt végiggondolni az összes előforduló esetet. Világos, hogy van nem definiált lehetőség is, nevezetesen a $(+\infty) - (+\infty)$ operációt ugyanúgy értelmetlennek gondoljuk, mint a zérussal való osztást. Ennek lehetősége soksok kellemetlenség forrása, hiszen minden egyes összeadás művelet végrehajtásakor meg kell gondolnunk, hogy nem a $(+\infty) - (+\infty)$ vagy a $(-\infty) - (-\infty)$ esetről van-e szó. Az előnye viszont az, hogy ezzel elkerülhetünk sok-sok olyan esetszétválasztást, amelyek végül is azonos eredményre vezetnének.

A témával itt ismerkedő olvasó számára furcsa lehet a következő, de az egész könyvben sűrűn és következetesen használt konvenció: A $0 \cdot (+\infty)$ szorzatot értelmesnek tekintjük, és értékét

$$0 \cdot (+\infty) = 0$$

módon definiáljuk. Az elemi analízist értve ez a definíció mindenképpen problémás, mivel cseppet sem intuitív. Itt csak annyit jegyzek meg, hogy új konvenciónk oka az, hogy például az egész számegyenesen értelmezett konstans 1 értékű függvény integrálját $+\infty$-nek szeretnénk látni, közvetlenül, az improprius integrál fogalmát teljesen kikerülve, viszont elvárjuk azt is, hogy a konstans 0 függvény integrálja a 0 valós szám legyen.

Fontos látnunk, hogy ez az első látásra bizarr konvenció nem zavarja meg a *disztributív szabályt* feltéve, hogy csak a nem negatív kiterjesztett valós számok jöhetnek szóba. Világos ugyanis, hogy $\alpha, \beta, \gamma \in \overline{\mathbb{R}}$, de $\alpha, \beta, \gamma \geq 0$ esetén $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

0.1.2.

Az X halmaz *hatványhalmazát* $\mathcal{P}(X)$-szel jelöljük. A szokásos naiv halmazelméleti jelölésrendszert követjük, tehát $A \cap B$ az $A, B \subseteq X$ halmazok *közös része*; az $A \cup B$ e két halmaz *egyesítése*; és $A \setminus B$ a *különbségük*. Ha világos, hogy mi az X alaphalmaz, akkor az A halmaz *komplementerét* $A^c = X \setminus A$ módon jelöljük. A *halmazrendszer* szón mindig egy halmaz hatványhalmazának egy bizonyos részhalmazát értjük. Sokszor fordul elő az szóhasználat, hogy egy $\mathcal{M} \subseteq \mathcal{P}(X)$ halmazrendszer *zárt* egy bizonyos műveletre. Ez azt jelenti, hogy ha a műveletet alkalmazzuk a halmaz elemére vagy elemeire, akkor a művelet eredménye is az $\mathcal{M}$ halmazrendszer eleme marad. Például az $\mathcal{M}$ halmazrendszer metszetzárt, ha $A_1, A_2 \in \mathcal{M}$ esetén $A_1 \cap A_2 \in \mathcal{M}$ teljesül; komplementer zárt, ha az $A \in \mathcal{M}$ tény implikálja az $A^c \in \mathcal{M}$ teljesülését.

Ha $\mathcal{M} \subseteq \mathcal{P}(X)$ egy halmazrendszer, akkor az

$$\bigcup \mathcal{M} = \bigcup \{A : A \in \mathcal{M}\} = \bigcup_{A \in \mathcal{M}} A$$

szinonim jelölések. Hasonló szinonimákat használunk a metszetre de más halmazelméleti műveletre is.

Egy halmazról azt mondjuk, hogy az *legfeljebb megszámlálható számosságú*, ha véges vagy megszámlálhatóan végtelen számosságú.

Az $f : X \to Y$ jelölés azt jelenti, hogy f egy olyan függvény, amelynek *értelmezési tartománya* a teljes X halmaz és *értékkészlete* az Y halmaz egy részhalmaza. Az f maga a függvény, $f(x)$ az x pontban felvett Y-beli érték. Ha a formula olyan, hogy nem világos, mi is a függvény argumentuma, akkor a *pont-jelölést* használom. Például, ha $f : X_1 \times X_2 \to Y$, akkor $f(x_1, \cdot) : : X_2 \to Y$ a rögzített x_1 melletti második változó szerinti parciális függvény, tehát az a függvény, amelynek egy $x_2 \in X_2$ pontban az értéke $f(x_1, x_2)$.

Az $\mathcal{R}(f)$ szimbólum az f függvény értékkészletét jelöli. Egy $f : X \to \overline{\mathbb{R}}$ függvény esetén $X(f > \alpha)$ jelöli az $\{x \in X : f(x) > \alpha\} = f^{-1}((\alpha, +\infty])$ *ősképet*. A fentivel analóg módon értelmezendők az $X(f \geq \alpha), X(f < \alpha), X(f \leq \alpha)$ stb. halmazok is.

Legyen adva egy $f_n : X \to \overline{\mathbb{R}}$ függvénysorozat. Azt mondjuk, hogy f_n *pontonként konvergál* az $f : X \to \overline{\mathbb{R}}$ függvényhez, ha minden $x \in X$ mellett $\lim_{n\to\infty} f_n(x) = f(x)$.

Hasonlóan a *pontonkénti szuprémum*: $(\sup_{n\in\mathbb{N}} f_n)(x) = \sup_{n\in\mathbb{N}} f_n(x)$.

Ha $\sup_X |f_n - f| \to 0$, midőn $n \to \infty$, akkor azt mondjuk, hogy az f_n *egyenletesen konvergál* az f határfüggvényhez.

Fontos szóhasználat a majoráns fogalma. Ha a h függvényre $f_n \le |h|$ teljesül minden $n \in \mathbb{N}$ mellett, akkor h függvény az f_n sorozat *majoránsa*.

0.1.3.

Egy *topologikus teret* annak nyílt halmazai segítségével definiálunk és ennek megfelelően (X, τ)-val fogunk jelölni. A $\tau \subseteq \mathcal{P}(X)$ halmazrendszer elemeit *nyílt halmazoknak* deklaráljuk. Egy (X, d) *metrikus térben* $B^\circ(u, r) = \{x \in X : d(x, u) < r\}$ jelöli az u középpontú r sugarú *nyílt gömböt* és $B(u, r) = \{x \in X : d(x, u) \le r\}$ az ugyanilyen sugarú *zárt gömböt*.

Az n-dimenziós $\mathbb{R}^n$ térben egy halmazt *balról zárt, jobbról nyílt intervallumnak* mondunk, ha az $[\alpha_1, \beta_1) \times [\alpha_2, \beta_2) \times \ldots \times [\alpha_n, \beta_n)$ alakú, ahol $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{R}$. Analóg módon definiáljuk az n-dimenziós térbeli zárt, nyílt stb. intervallumokat is.

Az $\overline{\mathbb{R}}$ halmazon használt topológiában egy $x \in \mathbb{R}$ valós pont környezetei azok a halmazok, amelyek tartalmaznak egy $(x - \varepsilon, x + \varepsilon)$ intervallumot valamilyen $\varepsilon > 0$ mellett. A $+\infty$ szimbólum környezete az összes olyan halmaz, amely tartalmaz $(\alpha, +\infty]$ intervallumot valamely $\alpha \in \mathbb{R}$-rel. Hasonlóan, a $-\infty$ környezetei azon halmazok amelyek tartalmaznak $[-\infty, \alpha)$ alakú intervallumot alkalmas $\alpha \in \mathbb{R}$ megválasztásával. Az $\overline{\mathbb{R}}$ nyílt halmazai azok a halmazok, amelyek valamennyi pontjuknak környezetei is.

0.1.4.

Emlékezzünk arra, hogy egy $\mathbb{R}$ feletti vektorteret *skaláris szorzatos térnek*, vagy *Euklideszi-térnek* nevezünk, ha létezik egy

$$\langle \cdot ; \cdot \rangle : X \times X \to \mathbb{R}$$

függvény, amelyre:

(i) $\langle x, x\rangle \ge 0$ minden $x \in X$ és $\langle x, x\rangle = 0$ pontosan akkor, ha $x = 0$; (ii) $\langle x, y\rangle = \langle y, x\rangle$ minden $x, y \in X$ esetén; (iii) minden rögzített $y \in X$ mellett az $\langle \cdot ; y\rangle : X \to \mathbb{R}$ függvények lineáris funkcionálok.

Egy $\mathbb{R}$ feletti vektorteret *normált térnek* nevezünk, ha a vektortéren értelmezve van egy

$$\|\cdot\| : X \to \mathbb{R}_+$$

függvény, amelyre: (i) $\|x\| = 0$ pontosan akkor, ha $x = 0$; (ii) minden $\alpha \in \mathbb{R}$ skalár és minden $x \in X$ vektor esetén $\|\alpha x\| = |\alpha| \, \|x\|$; (iii) fennáll az úgynevezett *háromszög-egyenlőtlenség*, azaz tetszőleges $x, y \in X$ mellett $\|x + y\| \le \|x\| + \|y\|$. E normált teret $(X, \|\cdot\|)$ módon jelöljük.

Azt mondjuk, hogy az $x_n \in X$ sorozat az $\|\cdot\|$ *normában konvergál* az $x \in X$ ponthoz, ha $\|x_n - x\| \to 0$. A háromszög-egyenlőtlenség szerint $|\|x_n\| - \|x\|| \leq \leq \|x_n - x\| \to 0$, tehát a norma-konvergenciából a sorozat elemek normáinak konvergenciája következik, amit úgy is kifejezhetünk, hogy a $\|\cdot\| \to \mathbb{R}$ norma-leképezés folytonos.

Egy $(X, \langle\cdot,\cdot\rangle)$ skaláris szorzatos térben az x vektor hossza, vagy *normája*:

$$\|x\| = \sqrt{\langle x; x\rangle}.$$

A háromszög-egyenlőtlenségtől eltekintve a normaaxiómák nyilvánvaló módon teljesülnek. De minden $x, y \in X$ mellett

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$$

(*Schwartz-egyenlőtlenség*), ezért a fent definiált függvény valóban normát definiál X-en. Ugyanis: $0 \leq \langle x - ry, x - ry\rangle = \|x\|^2 + r^2\|y\|^2 - 2r\langle x, y\rangle$ fennáll minden $r \in \mathbb{R}$ esetén. Ha $y \neq 0$, akkor ezt mint r kvadratikus függvényét tekintve a diszkrimináns nem lehet pozitív, azaz $4\langle x, y\rangle^2 - 4\|y\|^2\|x\|^2 \leq 0$, ami épp a Schwartz-egyenlőtlenség. Innen már a háromszög-egyenlőtlenség $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$ módon könnyen következik.

A skalárszorzat linearitását használó számolás mutatja, hogy egy skaláris szorzatos térben fennáll a *paralelogramma szabály*: $\|x + y\|^2 + \|x - y\|^2 = = 2\left(\|x\|^2 + \|y\|^2\right)$. Azt is meg lehet mutatni, hogy egy normált tér éppen akkor skaláris szorzatos tér, ha teljesül a paralelogramma szabály.

0.1.5.

Legyen adva egy $(X, \|\cdot\|)$ normált tér. Az $x_n \in X$ sorozatot *Cauchy-sorozatnak* mondjuk, ha minden $\varepsilon > 0$-hoz létezik $N \in \mathbb{N}$ index, hogy bármely $n, m > N$ mellett $\|x_n - x_m\| < \varepsilon$. Könnyű gyakorlat, hogy minden konvergens sorozat egyben Cauchy-sorozat is. Azt is gondoljuk meg, hogy amennyiben egy Cauchy-sorozatnak van konvergens részsorozata, akkor a sorozat maga is konvergens, és persze határértéke azonos a konvergens részsorozatának határértékével.

Ha egy normált-térben minden Cauchy-sorozat konvergens is, akkor ezt a normált teret *teljes normált térnek* vagy, a rövidség okán, *Banach-térnek* nevezzük. Korábbi ismereteinkből tudjuk, hogy minden véges dimenziós normált tér Banach-tér.

0.1.6.

Legyen $(X, \|\cdot\|)$ egy normált tér és $a_n \in X$ egy sorozat. Azt mondjuk, hogy a $\sum_{n=1}^{\infty} a_n$ sor *konvergens*, ha létezik $a \in X$ vektor, amely az $s_n = \sum_{k=1}^{n} a_k$ részletösszeg-sorozat határértéke, azaz $\|a - s_n\| \to 0$, mikor $n \to \infty$. Ilyenkor az a vektort mondjuk a sor *összegének*, amelyet $a = \sum_{n=1}^{\infty} a_n$ módon jelöljük.

A $\sum_{n=1}^{\infty} a_n$ sor a normált tér *abszolút konvergens* sora, ha az elemek normájából álló valós sor konvergens, azaz ha $\sum_{n=1}^{\infty} \|a_n\| < \infty$.

Nagyon fontos, ezért itt igazoljuk is, hogy egy normált tér teljessége megfogalmazható a Cauchy-sorozat fogalma nélkül is, pusztán az abszolút konvergens sor fogalmának segítségével.

0.1.7. tétel. *Legyen $(X, \|\cdot\|)$ normált tér. E normált tér pontosan akkor teljes, tehát Banach-tér, ha minden abszolút konvergens sora konvergens is.*

Bizonyítás. Tegyük fel először, hogy $(X, \|\cdot\|)$ egy Banach-tér. Tekintsük a $\sum_{n=1}^{\infty} a_n$ abszolút konvergens sor $s_n = \sum_{k=1}^{n} a_k$ részletösszeg-sorozatát. Világos, hogy például $n > m$ mellett

$$\|s_n - s_m\| = \Big\| \sum_{k=m+1}^{n} a_k \Big\| \leq \sum_{k=m+1}^{n} \|a_k\|.$$

No de $\sum_{k=1}^{\infty} \|a_k\| < \infty$ miatt minden $\varepsilon > 0$-hoz létezik N index, hogy minden $m, n > N$ mellett

$$\sum_{k=m+1}^{n} \|a_k\| < \varepsilon,$$

ami azt jelenti, hogy az s_n részletösszegek sorozata egy Cauchy-sorozat. Az $(X, \|\cdot\|)$ normált tér teljessége szerint s_n konvergens, ami éppen a $\sum_{n=1}^{\infty} a_n$ sor konvergenciáját jelenti.

Megfordítva, tegyük fel, hogy minden abszolút konvergens sor konvergens is, és tekintsünk egy $a_n \in X$ Cauchy-sorozatot. Az $\varepsilon = 1/2$ számhoz létezik $n_1 \in \mathbb{N}$ index, hogy bárhogyan is választjuk majd meg az $n_2 > n_1$ számot az $\|a_{n_1} - a_{n_2}\| < 1/2$ egyenlőtlenség fennáll. Most tekintsük az $\varepsilon = 1/4$ számot. A Cauchy-sorozat definíciója szerint létezik olyan $n_2 > n_1, n_2 \in \mathbb{N}$ index, hogy bármely később definiált $n_3 > n_2$ mellett $\|a_{n_2} - a_{n_3}\| < 1/4$ teljesül. Az eljárást folytatva olyan a_{n_k} részsorozatot nyerünk, melynek minden $k \in \mathbb{N}$ elemére fennáll az

$$\|a_{n_k} - a_{n_{k+1}}\| < \frac{1}{2^k}$$

egyenlőtlenség. A geometriai sor konvergenciája tehát biztosítja a

$$\sum_{k=1}^{\infty} \big(a_{n_k} - a_{n_{k+1}}\big)$$

sor abszolút konvergens voltát. Feltételünk szerint e sor konvergens is. No de a K-adik részletösszeget kiszámolva:

$$s_K = (a_{n_2} - a_{n_1}) + (a_{n_3} - a_{n_2}) + (a_{n_4} - a_{n_3}) + \cdots + \big(a_{n_{K+1}} - a_{n_K}\big) = a_{n_{K+1}} - a_{n_1}.$$

Az $(s_K)_{K \in \mathbb{N}}$ sorozat konvergenciája ezért az a_{n_K} részsorozat konvergenciájával ekvivalens. Összefoglalva: azt kaptuk, hogy az $(a_n)_{n \in \mathbb{N}}$ Cauchy-sorozatnak van $(a_{n_K})_{K \in \mathbb{N}}$ részsorozata, amely konvergens. Ekkor az eredeti a_n sorozat is konvergens. Ezt kellett belátni. $\square$

Az anyagban előrehaladva látjuk majd, hogy bizonyos függvénytereknek mint normált tereknek a teljessége kulcsfontosságú tény. Érdekes módon valós függvénytanban és funkcionálanalízisben is sokszor fordul elő, hogy egy bizonyos függvénytér teljessége e kritériummal egészen természetes módon ragadható meg. Erre látni fogunk egy kulcsfontosságú példát az 5.1.12 Riesz–Fischer-tételben.

0.1.8.
Buroktéren egy $(X, \mathcal{H})$ párt értünk, ahol X egy rögzített halmaz, $\mathcal{H} \subseteq \mathcal{P}(X)$ pedig olyan halmazrendszer, amely elemként tartalmazza az X halmazt, továbbá akárhány $\mathcal{H}$-beli halmaz közös része is $\mathcal{H}$-beli. A $\mathcal{H}$-beli halmazokat *burok-zárt* halmazoknak is mondjuk. Tetszőleges $A \subseteq X$ mellett

$$\mathrm{cl}\,(A) = \bigcap \{H \in \mathcal{H} : A \subseteq H\}$$

a buroktér *burok-operációja*, vagy *lezárási operációja.* A $\mathrm{cl}\,(A)$ halmaz tehát az A halmazt tartalmazó legszűkebb $\mathcal{H}$-beli, azaz burok-zárt halmaz. Ha a szövegkörnyezetből világos, hogy mi a burok-operáció, akkor azt mondjuk, hogy $\mathrm{cl}\,(A)$ az A által *generált burok-halmaz.* Világos, hogy e burok-operáció rendelkezik az alábbi tulajdonságokkal: (i) $A \subseteq \mathrm{cl}\,(A)$; (ii) $A \subseteq B \Rightarrow \mathrm{cl}\,(A) \subseteq \subseteq \mathrm{cl}\,(B)$; (iii) $\mathrm{cl}\,(\mathrm{cl}\,(A)) = \mathrm{cl}\,(A)$. A burok-operáció fenti definíciója szerint minden $H \in \mathcal{H}$ burok-zárt halmaz esetén $\mathrm{cl}\,(H) = H$.

Ha $B = \mathrm{cl}\,(A)$, akkor szokás azt mondani, hogy B halmazt az A halmaz *generálja*, vagy ami ugyanaz: A a B-nek *generátora.* Egy burok-zárt halmaznak sok-sok generáló halmaza lehetséges, tehát a generáló halmaz nem egyértelmű fogalom. Az iménti esetben például A, de B is generáló halmaza B-nek.

A következő típusú érvelést nagyon sokszor használjuk. Ha $B \in \mathcal{H}$ egy burok-zárt halmaz, és a $A \subseteq B$, akkor a fenti (ii) tulajdonság szerint $\mathrm{cl}\,(A) \subseteq \subseteq \mathrm{cl}\,(B) = B$. De úgy is indokolhatnánk, hogy $\mathrm{cl}\,(A)$ fenti definíciójára figyelve B az egyik metszendő halmaz. No persze a metszet minden metszendő halmaznak részhalmaza.

Analízis és algebra tanulmányaink folyamán sok-sok példát láttunk már burok-operációra. Például: lineáris altér, konvexitás, affinitás, topológiai lezárás. Mondjuk a lineáris altér esetében ez azt jelenti, hogy ha $\mathcal{H}$ jelöli valamely rögzített X vektortér összes altereinek halmazát, akkor $(X, \mathcal{H})$ egy buroktér, hiszen akárhány altér metszete is altér. E buroktér lezárási operátora a lineáris burok operátor, amely a vektortér egy $H \subseteq X$ részhalmazához hozzárendeli annak $\mathrm{lin}\,H$ lineáris burkát. A lezárási operátor definíciója szerint ez nem más, mint a H halmazt tartalmazó összes lineáris alterek közös része. Ezt neveztük generált altérnek. Szokásos szóhasználat még, hogy a generált altér *külső reprezentációját* kapjuk a H halmazt tartalmazó összes altér metszeteként.

Teljesen analóg módon, a generált konvex(affin) halmaz külső reprezentációja, a halmazt tartalmazó összes konvex(affin) halmazok metszete.

Az itt említett példákban jól ismerjük a *belső reprezentációt* is. Vektorterek esetében $\operatorname{lin} H$ az összes H-beli elem összes lineáris kombinációja, míg a konvex(affin) halmazok alkotta buroktér esetében H halmaz $\operatorname{co} H$ (affin esetben $\operatorname{aff} H$) konvex(affin)-burka az összes H-beli elem konvex(affin) kombinációinak halmaza.

Ezt a szóhasználatot megtartjuk általában a burokterekre is. Ha cl jelöli az $(X, \mathcal{H})$ burok-operációját, akkor $\operatorname{cl} A$, mint az $A \subseteq X$ halmazt tartalmazó legszűkebb cl-zárt halmaz, az A-t tartalmazó összes $\mathcal{H}$-beli halmaz közös része, a generált $\operatorname{cl}(A)$ halmaz külső reprezentációja. A belső reprezentációra nincs ilyen egyszerű általános szabály. Hogy hogyan tudjuk felépíteni a generált halmazt az adott halmazból, az persze buroktérről buroktérre más és más. Látni fogunk olyan burokteret is az anyagban, amikor csak a külső reprezentációt használva tudjuk a generált halmazt leírni.

0.1.9. (Cantor-halmaz)
Ismert, hogy a $[0{,}1]$ intervallumban minden valós szám felírható 3-as számrendszerben. Arról van szó, hogy ha $x \in [0{,}1]$, akkor léteznek $a_j \in \{0{,}1{,}2\}$, $j \in \mathbb{N}$ számjegyek, amelyekre $x = \sum_{j=1}^{\infty} a_j 3^{-j}$. E felírás gyenge pontja, hogy vannak olyan számok, amelyeknek több ilyen előállítása is lehetséges, a

$$\sum_{j=1}^{\infty} \frac{2}{3^{j+k}} = \frac{1}{3^k}$$

egyenlőség szerint. Ha egy szám véges sok – például k darab – jegy segítségével is reprezentálható, azaz éppen $p3^{-k}$ alakú valamely $k, p \in \mathbb{N}$ mellett, akkor két különböző felírás is van, egyébként a felírás egyértelmű. Például, ha $x = 14 \cdot 3^{-4}$, akkor

$$x = \begin{cases} 0.0112 \\ 0.011122222\dots \end{cases}$$

A Cantor-halmazt a hármas számrendszerben felírt alak segítségével definiáljuk, ezért szükségünk van egy egyértelmű reprezentáció rögzítésére, amely a következő legyen. Ha x nem $p \cdot 3^{-k}$ alakú, akkor a felírás egyértelmű, és mi is ezt használjuk. Ha $x = p \cdot 3^{-k}$ alakú, ahol 3 már nem osztója p-nek, akkor p-nek 3-mal való osztási maradéka 1 vagy 2. Ha az osztási maradék 1, akkor két felírása van x-nek, de mi válasszuk ki a másodikat

$$x = \begin{cases} 0.a_1 a_2 \dots a_{k-1} 1 \\ \boxed{0.a_1 a_2 \dots a_{k-1} 02222\dots} \end{cases}$$

Míg ha az osztási maradék 2, akkor a két felírás

$$x = \begin{cases} \boxed{0.a_1 a_2 \dots a_{k-1} 2} \\ 0.a_1 a_2 \dots a_{k-1} 12222\dots, \end{cases}$$

de mi válasszuk az elsőt. Úgy tudunk egyszerűen fogalmazni, hogy a $p \cdot 3^k$ alakú szám lehetséges felírásai közül azt válasszuk ki az egyértelmű reprezentáció céljára, amelyiknek a k-adik helyén lévő a_k számjegyére $a_k \neq 1$ teljesül. Ezt a konvenciót betartva elmondhatjuk, hogy minden szám egyféleképpen írható fel hármas számrendszerben.

Az egyértelmű reprezentációknak az az előnye, hogy így a szokásos rendezés megegyezik a *lexikografikus rendezéssel.* Ez azt jelenti, hogy ha $x = 0.a_1a_2a_3\ldots$ és $y = 0.b_1b_2b_3\ldots$ valós számok 3-as számrendszerben felírva, akkor $x < y$ pontosan akkor teljesül, ha létezik $n \in \mathbb{N}$ melyre $a_n < b_n$, és $a_j = b_j$ minden $j < n, j \in \mathbb{N}$ esetén.

Megegyezésünk szerint

$$\frac{1}{3} = 0.0222\ldots, \quad \frac{2}{3} = 0.2$$

így az első számjegyre $a_1 = 1$ pontosan akkor teljesül, ha $x \in (1/3, 2/3)$. Hasonló okból $a_1 \neq 1$ mellett $a_2 = 1$ pontosan akkor teljesül, ha x a $[0,1] \setminus (1/3, 2/3)$ halmazt alkotó két zárt intervallum egyikének belső nyílt harmadába esik, azaz $x \in (1/9, 2/9) \cup (7/9, 8/9)$. Ugyanígy $a_1 \neq 1, a_2 \neq 1$ mellett $a_3 = 1$ pontosan akkor teljesül, ha x a

$$[0,1] \setminus ((1/3, 2/3) \cup (1/9, 2/9) \cup (7/9, 8/9))$$

halmazt alkotó négy zárt intervallum valamelyikének belső nyílt harmadába esik. Így tovább az összes számjegyre.

A [0,1] intervallum azon pontjait, melyek fenti alakú hármas számrendszerbeli alakjában az összes számjegy 1-től különböző, a következő rekurzióval adhatjuk meg. Legyen C_1 a [0,1] zárt intervallum középső nyílt harmada, azaz $C_1 = (1/3, 2/3)$. Most tekintsük a $[0,1] \setminus C_1$ halmazt. Ez két zárt intervallum egyesítése. Legyen C_2 ezen zárt intervallumok középső nyílt harmadainak egyesítése, C_2 tehát két darab 1/9 hosszú nyílt intervallum uniója. Most tekintsük a $[0,1] \setminus (C_1 \cup C_2)$ halmazt. Ez négy darab zárt intervallum egyesítése. Álljon C_3 ezen zárt intervallumok középső nyílt harmadainak egyesítéséből, azaz C_3 négy darab 3^{-3} hosszú nyílt intervallum uniója. Az n-edik lépésben a

$$[0,1] \setminus \bigcup_{k=1}^{n-1} C_k$$

halmaz 2^{n-1} darab zárt intervallum egyesítése. Minden egyes zárt intervallumnak vegyük a középső nyílt harmadát, majd jelölje ezen nyílt intervallumok egyesítését C_n. Ekkor tehát C_n halmaz 2^{n-1} darab egyenként 3^{-n} hosszú nyílt intervallumocska egyesítése. A C_n halmazt alkotó nyílt intervallumok végpontjának 3-as számrendszerbeli alakjában az n-edik előtti jegyek az 1-est nem tartalmazzák. Egy jobb végpont esetében az n-edik jegye 2, az összes többi jegy 0. Egy bal végpont esetében pedig az n-edik jegy 0 és az összes többi jegy 2. Így tovább minden $n \in \mathbb{N}$-re.

Összefoglalva az eddigieket kapjuk a Cantor-halmaz definícióját:

0.1.10. definíció (Cantor-halmaz)**.** A [0,1] intervallum számait írjuk fel a fenti megállapodás szerinti egyértelműen reprezentált 3-as számrendszerbeli alakban. A *Cantor-halmazt* azok a pontok alkotják, amelyeknek van olyan 3-as számrendszerbeli felírása, amiben az 1 számjegy nem fordul elő. Ezek szerint

$$C = \left\{ x \in [0{,}1] : x = \sum_{j=1}^{\infty} a_j 3^{-j}, a_j \in \{0{,}2\}\, \forall j \in \mathbb{N} \right\}$$

A fent definiált C_j nyílt halmazokkal kifejezve:

$$C = [0{,}1] \setminus \left(\cup_{j=1}^{\infty} C_j\right) = \bigcap_{j=1}^{\infty} ([0{,}1] \setminus C_j)$$

A C halmazt nevezzük *Cantor-halmaznak*.

0.1.11. (Cantor-halmaz kicsi is, nagy is)
A C_n halmaz 2^{n-1} darab 3^{-n} hosszú nyílt intervallum egyesítése, így összhossza $\frac{2^{n-1}}{3^n}$. Az összes elhagyott, egymástól diszjunkt nyílt intervallum hosszainak összege:

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = \frac{1}{3} \frac{1}{1 - \frac{2}{3}} = 1,$$

ami azt jelenti, hogy a Cantor-halmaz mértékelméleti értelemben kicsi.

No de számossági értelemben nagy, hiszen kontinuum számosságú halmaz. Tekintsük ugyanis a Cantor-halmaz egy pontját, annak 3-as számrendszerben felírt alakjában. Cseréljük ki az összes 2-es számjegyet 1-re, és az így keletkező számot tekintsük egy 2-es számrendszerben felírt számnak. Világos, hogy ilyen módon [0,1] intervallum minden y eleme előáll, hiszen a Cantor-halmaz azon pontjának a képe, amelyiknek számjegyeit az y eredeti 2-es számrendszerben felírt számjegyeinek kétszereseként kapjuk éppen y. Szürjekciót definiáltunk tehát a Cantor-halmazról a [0,1] intervallumra, ezért a Cantor-halmaz számossága legalább kontinuum. Az viszont a $C \subseteq [0{,}1]$ tartalmazásból nyilvánvaló, hogy legfeljebb kontinuum számosságú halmaz a Cantor-halmaz, tehát a *Bernstein-tétel* szerint a Cantor-halmaznak és a [0,1] intervallumnak azonos a számossága.

0.1.12. (Cantor-függvény)
Érdemes kicsit alaposabban is szemügyre venni a fenti $f : C \to [0{,}1]$

$$\sum_{j=1}^{\infty} a_j 3^{-j} \mapsto \sum_{j=1}^{\infty} \alpha_j 2^{-j}, \qquad \text{ahol} \;\; \alpha_j = a_j/2$$

szürjektív leképezést. Ha $x, y \in C$ a Cantor-halmaz két különböző pontja, akkor a 3-as számrendszerbeli $x = \sum_{j=1}^{\infty} a_j 3^{-j}$ és $y = \sum_{j=1}^{\infty} b_j 3^{-j}$ előállításra igaz, hogy

$$x < y \iff \exists n \in \mathbb{N}, \text{ hogy } a_n = 0 < 2 = b_n, \text{ de } a_k = b_k \forall k < n \text{ mellett.}$$

A pontok képére tehát $f(x) = \sum_{j=1}^{\infty} \alpha_j 3^{-j}$ és $f(y) = \sum_{j=1}^{\infty} \beta_j 3^{-j}$, ahol $\alpha_j = a_j/2$ és $\beta_j = b_j/2$ minden $j \in \mathbb{N}$ mellett. Az $x < y$ esetben tehát

$$\exists n \in \mathbb{N}, \text{ hogy } \alpha_n = 0 < 1 = \beta_n, \text{ de } \alpha_k = \beta_k \quad \forall k < n \text{ mellett.}$$

Ebből azonnal következik, hogy f monoton növő, hiszen

$$f(y) = \sum_{j=1}^{\infty} \beta_j 2^{-j} \geq \sum_{j=1}^{n} \beta_j 2^{-j} = \sum_{j=1}^{n-1} \alpha_j 2^{-j} + 2^{-n} = \sum_{j=1}^{n-1} \alpha_j 2^{-j} + \sum_{j=n+1}^{\infty} 2^{-j} \geq$$
$$\geq \sum_{j=1}^{\infty} \alpha_j 2^{-j} = f(x).$$

Az is látszik, hogy $f(y) = f(x)$ akkor és csak akkor fordul elő, ha $\beta_j = 0$ és $\alpha_j = 1$ minden $j > n$ mellett, ami a $b_j = 0$ és $a_j = 2$ esetnek felel meg, amikor is x, y a Cantor-halmaz két olyan pontja, amely az n-edik helyükön 1-es számjegyet tartalmazó pontok alkotta egyik nyílt intervallum két végpontját alkotják. Ilyenkor a két képpont ugyanannak a számnak a két különböző 2-es számrendszerbeli előállítása, úgy hogy mindig a jobb végpont képe a véges előállítás. Ha elhagynánk a Cantor-halmazból az elhagyott nyílt intervallumok végpontjainak egyikét is, akkor a fenti leképezés már bijekció is lenne a Cantor-halmaz és a [0,1] intervallum közt.

Az a tény, hogy egy kimaradó nyílt intervallum két végpontjában a fenti hozzárendelés azonos értéket ad lehetővé teszi, hogy az $f : C \to [0,1]$ függvényt kiterjesszük a teljes [0,1] zárt intervallumra. Legyen ugyanis $F(u) = f(u)$, ha $u \in C$. Ha $u \notin C$, akkor u az egyik elhagyott nyílt intervallumba esik. Ennek két végpontja a Cantor-halmaz olyan x, y eleme, amelyeken f értéke azonos. Legyen $F(u) = f(x) = f(y)$. Az F függvényt *Cantor-függvénynek* nevezzük.

A Cantor-függvény tehát olyan $F : [0,1] \to [0,1]$ monoton növekedő függvény, amelyre $F(0) = 0, F(1) = 1$, de F konstans – persze más és más konstans – megszámlálhatóan sok diszjunkt nyílt intervallumon, amelyeknek összhossza 1. Egy monoton növekedő függvénynek minden pontjában van bal és jobb oldali határértéke is. Ha F nem lenne folytonos egy pontban, akkor ott a bal és a jobb oldali határérték sem lenne azonos, ezért az ezek alkotta nyílt intervallum nem lenne üres és kimaradna F értékkészletéből. Mivel F szürjekció, ezért kimaradó érték nincs, így szakadási pont sem lehetséges.

A Cantor-függvény tehát folytonos, monoton növekedő függvény, amely a Cantor-halmaz komplementerét alkotó nyílt intervallumokon konstans, persze itt differenciálható is és deriváltja zérus.

0.1.13. (Cantor-halmaz kompakt, belseje üres, nincs izolált pontja)
A Cantor-halmaz kompakt, hiszen korlátos és zárt. Belseje üres, hiszen tetszőleges pontjának tetszőleges környezete tartalmaz olyan számot, amelynek számjegyei közt az 1-es szerepel. Viszont tetszőleges pontjának tetszőleges környezete tartalmaz olyan másik számot is, amelynek számjegyei közt az 1 nem szerepel, ezért a Cantor-halmaznak nincs izolált pontja. Így egyetlen pontja sem lehet egyszerre bal és jobb végpontja is egy-egy a Cantor-halmazból kimaradó intervallumnak.

I. rész

Mértékelmélet

1. fejezet

Mérhetőség

1.1. Félgyűrű, gyűrű, σ-algebra és monoton osztály

1.1.1. definíció (félgyűrű)**.** Egy X-beli nem üres $\mathcal{P}$ halmazrendszert *félgyűrűnek* nevezünk, ha zárt a metszetre és bármely két elem különbsége előáll véges sok diszjunkt $\mathcal{P}$-beli elem egyesítéseként.

Kicsit formálisabban, a $\mathcal{P}$ halmazrendszerről, az alábbi két tulajdonságot követeljük meg:

1. bármely két $A, B \in \mathcal{P}$ esetén $A \cap B \in \mathcal{P}$;
2. bármely két $A, B \in \mathcal{P}$-hez létezik $n \in \mathbb{N}$ egész és léteznek diszjunkt $C_1, \ldots, C_n \in \mathcal{P}$ halmazok, amelyekre $A \setminus B = \bigcup_{i=1}^{n} C_i$

A $\mathcal{P}$ nem ürességének feltétele csak az érdektelen eset mellőzését jelenti. Világos, hogy $\varnothing \in \mathcal{P}$ mindig fennáll a fenti 2. tulajdonság szerint.

1.1.2.
Ilyen például $\mathbb{R}$-en az összes intervallumok halmaza, vagy az összes balról zárt jobbról nyílt intervallumok halmaza. Látható, hogy két ilyen intervallum metszete is ilyen. Az is egészen szemléletes, hogy két ilyen intervallum különbsége előáll legfeljebb két balról zárt, jobbról nyílt diszjunkt intervallum egyesítéseként.

Amint a következő állítás mutatja, az $\mathbb{R}^n$-beli balról zárt, jobbról nyílt halmazok rendszere is félgyűrűt alkot. A fogalom jobb megértése kedvéért gondoljuk végig, hogy mondjuk a 2- vagy 3-dimenziós euklideszi tér balról zárt, jobbról nyílt intervallumai miért alkotnak félgyűrűt. Érdemes meggondolni, hogy vajon legfeljebb hány darab diszjunkt halmazra van szükség két ilyen intervallum különbségének előállításához. Láttuk, hogy 1 dimenzióban két halmaz elég. Kis rajzolgatás után rájövünk, hogy 2 dimenzióban 4 ez a szám. Akinek jó térlátása van rájön, hogy 3-dimenziós térben 6 halmaz elegendő.[1] De vajon mi a helyzet magasabb dimenzióban? A térlátásunk itt bizonyára cserbenhagy minket, és nem könnyű látni, hogy hogyan kell előállítani diszjunkt intervallumok egyesítéseként két intervallum különbségét.

Nem meglepő, hogy ha nem látni, hanem kiszámolni akarjuk a jelenséget, akkor mennyivel könnyebb helyzetben vagyunk.

1.1.3. állítás (félgyűrűk szorzata félgyűrű)**.** *Legyen* $(X,\mathcal{P})$ *és* $(Y,\mathcal{Q})$ *egy-egy félgyűrű. Jelölje* $\mathcal{P}\times\mathcal{Q}$ *az* $X\times Y$ *halmaz következő halmazrendszerét:*

$$\mathcal{P}\times\mathcal{Q}=\{P\times Q : P\in\mathcal{P}, Q\in\mathcal{Q}\}\,.$$

Ekkor az $(X\times Y,\mathcal{P}\times\mathcal{Q})$ *pár is félgyűrűt alkot.*

Bizonyítás. A metszetre való zártságot igazolja az

$$(P_1\times Q_1)\cap(P_2\times Q_2)=(P_1\cap P_2)\times(Q_1\cap Q_2)$$

azonosság, ugyanis a $\mathcal{P}$ félgyűrű metszetzártsága miatt $P_1,P_2\in\mathcal{P}$ esetén $P_1\cap P_2\in\mathcal{P}$ is teljesül. Teljesen hasonló módon a $\mathcal{Q}$ halmaz metszetzártsága szerint $Q_1,Q_2\in\mathcal{Q}$ a $Q_1\cap Q_2\in\mathcal{Q}$ tartalmazást implikálja.

A két szorzatelem különbségére vonatkozó állítás pedig az

$$(P_1\times Q_1)\setminus(P_2\times Q_2)=(P_1\times Q_1)\cap(P_2\times Q_2)^c=$$
$$=(P_1\times Q_1)\cap((P_2^c\times Y)\cup(P_2\times Q_2^c))=((P_1\setminus P_2)\times Q_1)\cup((P_1\cap P_2)\times(Q_1\setminus Q_2))$$

azonosságból következik, hiszen a fenti jelölések mellett $P_1\setminus P_2$ és $Q_1\setminus Q_2$ is előáll mint diszjunkt $\mathcal{P}$-, illetve $\mathcal{Q}$-beli halmazok egyesítése. Mivel $P_1\setminus P_2$ diszjunkt a $P_1\cap P_2$ halmaztól, ezért a $(P_1\setminus P_2)\times Q_1$ halmaz is diszjunkt a $(P_1\cap P_2)\times(Q_1\setminus Q_2)$ halmaztól. Ezt kellett belátni. □

Az előző állítást használó indukcióval már látható is az egyik legfontosabb konkrét félgyűrűnk.

1.1.4. következmény. *Az* $\mathbb{R}^n$ *tér balról zárt, jobbról nyílt intervallumainak rendszere félgyűrűt alkot.*

Figyelmesen visszatekintve a bizonyításra, a különbség előállítását igazoló halmaz azonosságból az is rögtön adódik, hogy az $\mathbb{R}^n$ térben két balról zárt, jobbról nyílt intervallum különbségéhez elég legfeljebb 2^n darab ilyen diszjunkt halmaz. Ez igaz is, az persze más kérdés, hogy a fenti egyszerű érvelés nem elég éles ahhoz, hogy a legjobb becslést megkapjuk, ahogyan azt a 3-dimenziós esetben már meggondoltuk.

[1] Képzeljünk el egy kockaodvas-kockát, azaz olyan kockát, amelynek a belsejéből kihagyunk egy kisebb kockát.

A lényeg nem is az, hogy hány diszjunkt intervallumot kell találnunk két halmaz különbségének előállításához, hanem csak annyi, hogy mindig található véges sok ilyen intervallum.

A következő halmazstruktúra szoros kapcsolatban van a félgyűrű fogalmával. Amikor halmazokkal számolunk, kényelmes, ha a különbség művelet nem vezet ki a struktúrából. Sajnos az intervallumok rendszere nem ilyen, hiszen két intervallum különbsége nem feltétlen intervallum.

1.1.5. definíció (gyűrű, σ-gyűrű)**.** Egy $\mathcal{M} \subseteq \mathcal{P}(X)$ halmazrendszert *gyűrűnek* vagy *halmaz gyűrűnek* nevezünk, ha az zárt a halmaz egyesítés és a halmaz különbség műveletekre. Ha zárt a megszámlálható egyesítésre is, akkor *σ-gyűrűnek* nevezzük.

Formálisabban tehát $\mathcal{M}$ gyűrű, ha $A_1, A_2 \in \mathcal{M}$ esetén

1. $A_1 \cup A_2 \in \mathcal{M}$;
2. $A_1 \setminus A_2 \in \mathcal{M}$.

Ha még az is fennáll, hogy egy megszámlálható $\{A_n \in \mathcal{M} : n \in \mathbb{N}\}$ halmazrendszerre

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{M},$$

akkor a gyűrűt σ-gyűrűnek mondjuk.

1.1.6.

Egy gyűrű zárt a véges egyesítésre is, így minden σ-gyűrű egyben gyűrű is. Ha $A, B \in \mathcal{M}$ és $\mathcal{M}$ egy gyűrű, akkor $A \setminus B \in \mathcal{M}$ is teljesül. Újra alkalmazva a különbségre nézve zártságot, de most már az A és az $A \setminus B$ halmazokra kapjuk, hogy

$$A \cap B = A \setminus (A \setminus B) \in \mathcal{M}.$$

Azt látjuk tehát, hogy a különbségre zártság implikálja a metszetzártságot, ezért minden gyűrű egyben félgyűrű is.

Természetesen egy $\mathcal{M}$ gyűrű zárt a véges metszetre is, tehát ha $A_1, ..., A_n \in \in \mathcal{M}$, akkor $\cap_{k=1}^{n} A_k \in \mathcal{M}$ is teljesül. Ha viszont $\mathcal{M}$ egy σ-gyűrű, akkor nem csak a véges, de a megszámlálható metszetre zártság is teljesül. Legyenek ugyanis $A_n \in \mathcal{M}$ minden $n \in \mathbb{N}$ mellett. Ha $Y = \cup_{n=1}^{\infty} A_n$, akkor a megszámlálható egyesítésre való zártság szerint $Y \in \mathcal{M}$. Viszont

$$\cap_{n=1}^{\infty} A_n = Y \setminus \cup_{n=1}^{\infty} (Y \setminus A_n),$$

ami azt jelenti, hogy $\cap_{n=1}^{\infty} A_n \in \mathcal{M}$ is teljesül.

Egy gyűrűelemet sokszor kényelmes úgy tekinteni, mint más gyűrűelemek diszjunkt egyesítése. Például, ha $A_1, A_2 \in \mathcal{M}$ és a szóban forgó gyűrűelem a

$B = A_1 \cup A_2 \in \mathcal{M}$ halmaz, akkor világos, hogy $B = A_1 \cup (A_2 \setminus A_1)$ egy gyűrűbeli, de már diszjunkt előállítás. A következő lemma azt biztosítja, hogy ezt nem csak kételemű egyesítés, hanem akármilyen legfeljebb megszámlálható egyesítés esetén is meg lehet csinálni.

1.1.7. lemma (diszjunktizáció gyűrűben)**.** *Legyen adva az $\mathcal{M}$ gyűrűben egy $\{A_n \in \mathcal{M} : n \in \mathbb{N}\}$ megszámlálható halmazrendszer. Ekkor léteznek $B_n \subseteq A_n, B_n \in \mathcal{M}$ halmazok, amelyek diszjunktak, $\cup_{n=1}^{N} A_n = \cup_{n=1}^{N} B_n$ minden $N \in \mathbb{N}$ mellett, így persze $\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n$.*

Bizonyítás. Legyen

$$B_1 = A_1, \quad B_2 = A_2 \setminus B_1, \dots, \quad B_n = A_n \setminus \cup_{k=1}^{n-1} B_k.$$

Mivel $\mathcal{M}$ gyűrű, ezért $B_n \in \mathcal{M}$, melyekre $B_n \subseteq A_n$ minden $n \in \mathbb{N}$ mellett. A B_n halmazok definíciójuk szerint egymástól diszjunktak. Most N szerinti indukcióval megmutatjuk, hogy $\cup_{n=1}^{N} A_n = \cup_{n=1}^{N} B_n$ is fennáll $\forall N \in \mathbb{N}$ esetén. Az $N = 1$ feltevés mellett ez B_1 konstrukciója. Ha N-re igaz az indukciós feltevés, akkor $N + 1$-re a következő azonosságot írjuk fel.

$$\begin{aligned} \cup_{n=1}^{N+1} A_n &= \left(\cup_{n=1}^{N} A_n\right) \cup A_{N+1} = \left(\cup_{n=1}^{N} B_n\right) \cup A_{N+1} = \\ &= \left(\cup_{n=1}^{N} B_n\right) \cup \left(A_{N+1} \setminus \cup_{n=1}^{N} B_n\right) = \left(\cup_{n=1}^{N} B_n\right) \cup B_{N+1} = \cup_{n=1}^{N+1} B_n. \end{aligned}$$

Ezt kellett belátni. □

Vajon tudunk-e hasonló állítást igazolni abban az esetben, ha $\mathcal{M}$ nem gyűrű, hanem csak félgyűrű! Az olvasó ne nyugodjon addig, amíg a feltett kérdésre igenlő választ nem talál.

Most tovább specializáljuk a gyűrű fogalmát.

1.1.8. definíció (algebra, σ-algebra)**.** Egy $\mathcal{M} \subseteq \mathcal{P}(X)$ halmazrendszert *algebrának* vagy *halmaz algebrának* nevezünk, ha $\mathcal{M}$ olyan gyűrű, amelyre $X \in \mathcal{M}$ is teljesül. Ha $\mathcal{M}$ olyan σ-gyűrű, ami algebra is, akkor $\mathcal{M}$-et *σ-algebrának* nevezzük.

Sokszor kényelmesebb és persze ekvivalens definíció is lehetne az alábbi megfogalmazás.

1.1.9. állítás. *Az $\mathcal{M} \subseteq \mathcal{P}(X)$ halmazrendszer pontosan akkor algebra, ha az* 1.*,* 2. *és a* 3. *feltevések*

1. $X \in \mathcal{M}$;
2. $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$;
3. $A_n \in \mathcal{M}, N \in \mathbb{N}, n = 1, \dots, N \Rightarrow \cup_{n=1}^{N} A_n \in \mathcal{M}$;

3′. $A_n \in \mathcal{M}, n \in \mathbb{N} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{M}$

egyszerre teljesülnek. Hasonlóan az $\mathcal{M}$ pontosan akkor σ-algebra, ha az 1., 2. *és a* $3'$. *feltételek egyszerre teljesülnek.*

Bizonyítás. Tegyük fel először, hogy $\mathcal{M}$ algebra. Tudjuk, hogy M gyűrűként a különbség műveletre zárt (1.1.6), így

$$A^c = X \smallsetminus A \in \mathcal{M}.$$

Most tegyük fel, hogy az állításban előírt három feltételt teljesíti $\mathcal{M}$. Ekkor $\mathcal{M}$ metszetzárt is, hiszen ha $A, B \in \mathcal{M}$, akkor $A^c, B^c \in \mathcal{M}$ a komplementer zártság miatt, majd $A^c \cup B^c \in \mathcal{M}$ az egyesítés művelet zártsága miatt, végül

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{M}$$

újra a komplementer zártság szerint. A különbségre való zártság, a már igazolt metszetzártságból és az

$$A \smallsetminus B = A \cap B^c$$

azonosságból látható.

A σ-algebrára vonatkozó állítás az eddigiekből már nyilvánvaló. □

1.1.10.
Tegyük fel, hogy adva van gyűrűknek egy tetszőleges számosságú halmaza. Legyen tehát rögzítve egy Γ indexhalmaz, és minden $\gamma \in \Gamma$ mellett $\mathcal{M}_\gamma \subseteq X$ egy-egy gyűrű. Képezzük ezen gyűrűk közös részét, azaz definiálja

$$\mathcal{M} = \{A \subseteq X : A \in \mathcal{M}_\gamma, \forall \gamma \in \Gamma\} = \bigcap_{\gamma \in \Gamma} \mathcal{M}_\gamma.$$

Látható, hogy $\mathcal{M}$ is gyűrű. Ugyanis, ha $A, B \in \mathcal{M}$, akkor minden $\gamma \in \Gamma$ mellett $A, B \in \mathcal{M}_\gamma$. No de minden egyes γ-ra $\mathcal{M}_\gamma$ egy gyűrű, emiatt $A \cup B \in$ $\in \mathcal{M}_\gamma$ is teljesül. Mivel ez minden $\gamma \in \Gamma$ mellett igaz, ezért $A \cup B \in \mathcal{M}$ is fennáll, ami azt jelenti, hogy $\mathcal{M}$ zárt az egyesítés műveletre. Analóg érveléssel látjuk azt is, hogy $\mathcal{M}$ a különbség műveletre is zárt, emiatt valóban gyűrű.

Azt láttuk tehát, hogy a

$$(\mathcal{P}(X), \{\mathcal{M} : \mathcal{M} \subseteq \mathcal{P}(X), \mathcal{M} \text{ gyűrű}\})$$

pár egy burokteret definiál. Ez az X halmaz feletti *gyűrűk buroktere*. Emlékezzünk arra, hogy egy buroktérben egy adott halmazt tartalmazó legszűkebb burok-zárt halmaz a halmazt tartalmazó valamennyi burok-zárt halmaz közös része.

1.1.11. definíció (generált gyűrű)**.** Jelölje r az X feletti gyűrűk burokterének lezárási operátorát. Ha tehát $\mathcal{P} \subseteq \mathcal{P}(X)$ egy tetszőleges halmazrendszer[2],

[2] Talán nem szerencsés jelölés, de figyeljünk arra, hogy $\mathcal{P}(X)$ a hatványhalmazt jelöli, és $\mathcal{P}$ ennek egy részhalmaza. Általában a *halmazrendszer* szó alatt valamely adott halmaz hatványhalmazának egy részhalmazát értjük.

akkor $\mathrm{r}(\mathcal{P})$ a $\mathcal{P}$ halmazrendszert tartalmazó *legszűkebb gyűrű*. Ezt az $\mathrm{r}(\mathcal{P})$ halmazrendszert nevezzük a $\mathcal{P}$ halmazrendszer által *generált gyűrűnek*.

A generált gyűrű külső reprezentációja, mint minden buroktér külső reprezentációja teljesen világos: Ha $\mathcal{P}$ egy tetszőleges halmazrendszere az X halmaznak, akkor

$$\mathrm{r}(\mathcal{P}) = \bigcap \{\mathcal{M} \subseteq \mathcal{P}(X) : \mathcal{P} \subseteq \mathcal{M}, \mathcal{M} \text{ gyűrű}\}.$$

De vajon tudunk-e belső reprezentációt is adni? Erre a kérdésre igen a válasz feltéve, hogy a generáló $\mathcal{P}$ halmaz egy félgyűrű.

1.1.12. állítás (félgyűrű generálta gyűrű belső reprezentációja)**.** *Legyen $\mathcal{P}$ egy X-beli halmazokból álló félgyűrű. Ekkor*

$$\mathrm{r}(\mathcal{P}) = \left\{\cup_{n=1}^{N} A_n : A_n \in \mathcal{P}, i \neq j \textit{ mellett } A_i \cap A_j = \varnothing,\ N \in \mathbb{N}\right\}, \tag{1.1}$$

azaz a generált gyűrű a $\mathcal{P}$-beli diszjunkt véges uniók halmaza.

Bizonyítás. Jelölje $\mathcal{H}$ az (1.1) képlet jobb oldalán álló halmazt. Megmutatjuk, hogy $\mathcal{H}$ egy gyűrű. A következő sorrendben érdemes haladnunk: $\mathcal{H}$ halmazrendszer zárt

1. a diszjunkt egyesítésre;
2. a metszetre;
3. a különbség képzésre
4. és az egyesítésre.

Igazoljuk sorjában:
1. A diszjunkt egyesítésre zártság nyilvánvaló.
2. A metszetre zártság az

$$(\cup_n A_n) \cap (\cup_m B_m) = \cup_{n,m}(A_n \cap B_m) \tag{1.2}$$

egyenlőség következménye, hiszen $\mathcal{P}$ félgyűrű volta miatt $A_n \cap B_m \in \mathcal{P}$. Így a már meggondolt diszjunkt egyesítésre való zártságot alkalmazva készen is vagyunk.
3. A különbségre zártsághoz először tekintsük az alábbi átalakítást.

$$\begin{aligned}(\cup_n A_n) \setminus (\cup_m B_m) &= (\cup_n A_n) \cap (\cup_m B_m)^c = (\cup_n A_n) \cap (\cap_m B_m^c) = \\ &= \cup_n (A_n \cap (\cap_m B_m^c)) = \cup_n \cap_m (A_n \cap B_m^c) = \cup_n \cap_m (A_n \setminus B_m).\end{aligned}$$

Mivel $\mathcal{P}$ egy félgyűrű, ezért minden rögzített n és m mellett $A_n \setminus B_m \in \in \mathcal{H}$. A már igazolt metszet zártság szerint $\cap_m (A_n \setminus B_m) \in \mathrm{r}(P)$ fennáll

minden rögzített n mellett. Ilyenek véges diszjunkt egyesítése is $\mathcal{H}$-ban marad 1. szerint.
4. Az egyesítésre zártság innen már könnyű, hiszen $A, B \in \mathcal{H}$ mellett $A \cup \cup B = A \cup (B \setminus A)$. Na most, $B \setminus A \in \mathcal{H}$ a 3. szerint, és az 1. szerint a diszjunkt egyesítés művelet sem vezet ki $\mathcal{H}$-ból.

A fenti pontok igazolása után látjuk, hogy $\mathcal{H}$ valóban egy $\mathcal{P}$-t tartalmazó gyűrű. Amennyiben adott egy másik $\mathcal{P}$-t tartalmazó gyűrű, akkor annak $\mathcal{H}$-t is tartalmaznia kell, hiszen egy gyűrű zárt az egyesítésre. Így azt kapjuk, hogy a $\mathcal{H}$ gyűrű minden $\mathcal{P}$-t tartalmazó gyűrűnek része, ami éppen azt jelenti, hogy $\mathrm{r}(\mathcal{P}) = \mathcal{H}$. □

Rögzítsük most a *mérhető halmaz* szóhasználatot.

1.1.13. definíció (mérhető tér)**.** Az $(X, \mathcal{M})$ párost *mérhető térnek* nevezzük, ha $\mathcal{M}$ egy X feletti σ-algebra. Az $\mathcal{M}$ halmaz elemeit *mérhető halmazoknak* is nevezzük.

Ezek szerint *mérhető halmaznak* lenni semmi mást nem jelent, mint egy adott σ-algebrához való tartozást. Később ettől eltérőnek tűnő koncepciók is megjelennek majd, amelyek voltaképpen ennek speciális esetei, azaz egy-egy speciális σ-algebrához való tartozást fejeznek ki. Ilyenek például egy halmaz *Caratheodory-mérhetősége*, *Borel-mérhetősége* vagy *Lebesgue-mérhetősége*.

Két új burokfogalmat szeretnénk bevezetni, a σ-algebra burkot és a monotonosztály-burkot. Ehhez először a monoton osztály fogalmát kell megértenünk.

1.1.14. definíció (monoton osztály)**.** Az $\mathcal{M} \subseteq \mathcal{P}(X)$ halmazrendszert *monoton osztálynak* mondjuk, ha $\mathcal{M}$ zárt a monoton bővülő halmazok megszámlálható egyesítésére és zárt a monoton szűkülő halmazok megszámlálható metszetére. Formálisabban, az alábbi két feltételnek kell teljesülnie:

1. Ha $A_n \in \mathcal{M}, A_n \subseteq A_{n+1}$ minden $n \in \mathbb{N}$ esetén, akkor $\cup_{n=1}^{\infty} A_n \in \mathcal{M}$ is teljesül.

2. Ha $A_n \in \mathcal{M}, A_n \supseteq A_{n+1}$ minden $n \in \mathbb{N}$ esetén, akkor $\cap_{n=1}^{\infty} A_n \in \mathcal{M}$ is teljesül.

1.1.15. állítás. *Egy halmazrendszer pontosan akkor σ-gyűrű, ha az egyszerre gyűrű és monoton osztály. Hasonlóan, egy halmazrendszer pontosan akkor σ-algebra, ha az egyben algebra és monoton osztály is.*

Bizonyítás. Ha az $\mathcal{M}$ halmazrendszer σ-gyűrű, akkor egyrészt gyűrű, másrészt zárt a megszámlálható egyesítésre és a megszámlálható metszetre is – 1.1.6.

Ha $\mathcal{M}$ egyszerre gyűrű és monoton osztály, akkor tekintsük az

$$\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} \left(\cup_{i=1}^{n} A_i\right)$$

azonosságot. Itt a megszámlálható egyesítés monoton növő gyűrűbeli halmazok egyesítéseként áll elő. Ezt kellett belátni. □

Most bevezetjük a két új burok-operációt.

1.1.16. definíció (generált σ-algebra és monoton osztály)**.** Hasonlóan ahhoz, ahogy 1.1.10.-ben láttuk, hogy a gyűrűk halmaza burokteret alkot, az is könnyen látható, hogy akárhány σ-algebra vagy akárhány monoton osztály metszete is σ-algebra, illetve monoton osztály. Így a σ-algebrának lenni, vagy monoton osztálynak lenni is egy-egy burok-fogalom. Jelölje σ, illetve m a megfelelő burok-operációkat, azaz $\sigma(\mathcal{A})$ az $\mathcal{A}$ halmazrendszert tartalmazó *legszűkebb σ-algebrát* és $\mathrm{m}(\mathcal{M})$ az $\mathcal{M}$-et tartalmazó *legszűkebb monoton osztályt*.

Persze a külső reprezentáció nyilvánvaló: $\sigma(\mathcal{A})$ az $\mathcal{A}$ halmazrendszert tartalmazó összes σ-algebra közös része, és hasonlóan $m(\mathcal{H})$ a $\mathcal{H}$ halmazrendszert tartalmazó összes monoton osztály közös része. Ezen a ponton a természetes kérdés: Hogyan állnak elő a generált σ-algebra elemei a generáló $\mathcal{A}$ halmazrendszer elemeiből halmazelméleti műveletek segítségével? Mi $\sigma(\mathcal{A})$ belső reprezentációja? Lehangoló tény, de tudomásul kell venni, hogy a válasz komplikáltabb, mint gondolnánk, mert $\sigma(\mathcal{A})$-ra és $m(\mathcal{H})$-ra általánosságban nem tudunk kényelmesen használható belső reprezentációt adni. Halmazelméleti ismereteink hiányában nem tudunk olyan „kényelmes algoritmust” adni, amely tetszőleges $\mathcal{A}$ halmazrendszerből kiindulva halmazelméleti műveletekkel felépítené az $\mathcal{A}$ által generált $\sigma(\mathcal{A})$ σ-algebra valamennyi elemét. Definiálja $\mathcal{A}_1$ az $\mathcal{A}$ halmazrendszer elemeiből alkotott megszámlálható egyesítések és azok komplementereinek halmazát. Hasonlóan, $\mathcal{A}_2$ az $\mathcal{A}_1$-beli halmazokból alkotott megszámlálható egyesítések és azok komplementereinek halmazát és így tovább. Az

$$\mathcal{A}_\omega = \bigcup_{n=1}^{\infty} \mathcal{A}_n$$

halmazrendszer ugyan jó nagy részhalmaza $\sigma(\mathcal{A})$-nak de sajnos nem elég nagy. Nem biztos ugyanis, hogy σ-algebra. Ha $E_n \in \mathcal{A}_n \setminus \mathcal{A}_{n-1}$, akkor semmi oka, hogy az $\cup_{n=1}^{\infty} E_n$ halmaz is a fent kiemelt $\mathcal{A}_\omega$ halmazrendszerhez tartozzon.

Kicsivel több halmazelméleti ismerettel meg tudnánk adni a belső reprezentációt is. A fenti gondolatot kell ugyanis folytatni. Jelölje α megszámlálható számosság mellett $\mathcal{A}_\alpha$ az $\mathcal{A}_\beta$ halmazrendszer elemeiből alkotott megszámlálható egyesítések és azok komplementereinek halmazát, ha van α-t közvetlenül megelőző $\beta < \alpha$ számosság. Ha α olyan megszámlálható számosság, amelynek nincs közvetlen megelőzője, akkor legyen

$$\mathcal{A}_\alpha = \bigcup_{\beta<\alpha} \mathcal{A}_\beta,$$

ahol az unióban az összes α számosságnál kisebb megszámlálható számosság jön szóba. Ekkor transzfinit indukcióval megmutatható, hogy

$$\sigma(\mathcal{A}) = \bigcup_{\alpha\in\Omega} \mathcal{A}_\alpha,$$

halmaz a generált σ-algebra belső reprezentációját írja le. Itt Ω az összes megszámlálható számosságok halmaza.

Látni fogjuk, hogy a σ-algebra belső reprezentációjának hiánya nem okoz tárgyalásunk szempontjából igazi nehézséget. Mi több, a belső reprezentáció csak nagyon kevés ponton ad valamit a helyes szemlélet kialakításához. Felesleges lenne pusztán ez okból olyan halmazelméleti kitérőt tenni, amely a fent vázolt belső reprezentációt minden részletében pontosítja. Továbbra is megelégszünk ezért a naiv halmazelméleti ismeretekkel és kerüljük a σ-algebrák belső reprezentációját. Az érdeklődő olvasó [Folland (1999)]-ben pontos leírást talál a szükséges halmazelméleti tudnivalókról.

1.1.17.

Természetesen merül fel, hogy mi a kapcsolat a két imént bevezetett burokoperáció közt. Az első lépésként vegyük észre, hogy az

$$\mathrm{m}(\mathcal{A}) \subseteq \sigma(\mathcal{A})$$

tartalmazás triviálisan teljesül minden $\mathcal{A}$ halmazrendszerre, hiszen minden σ-algebra egyben monoton osztály is (1.1.15), de $\mathrm{m}(\mathcal{A})$ a legszűkebb monoton osztály.

Az előző gondolat fényében különösen érdekes a következő állítás, amely szerint egy gyűrű monoton osztály burka gyűrű marad. Hasonlóan például ahhoz, ahogyan egy konvex halmaz topológiai lezártja is mindig konvex marad.

1.1.18. állítás (gyűrű monoton osztály burka gyűrű (Dynkin))**.** *Legyen* $\mathcal{A} \subseteq \mathcal{P}(X)$ *egy gyűrű. Ekkor* $\mathrm{m}(\mathcal{A})$ *egy gyűrű, ezért egy* σ*-gyűrű is. Ha* $\mathcal{A}$ *egy olyan gyűrű, amelyre* $X \in \mathrm{m}(\mathcal{A})$*, akkor* $\mathrm{m}(\mathcal{A})$ *egy* σ*-algebra, így*

$$\sigma(\mathcal{A}) = \mathrm{m}(\mathcal{A}).$$

Bizonyítás. A gyűrű definíciójának megfelelve azt mutatjuk meg, hogy bármely két $\mathrm{m}(\mathcal{A})$-beli halmaz különbsége és egyesítése is $\mathrm{m}(\mathcal{A})$-beli. Ehhez tetszőlegesen rögzített $B \subseteq X$ mellett tekintsük az

$$\mathcal{A}_B = \{C \subseteq X : B \setminus C, C \setminus B, C \cup B \in \mathrm{m}(\mathcal{A})\}$$

halmazrendszert. Azt kell megmutatnunk, hogy bárhogy rögzítünk $C \in \mathrm{m}(\mathcal{A})$ halmazt, $\mathrm{m}(\mathcal{A}) \subseteq \mathcal{A}_C$ fennáll. Ehhez először is vegyük észre, hogy

1. $C \in \mathcal{A}_B$ pontosan akkor, ha $B \in \mathcal{A}_C$, minden $B, C \subseteq X$ mellett;
2. $\mathcal{A}_B$ monoton osztály minden $B \subseteq X$ mellett;
3. $\mathcal{A} \subseteq \mathcal{A}_B$ minden $B \in \mathcal{A}$ mellett.

Sorjában az indoklások:
1. A nyilvánvaló szimmetria következménye.
2. Az $\mathrm{m}(\mathcal{A})$ monoton osztály tulajdonsága miatt áll fenn. Ugyanis, ha $C_n \in \in \mathcal{A}_B$ egy monoton bővülő halmaz-sorozat, akkor a $B \setminus C_n$ sorozat monoton szűkülő; a $C_n \setminus B$ és a $C_n \cup B$ sorozat monoton bővülő. Persze

$$B \setminus (\cup_{n=1}^{\infty} C_n) = \cap_{n=1}^{\infty} (B \setminus C_n),$$
$$(\cup_{n=1}^{\infty} C_n) \setminus B = \cup_{n=1}^{\infty} (C_n \setminus B),$$
$$(\cup_{n=1}^{\infty} C_n) \cup B = \cup_{n=1}^{\infty} (C_n \cup B).$$

No de $\mathrm{m}(\mathcal{A})$ monoton osztályként zárt a szűkülő metszetre és a bővülő egyesítésre, ezért $\cup_{n=1}^{\infty} C_n \in \mathrm{m}(\mathcal{A})$ is teljesül.
A $\cap_{n=1}^{\infty} C_n \in \mathrm{m}(\mathcal{A})$ hasonlóan adódik. Itt azt kell feltennünk, hogy a C_n halmaz-sorozat szűkülő és az

$$B \setminus (\cap_{n=1}^{\infty} C_n) = \cup_{n=1}^{\infty} (B \setminus C_n),$$
$$(\cap_{n=1}^{\infty} C_n) \setminus B = \cap_{n=1}^{\infty} (C_n \setminus B),$$
$$(\cap_{n=1}^{\infty} C_n) \cup B = \cap_{n=1}^{\infty} (C_n \cup B)$$

azonosságokat kell figyelembe vennünk.
3. Az utolsó tulajdonság pedig azért igaz, mert $\mathcal{A}$ gyűrű. Ugyanis $C, B \in \mathcal{A}$ mellett

$$B \setminus C, C \setminus B, C \cup B \in \mathcal{A} \subseteq \mathrm{m}(\mathcal{A})$$

fennáll.

A fenti három tulajdonságból mindent rögzített $C \in \mathrm{m}(\mathcal{A})$ mellett $\mathrm{m}(\mathcal{A}) \subseteq \subseteq \mathcal{A}_C$ már könnyen adódik. Ugyanis 3. és 2. miatt

$$\mathrm{m}(\mathcal{A}) \subseteq \mathrm{m}(\mathcal{A}_B) = \mathcal{A}_B$$

minden $B \in \mathcal{A}$ mellett, azaz $\forall C \in \mathrm{m}(\mathcal{A})$ és $\forall B \in \mathcal{A}$ esetén $C \in \mathcal{A}_B$, így 1. miatt $B \in \mathcal{A}_C$, azaz $\forall C \in \mathrm{m}(\mathcal{A})$ esetén

$$\mathcal{A} \subseteq \mathcal{A}_C.$$

Újra alkalmazva 2.-t kapjuk, hogy

$$\mathrm{m}(\mathcal{A}) \subseteq \mathcal{A}_C$$

fennáll $\forall C \in \mathrm{m}(\mathcal{A})$ mellett, és épp ezt kellett belátnunk a gyűrű axiómák teljesülésének ellenőrzéséhez.

Ha tehát $\mathrm{m}(\mathcal{A})$ gyűrű, akkor egyben gyűrű is és monoton osztály is, ergo σ-gyűrű is, amint azt már láttuk 1.1.15.-ben.

Ha még az is fennáll, hogy $X \in \mathrm{m}(\mathcal{A})$, akkor $\mathrm{m}(\mathcal{A})$ egy σ-algebra. Mivel az $\mathcal{A}$ halmazrendszert tartalmazó legszűkebb σ-algebra definíció szerint $\sigma(\mathcal{A})$, ezért

$$\sigma(\mathcal{A}) \subseteq \mathrm{m}(\mathcal{A}).$$

Az ellenkező irányú tartalmazás triviálisan teljesül (1.1.17). □

A Dynkin-tételt annak alábbi következményében fogjuk használni.

1.1.19. következmény. *Legyen $\mathcal{P}$ az X alaphalmaz egy olyan halmazrendszere, amelyre $X \in \mathrm{m}(\mathrm{r}(\mathcal{P}))$. Tegyük fel, hogy $\mathcal{O}$ egy olyan monoton osztály, amelyre $\mathrm{r}(\mathcal{P}) \subseteq \mathcal{O}$. Ekkor*

$$\sigma(\mathcal{P}) \subseteq \mathcal{O}.$$

Bizonyítás. Először is $\mathcal{P} \subseteq \mathrm{r}(\mathcal{P})$-ből $\sigma(\mathcal{P}) \subseteq \sigma(\mathrm{r}(\mathcal{P}))$ következik. A $\sigma(\mathrm{r}(\mathcal{P})) = \mathrm{m}(\mathrm{r}(\mathcal{P}))$ egyenlőség a Dynkin-tétel következménye az $\mathrm{r}(\mathcal{P})$ gyűrűre alkalmazva. Mivel $\mathrm{r}(\mathcal{P}) \subseteq \mathcal{O}$, ezért $\mathrm{m}(\mathrm{r}(\mathcal{P})) \subseteq \mathrm{m}(\mathcal{O})$. Végül $\mathcal{O}$ monoton osztály volta szerint $\mathrm{m}(\mathcal{O}) = \mathcal{O}$. Összefoglalva tehát

$$\sigma(\mathcal{P}) \subseteq \sigma(\mathrm{r}(\mathcal{P})) = \mathrm{m}(\mathrm{r}(\mathcal{P})) \subseteq \mathrm{m}(\mathcal{O}) = \mathcal{O}.$$

Ezt kellett belátni. □

1.1.20. (σ-algebra indukció)
Tegyük fel, hogy adott egy $\mathcal{M}$ σ-algebra és egy T tulajdonság, amely a σ-algebra halmazain van értelmezve. A T vagy igaz egy $A \in \mathcal{M}$ halmazon vagy sem. Most egy sokat használt bizonyítási módszert gondolunk át, melyet nevezhetnénk a teljes indukció analógiájára mértékelméleti indukciónak is. Tegyük fel, hogy az állításunk az, hogy a T tulajdonság igaz az $\mathcal{M}$ σ-algebra minden elemére. Legyen $\mathcal{P}$ olyan halmazrendszer, amely generálja $\mathcal{M}$-et, azaz

$$\sigma(\mathcal{P}) = \mathcal{M}.$$

Most is két lépést kell végrehajtanunk.

(i) Mutassuk meg, hogy $\mathcal{P}$ minden halmazára T igaz.

(ii) Definiálja

$$\mathcal{H} = \{A \in \mathcal{M} : T \text{ igaz } A\text{-ra}\}.$$

Mutassuk meg, hogy $\mathcal{H}$ egy σ-algebra!

Látható, hogy ha a fenti (i) és (ii) teljesül, akkor T tulajdonság az $\mathcal{M}$ σ-algebra minden $A \in \mathcal{M}$ elemén is igaz. Ugyanis (i) azt jelenti, hogy $\mathcal{P} \subseteq \mathcal{H}$, és (ii) azt jelenti, hogy $\sigma(\mathcal{H}) = \mathcal{H}$. Összefoglalva

$$\mathcal{M} = \sigma(\mathcal{P}) \subseteq \sigma(\mathcal{H}) = \mathcal{H}.$$

Ez pontosan azt jelenti, hogy minden $A \in \mathcal{M}$ halmaz kielégíti $\mathcal{H}$ definícióját, azaz T tulajdonság igaz minden mérhető halmazra, az az $\mathcal{M}$ minden elemére.

1.1.21. (monoton osztály indukció)
Az 1.1.19. következmény segítségével ez a bizonyítási módszer erősíthető. Legyen most is T, egy az $\mathcal{M}$ σ-algebra elemein értelmezett tulajdonság, és $\mathcal{P}$ egy generáló halmazrendszer, tehát

$$\sigma(\mathcal{P}) = \mathcal{M}.$$

Tegyük fel még, hogy $X \in \mathrm{m}(\mathrm{r}(\mathcal{P}))$. A monoton osztályra módosított mértékelméleti indukció a következő két lépés végrehajtását jelenti.

(i) Mutassuk meg, hogy $\mathrm{r}(\mathcal{P})$ minden halmazára T igaz.

(ii) Definiálja
$$\mathcal{O} = \{A \in \mathcal{M} : T \text{ igaz } A\text{-ra}\}$$
Mutassuk meg, hogy $\mathcal{O}$ egy monoton osztály.

Most is látható, hogy ha a fenti (i) és (ii) teljesül, akkor T tulajdonság az $\mathcal{M}$ σ-algebra minden $A \in \mathcal{M}$ elemén is igaz. Ugyanis (i) azt jelenti, hogy $\mathrm{r}(\mathcal{P}) \subseteq \mathcal{O}$, ezt 1.1.19.-re alkalmazva azt kapjuk, hogy

$$\mathcal{M} = \sigma(\mathcal{P}) \subseteq \mathcal{O}.$$

Ez éppen azt jelenti, hogy minden $A \in \mathcal{M}$ halmaz kielégíti $\mathcal{O}$ definícióját, ezért a T tulajdonság igaz a σ-algebra minden elemére.

Összefoglalva az előző két bizonyítási eljárást az alábbi állításokat kapjuk. Mindkét állítás a σ-algebrák belső reprezentációjának pótlására szolgál. Az első állításhoz csak a lezárási operátor egyszerű tulajdonságai kellettek, míg a második állítás már a mélyebb 1.1.19. következményen alapszik.

1.1.22. állítás (σ-indukció)**.** *Legyen $\mathcal{M}$ egy σ-algebra és $\mathcal{P}$ egy olyan halmazrendszer az X alaphalmaz felett, amelyre*

$$\sigma(\mathcal{P}) = \mathcal{M}$$

teljesül. Legyen T egy olyan az $\mathcal{M}$ halmazain értelmezett tulajdonság, amelyre az alábbi feltevések fennállnak.

1. *A $\mathcal{P}$ halmazrendszer minden elemén T igaz.*
2. a) *Az üreshalmazra T igaz.*
   b) *Ha $A \in \mathcal{M}$ halmazra T igaz, akkor T az A^c halmazra is igaz.*
   c) *Ha T igaz az $A_n \in \mathcal{M}$, $n \in \mathbb{N}$ megszámlálhatóan sok halmazra, akkor azok $\cup_{n=1}^{\infty} A_n$ egyesítésére is igaz.*

Ekkor T az $\mathcal{M}$ σ-algebra minden elemére igaz.

A második állítás voltaképpen a fenti indukció erősítése. Itt fontos, hogy egy félgyűrűből kell kiindulnunk.

1.1.23. állítás (m-indukció)**.** *Legyen $\mathcal{M}$ egy σ-algebra és $\mathcal{P}$ egy olyan félgyűrű az X alaphalmaz felett, amelyre*

$$\sigma(\mathcal{P}) = \mathcal{M}$$

teljesül, és X előáll mint $\mathcal{P}$-beli halmazok diszjunkt, legfeljebb megszámlálható egyesítése. Legyen T egy olyan, az $\mathcal{M}$ halmazain értelmezett tulajdonság, amelyre az alábbi feltevések fennállnak.

1. *Véges sok diszjunkt, $\mathcal{P}$-beli halmaz egyesítésére T igaz.*
2. a) *Ha az $A_n \in \mathcal{M}$ monoton bővülő halmaz-sorozat minden elemére T igaz, akkor T az $\cup_{n=1}^{\infty} A_n$ egyesítésre is igaz.*
   b) *Ha az $A_n \in \mathcal{M}$ monoton szűkülő halmaz-sorozat minden elemére T igaz, akkor T a $\cap_{n=1}^{\infty} A_n$ metszetre is igaz.*

Ekkor T az $\mathcal{M}$ σ-algebra minden elemére igaz.

A σ-indukció első alkalmazásaként nézzük, hogy a szorzat σ-algebrának hogyan tudjuk megadni egy generátor rendszerét. Ha adottak az $\mathcal{M}$ és az $\mathcal{N}$ σ-algebrák, akkor ezek szorzata csak érdektelen speciális esetekben σ-algebra. Mivel minden σ-algebra egyben félgyűrű is, ezért csak annyit mondhatunk, hogy az $\mathcal{M} \times \mathcal{N}$ szorzat félgyűrűk szorzataként maga is félgyűrű (1.1.3).

1.1.24. definíció (mérhető terek szorzata)**.** Legyenek $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető terek. E mérhető térek szorzata az $(X \times Y, \sigma(\mathcal{M} \times \mathcal{N}))$ mérhető tér. A szorzat σ-algebra tehát az

$$\mathcal{M} \times \mathcal{N} = \{M \times N : M \in \mathcal{M}, N \in \mathcal{N}\}$$

mérhető téglák félgyűrűje által generált σ-algebra. A szorzat σ-algebrát szokás $\mathcal{M} \otimes \mathcal{N}$ módon is jelölni.

1.1.25. állítás (generátorok szorzata a szorzat generátora)**.** *Legyen* $(X, \mathcal{M})$ *és* $(Y, \mathcal{N})$ *mérhető tér, melyekre a* $\mathcal{P}$ *halmazrendszer az* $\mathcal{M}$ σ*-algebrát, míg a* $\mathcal{Q}$ *halmazrendszer az* $\mathcal{N}$ σ*-algebrát generálja. Ekkor a*

$$(\mathcal{P} \times \{Y\}) \cup (\{X\} \times \mathcal{Q}) = \{P \times Y : P \in \mathcal{P}\} \cup \{X \times Q : Q \in \mathcal{Q}\} \tag{1.3}$$

halmazrendszer a $\sigma(\mathcal{M} \times \mathcal{N})$ σ*-algebrát generálja. Ha feltesszük még, hogy* X *előáll legfeljebb megszámlálhatóan sok* $\mathcal{P}$*-beli, és* Y *legfeljebb megszámlálhatóan sok* $\mathcal{Q}$*-beli halmaz egyesítéseként, akkor a*

$$\mathcal{P} \times \mathcal{Q} = \{P \times Q : P \in \mathcal{P}, Q \in \mathcal{Q}\}$$

halmazrendszer is generálja $\sigma(\mathcal{M} \times \mathcal{N})$*-et. Ekkor tehát*

$$\sigma(\mathcal{P} \times \mathcal{Q}) = \sigma(\sigma(\mathcal{P}) \times \sigma(\mathcal{Q})).$$

Bizonyítás. Jelölje $\mathcal{F}$ a tétel kimondásában szereplő halmazrendszert – (1.3). Mivel $\mathcal{P} \subseteq \mathcal{M}$ és $Y \in \mathcal{N}$, ezért $\mathcal{P} \times \{Y\} \subseteq \mathcal{M} \times \mathcal{N}$. Hasonlóan $\{X\} \times \mathcal{Q} \subseteq \mathcal{M} \times \mathcal{N}$, ezért $\sigma(\mathcal{F}) \subseteq \sigma(\mathcal{M} \times \mathcal{N})$.

A fordított irányú tartalmazás igazolásához azt gondoljuk meg, hogy az $\mathcal{M} = \sigma(\mathcal{P})$ σ-algebra minden eleme az

$$\{E \subseteq X : E \times Y \in \sigma(\mathcal{P} \times \{Y\})\} \tag{1.4}$$

halmazhoz tartozik. Ezt σ-indukcióval mutatjuk meg (1.1.22).

1. A $\mathcal{P}$ halmazrendszer elemeire (1.4) nyilvánvalóan teljesül.
2. a) Világos, hogy $\varnothing \times Y = \varnothing \in \sigma(\mathcal{P} \times \{Y\})$.
   b) Ha az $A \subseteq X$ halmazra $A \times Y \in \sigma(\mathcal{P} \times \{Y\})$, akkor $A^c \times Y = (A \times Y)^c \in \sigma(\mathcal{P} \times \{Y\})$ is teljesül.
   c) Ha az $A_n \subseteq X$ halmazok az (1.4) halmaz elemei, akkor $(\cup_{n=1}^{\infty} A_n) \times Y = \cup_{n=1}^{\infty} (A_n \times Y)$ is a fenti halmazhoz tartozik.

Megmutattuk tehát, hogy

$$\sigma(\mathcal{P}) \times \{Y\} \subseteq \sigma(\mathcal{P} \times \{Y\}) \subseteq \sigma(\mathcal{F}).$$

Szimmetrikus okoskodás mutatja, hogy

$$\{X\} \times \sigma(\mathcal{Q}) \subseteq \sigma(\{X\} \times \mathcal{Q}) \subseteq \sigma(\mathcal{F})$$

is fennáll. Ha $M \in \mathcal{M}$ és $N \in \mathcal{N}$, akkor $M \times N = (M \times Y) \cap (X \times N)$. A gyűrűk metszetzártsága miatt az

$$\mathcal{M} \times \mathcal{N} \subseteq \sigma(\mathcal{F})$$

azonosság is fennáll, amiből $\sigma(\mathcal{M} \times \mathcal{N}) \subseteq \sigma(\mathcal{F})$ már könnyen következik.

Most tegyük fel, hogy $Y = \cup_{n=1}^{\infty} Q_n$, valamely $Q_n \in \mathcal{Q}$ halmazokra. Ekkor egy $P \in \mathcal{P}$ halmaz mellett $P \times Y = \cup_{n=1}^{\infty} (P \times Q_n) \in \sigma(\mathcal{P} \times \mathcal{Q})$. Hasonlóan, ha az X alaphalmaz $\mathcal{P}$-beli megszámlálható egyesítés, akkor $X \times Q \in \in \sigma(\mathcal{P} \times \mathcal{Q})$ minden $Q \in \mathcal{Q}$ mellett. Így $\mathcal{F} \subseteq \sigma(\mathcal{P} \times \mathcal{Q})$, ezért $\sigma(\mathcal{F}) \subseteq \subseteq \sigma(\mathcal{P} \times \mathcal{Q})$. Mivel $\mathcal{P} \times \mathcal{Q}$ minden eleme előáll két $\mathcal{F}$-beli elem metszeteként, ezért $\mathcal{P} \times \mathcal{Q} \subseteq \sigma(\mathcal{F})$, emiatt $\sigma(\mathcal{P} \times \mathcal{Q}) \subseteq \sigma(\mathcal{F})$ is fennáll. Ezt kellett belátni. □

1.1.26.
A fenti gondolat szinte szó szerinti ismétlésével megszámlálhatóan sok vagy véges sok σ-algebra szorzatára analóg állítást kapunk. Ha adottak az $(X_n, \mathcal{M}_n)$ mérhető terek a $\sigma(\mathcal{P}_n) = \mathcal{M}_n$ generáló halmazrendszerekkel, akkor a

$$\left\{ \prod_{n=1}^{\infty} P_n : P_n = X_n \forall n \neq k, P_k \in \mathcal{P}_k; k \in \mathbb{N} \right\}$$

típusú úgynevezett *cilinder halmazok* generálják a $\otimes_{n=1}^{\infty} \mathcal{M}_n = \sigma\left(\times_{n=1}^{\infty} \mathcal{M}_n\right)$ σ-algebrát. Ha még azt is feltesszük, hogy minden X_n előáll mint megszámlálhatóan sok $\mathcal{P}_n$-beli halmaz egyesítése, akkor a

$$\left\{ \prod_{n=1}^{\infty} P_n : P_n \in \mathcal{P}_n; n \in \mathbb{N} \right\}$$

téglák is generálját a szorzat σ-algebrát. Ha több, mint megszámlálhatóan sok σ-algebra szorzatára van szükség, akkor csak a cilinderekből álló halmazrendszerre támaszkodhatunk mint a szorzat σ-algebra generátor rendszerére.

Most az analízis szemszögéből nézett két legfontosabb σ-algebra közül a szűkebbet vezetjük be.

1.1.27. definíció (Borel-halmaz)**.** Legyen (X, τ) egy topologikus tér. A τ nyílt halmazok halmaza, mint halmazrendszer generálta σ-algebrát *Borel-halmazok* σ-algebrájának nevezzük. E σ-algebra elemeit *Borel-halmazoknak*, vagy *Borel-mérhető* halmazoknak mondjuk.

A Borel-halmaz kifejezés tehát csak úgy értelmes, ha a téren adott egy topológia is. Sokszor fordul elő, hogy a szövegkörnyezetben nyilvánvaló, hogy melyik topológiára gondolunk. Például a számegyenes Borel-halmazai kifejezés használatakor a szokásos euklideszi topológiára kell gondolnunk. Hasonlóan véges dimenziós vektorterek esetén is a Borel-halmazok halmaza a nyílt halmazokat tartalmazó legszűkebb σ-algebra.

1.1.28.
Látható, hogy egy topologikus tér nyílt halmazai és ugyanezen tér zárt halmazai azonos σ-algebrát generálnak. Ugyanis minden zárt halmaz egy nyílt

halmaz komplementereként Borel-halmaz, így a zártak generálta σ-algebra minden eleme is Borel-halmaz. Hasonlóan, minden nyílt halmaz egy alkalmas zárt halmaz komplementereként a zártak generálta σ-algebra egy eleme, ezért a Borel-halmazok sem eshetnek a zárt halmazok generálta σ-algebrán kívül. Ezt úgy is kifejezhetjük, hogy a Borel-halmazokat a zárt halmazok is generálják.

1.1.29. állítás (Borel-σ-algebrák szorzata is Borel-σ-algebra)**.** *Jelölje $\mathcal{B}_{\mathbb{R}^n}$ az n-dimenziós euklideszi tér Borel-halmazait. $n = s + t$. Ekkor*

$$\mathcal{B}_{\mathbb{R}^n} = \mathcal{B}_{\mathbb{R}^s} \otimes \mathcal{B}_{\mathbb{R}^t}, \textit{ speciálisan } \otimes_1^n \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^n}.$$

Bizonyítás. Tudjuk, hogy a szorzat σ-algebra egy generátora a generáló nyílt halmazok szorzata (1.1.25). Persze egy s- és egy t-dimenziós nyílt halmaz szorzata is nyílt részhalmaza az n-dimenziós térnek, ezért

$$\begin{aligned} \mathcal{B}_{\mathbb{R}^s} \otimes \mathcal{B}_{\mathbb{R}^t} = \sigma \left\{ U \times V : U \subseteq \mathbb{R}^s, V \subseteq \mathbb{R}^t, U, V \text{ nyílt} \right\} \subseteq \\ \subseteq \sigma \left(\{ W \subseteq \mathbb{R}^n : W \text{ nyílt} \} \right) = \mathcal{B}_{\mathbb{R}^n}. \end{aligned}$$

No de az $\mathbb{R}^n$ térben minden nyílt halmaz előáll mint racionális középpontú, és racionális oldalhosszú nyílt kockák egyesítése. Ilyenből legfeljebb megszámlálhatóan sok van, tehát minden nyílt halmaz hozzátartozik a nyílt kockák generálta σ-algebrához. Ebből fent kiemelt sor fordított irányú tartalmazása következik. □

1.2. Mértéktér és legegyszerűbb tulajdonságai

1.2.1. definíció (végesen additív, σ-additív halmazfüggvény)**.** Legyen $\mathcal{H} \subseteq \subseteq \mathcal{P}(X)$ egy nem üres halmazrendszer, és μ egy nem negatív valós értékeket felvevő halmazfüggvény, amelynek értelmezési tartománya $\mathcal{H}$, azaz

$$\mu : \mathcal{H} \to \overline{\mathbb{R}}_+.$$

Azt mondjuk, hogy μ halmazfüggvény

– *additív*, vagy *végesen additív*, ha fennáll az

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$$

egyenlőség minden olyan $A_1, A_2 \subseteq X$ halmazra, melyekre:

1. $A_1, A_2 \in \mathcal{H}$;
2. $A_1 \cup A_2 \in \mathcal{H}$;
3. $A_1 \cap A_2 = \varnothing$.

– *σ-additív*, vagy *megszámlálhatóan additív*, ha fennáll az

$$\mu\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right) \tag{1.5}$$

egyenlőség minden olyan $\{A_n \subseteq X : n \in \mathbb{N}\}$ halmazrendszerre, amelyre:

1. $A_n \in \mathcal{H}, \forall n \in \mathbb{N}$;
2. $\cup_{n=1}^{\infty} A_n \in \mathcal{H}$;
3. $A_n \cap A_m = \varnothing, \forall n, m \in \mathbb{N}, n \neq m$.

Fontos megértenünk, hogy az additivitás és a σ-additivitás definíciója semmit nem ír elő olyan diszjunkt halmazokra, amelyek egyesítése a $\mathcal{H}$ értelmezési tartományon kívül esik. A természetes értelmezési tartománya egy végesen additív halmazfüggvénynek egy gyűrű. Ekkor egyszerűen úgy fogalmazhatunk, hogy bármely két $A_1, A_2 \in \mathcal{H}$ diszjunkt halmaz mellett $\mu\left(A_1 \cup A_2\right) = \mu\left(A_1\right) + \mu\left(A_2\right)$. Hasonlóan, ha $\mathcal{H}$ legalább σ-gyűrű, akkor egy $A_n \in \mathcal{H}$ megszámlálható halmazsorozat mellett nincs szükség az $\cup_{n=1}^{\infty} A_n \in \mathcal{H}$ feltétel megkövetelésére, hiszen az automatikusan teljesül. A definíció így egyszerűbben is fogalmazható azzal, hogy a fenti 2. pontot el lehet hagyni.

A felsőhatár-axiómával kapcsolatos feltevésünknek (lásd 0.1.1) megfelelően az (1.5)-beli sor mindenképpen értelmes, hiszen a részletösszegek monoton növekvő sorozatot alkotnak, így a sorösszeg mint a részletösszegek szuprémuma, vagy egy véges valós szám, vagy $+\infty$. Tudjuk, hogy egy nem negatív tagú valós sor nem érzékeny a sorozat elemeinek átrendezésére, így akármilyen sorrendben is írjuk fel a halmazrendszer elemeit (1.5)-ben, mind a bal oldal, mind a jobb oldal ugyanazt a valós számot vagy $+\infty$-t adja.

1.2.2. definíció (mértéktér)**.** Az $(X, \mathcal{M}, \mu)$ hármast *mértéktérnek nevezzük*, ha $(X, \mathcal{M})$ mérhető tér, $\mu : \mathcal{M} \to \overline{\mathbb{R}}_+$ a konstans $+\infty$-től különböző, nem negatív, σ-additív halmazfüggvény. Ilyenkor a μ halmazfüggvényt *mértéknek* mondjuk. Ha az alaphalmaz mértéke véges, akkor azt mondjuk, hogy $(X, \mathcal{M}, \mu)$ *véges mértéktér*.

Azt mondhatjuk tehát, hogy az $(X, \mathcal{M}, \mu)$ hármas mértéktér, ha $\mathcal{M}$ az X alaphalmaz egy σ-algebrája; μ olyan függvény, melynek értelmezési tartománya $\mathcal{M}$; $A_n \in \mathcal{M}, A_n \cap A_m = \varnothing, n, m \in \mathbb{N}, n \neq m$ mellett $\mu\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right)$, és legalább egy $A \in \mathcal{M}$ halmazra $\mu(A) < \infty$. Látható, hogy a konstans halmazfüggvények közül csak a konstans $+\infty$ és a konstans zérus teljesíti a σ-additivitás feltételét. A konstans $+\infty$ esetet mint érdektelent zárjuk ki.

1.2.3.

Időnként szükség van a mérték fentinél általánosabb értelmezéseire is, de azokat mindig külön hangsúlyozzuk. Néha mértéknek mondunk egy olyan nem konstans halmazfüggvényt, amely σ-additív, de nem feltétlen egy σ-algebrán értelmezett. Például a 3. fejezetben félgyűrűn értelmezett mértéket terjesztünk ki, egy, a félgyűrűt tartalmazó σ-algebrára. Ha nem egy σ-algebra a mérték értelmezési tartománya, akkor azt explicit jelöljük.

Szükség lehet például arra is, hogy a μ halmazfüggvény értékeit az itt szereplő $\overline{\mathbb{R}}_+$ helyett egy másik halmazból vegyük. Például *előjeles mértéknek* nevezünk egy σ-algebrán értelmezett $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ halmazfüggvényt, ha a $+\infty$ és a $-\infty$ közül képként csak az egyiket állítja elő; legalább egy halmaz mértéke valós; valamint tetszőleges $A = \cup_{n=1}^{\infty} A_n, A_n \in \mathcal{H}$ diszjunkt egyesítés a $\mu(A) = \sum_{n=1}^{\infty} \mu(A_n)$ egyenlőséget implikálja. Hasonló a $\mu : \mathcal{M} \to \mathbb{C}$ *komplex mérték* definíciója is.

1.2.4.
Emlékezzünk arra a tényre, hogy ha $a_n \in \mathbb{C}$ komplex számok, akkor annak szükséges és elegendő feltétele, hogy minden $\pi : \mathbb{N} \to \mathbb{N}$ permutáció mellett a $\sum_{n=1}^{\infty} a_{\pi(n)}$ sor a π permutációtól függetlenül konvergens és értéke azonos legyen, éppen a $\sum_{n=1}^{\infty} a_n$ sor abszolút konvergenciája. Mivel egy halmazrendszer egyesítése független az elemek sorrendjétől, ezért komplex- vagy véges előjeles mérték esetén az

$$A = \cup_{n=1}^{\infty} A_n \implies \mu(A) = \sum_{n=1}^{\infty} \mu(A_n)$$

implikáció megkövetelése a

$$\sum_{n=1}^{\infty} |\mu(A_n)| < \infty$$

teljesülését is magában foglalja, ami (nem véges) mérték esetén nincs így.

Látjuk tehát, hogy a komplex mérték vagy az előjeles mérték, némileg ellentmondva a jelzős nyelvi szerkezetnek, nem speciális mérték.

Ha arra külön utalást nem teszünk, akkor mértéken egy σ-algebrán értelmezett, nem negatív, σ-additív halmazfüggvényt fogunk érteni, amelynek értékkészletében a $+\infty$ ugyan szerepelhet, de legalább egy halmaz mértéke ettől különböző.

1.2.5. állítás (mérték végesen additív, monoton és szubtraktív)**.** *Legyen* $(X, \mathcal{M}, \mu)$ *egy mértéktér. Ekkor*

1. $\mu(\varnothing) = 0$.
2. μ *végesen additív.*
3. $A, B \in \mathcal{M}, A \subseteq B$ *esetén* $\mu(A) \leq \mu(B)$, *azaz a mérték* monoton.

4. *Ha $A, B \in \mathcal{M}$, $A \subseteq B$ és $\mu(A) < \infty$, akkor $\mu(B \setminus A) = \mu(B) - \mu(A)$, azaz a mérték* szubtraktív.

Bizonyítás. Az egyes tulajdonságok sorban:

1. Az üreshalmaz előáll megszámlálhatóan sok diszjunkt üreshalmaz egyesítéseként, így a σ-additivitás szerint $\mu(\varnothing) = \sum_{n=1}^{\infty} \mu(\varnothing) = \lim_{n\to\infty}(n\mu(A_n))$. Ebből az következik, hogy $\mu(\varnothing) = 0$, vagy $\mu(\varnothing) = +\infty$. Ez utóbbi esetben tetszőleges A mérhető halmazra $A = A \cup \varnothing \ldots \cup \varnothing \ldots$ megszámlálhatóan sok diszjunkt halmazból álló egyesítés, ezért a σ-additivitás miatt $\mu(A) = = +\infty$ lenne.
2. Ha $A = \cup_{i=1}^{n} A_i$ véges sok diszjunkt halmaz egyesítése, akkor az is igaz, hogy $A = \cup_{i=1}^{n} A_i \cup \varnothing \ldots \cup \varnothing \ldots$ megszámlálhatóan végtelen sok diszjunkt halmaz egyesítése. Tudva, hogy $\mu(\varnothing) = 0$, kapjuk a $\mu(\cup_{i=1}^{n} A_i) = = \sum_{i=1}^{n} \mu(A_i)$ egyenlőséget.
3. Nyilván $B = A \cup (B \setminus A)$ diszjunkt egyesítés, így a végesen additivitást kihasználva $\mu(B) = \mu(A) + \mu(B \setminus A)$. No de μ nem negatív értékeket vesz csak fel, tehát $\mu(A) \leq \mu(B)$,
4. másrészt ha $\mu(A) < \infty$, akkor a $\mu(B) - \mu(A)$ kifejezés értelmes és egyenlő $\mu(B \setminus A)$-val. □

1.2.6. állítás (a mérték σ-szubadditív)**.** *Az $(X, \mathcal{M}, \mu)$ mértéktér tetszőleges $\{A_n \subseteq X : A_n \in \mathcal{M}, n \in \mathbb{N}\}$ halmazrendszerére*

$$\mu(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Bizonyítás. Legyen $B_n \subseteq A_n$ a diszjunktizált halmaz-sorozat (1.1.7). Ekkor a mérték σ-additivitása és monotonitása miatt:

$$\mu(\cup_{n=1}^{\infty} A_n) = \mu(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n). \qquad \square$$

1.2.7. (gyűrű esete, végesen additív eset)
Gondoljuk végig, hogy ez előző két állítás (1.2.5 és 1.2.6) szó szerint érvényben marad akkor is, ha a mérték értelmezési tartománya csak gyűrű. Ennek oka, hogy a diszjunktizációs lemmát (1.1.7) gyűrűben mondtuk ki.

A fentivel analóg módon látszik, hogy amennyiben $\mathcal{M}$ egy gyűrű, és $\mu : \mathcal{M} \to \overline{\mathbb{R}}_+$ egy végesen additív halmazfüggvény, akkor μ monoton és szubadditív is. Ez azt jelenti, hogy $A, B \in \mathcal{M}, A \subseteq B$ esetén $\mu(A) \leq \mu(B)$ és tetszőleges, de csak véges sok $A_1, \ldots, A_N \in \mathcal{M}$ esetén

$$\mu\left(\cup_{n=1}^{N} A_n\right) \leq \sum_{n=1}^{N} \mu(A_n).$$

A következő állítás a σ-additivitás egyik kulcs következménye, amit *monoton folytonosságnak* szokás nevezni.

1.2.8. állítás (mérték monoton folytonossága)**.** *Az $(X, \mathcal{M}, \mu)$ mértéktérben legyen $A_n \in \mathcal{M}$ mérhető halmazok egy sorozata.*

1. *Ha az A_n sorozat monoton bővülő, azaz $A_n \subseteq A_{n+1}$ minden $n \in \mathbb{N}$ mellett, akkor*
$$\mu\left(\cup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu\left(A_n\right).$$
2. *Ha az A_n sorozat monoton szűkülő, azaz $A_{n+1} \subseteq A_n$ minden $n \in \mathbb{N}$ mellett, és $\mu\left(A_1\right) < \infty$, akkor*
$$\mu\left(\cap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu\left(A_n\right).$$

Bizonyítás. Először a monoton bővülő eset, majd erre visszavezetve a szűkülő halmazok esete:

1. Legyen B_n a diszjunktizált halmaz-sorozat 1.1.7. szerint. Ekkor a végtelen sor definíciója, és az A_n halmaz-sorozat bővülése miatt:
$$\mu\left(\cup_{n=1}^{\infty} A_n\right) = \mu\left(\cup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu\left(B_n\right) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu\left(B_k\right) =$$
$$= \lim_{n\to\infty} \mu\left(\cup_{k=1}^{n} B_k\right) = \lim_{n\to\infty} \mu\left(\cup_{k=1}^{n} A_k\right) = \lim_{n\to\infty} \mu\left(A_n\right),$$
hiszen $\cup_{k=1}^{n} A_k = A_n$.
2. Most tegyük fel, hogy A_n halmaz-sorozat monoton szűkülő, így minden tagja véges mértékű a $\mu\left(A_n\right) \le \mu\left(A_1\right) < \infty$ szerint. Jelölje $B_n = A_1 \setminus A_n$. Világos, hogy $B_n \in \mathcal{M}$ és B_n monoton bővülő. Az is világos, hogy $\cup_n B_n = \cup_n (A_1 \setminus A_n) = A_1 \setminus \left(\cap_n A_n\right)$, ezért
$$\mu\left(B_n\right) \to \mu\left(A_1 \setminus \cap_n A_n\right).$$
No de $\mu\left(A_n\right) < \infty$ miatt a mérték szubtraktív, így
$$\mu\left(A_1\right) - \mu\left(A_n\right) = \mu\left(A_1 \setminus A_n\right) = \mu\left(B_n\right) \to$$
$$\to \mu\left(A_1 \setminus \cap_n \left(A_n\right)\right) = \mu\left(A_1\right) - \mu\left(\cap_n A_n\right).$$
Újra kihasználva $\mu\left(A_1\right)$ végességét, azt kapjuk, hogy $\mu\left(A_n\right) \to \mu\left(\cap_n A_n\right)$. □

A mérték monoton folytonossága, azaz a fenti tétel 1. pontja az a tulajdonság, amely a végesen additív halmazfüggvényt megkülönbözteti a σ-additív halmazfüggvénytől. Látható ugyanis, hogy egy σ-algebrán értelmezett halmazfüggvény pontosan akkor mérték, ha az egyszerre végesen additív és monoton folytonos.

A mérték konstrukciós eljárás kulcsfontosságú része az anyagnak. Ennek legelemibb része, hogy a félgyűrűről gyűrűre kiterjesszünk egy additív vagy σ-additív halmazfüggvényt. Ez a probléma lényegesen különbözik a generált σ-algebrára való kiterjesztéstől, hiszen itt ismerjük a generált gyűrű belső reprezentációját.

1.2.9. állítás (elemi kiterjesztési tétel)**.** *Legyen* $\mu : \mathcal{P} \to \overline{\mathbb{R}}_+$ *a* $\mathcal{P}$ *félgyűrűn értelmezett megszámlálhatóan additív halmazfüggvény. Ekkor* μ *egyértelműen terjeszthető ki a generált* $\mathrm{r}(\mathcal{P})$ *gyűrűre úgy, hogy a kiterjesztett* $\hat{\mu} : \mathrm{r}(\mathcal{P}) \to \overline{\mathbb{R}}_+$ *halmazfüggvény is megszámlálhatóan additív maradjon.*

Bizonyítás. Olyan $\hat{\mu} : \mathrm{r}(\mathcal{P}) \to \overline{\mathbb{R}}_+$ halmazfüggvényt keresünk, amely σ-additív és minden $A \in \mathcal{P}$ mellett

$$\hat{\mu}(A) = \mu(A).$$

A σ-additivitás szerint csak $\mu(\varnothing) = 0$, vagy $\mu(\varnothing) = +\infty$ lehetséges. Ez utóbbi esetben μ a konstans $+\infty$ függvény, amelyre az állítás nyilvánvaló.

Először megmutatjuk azt, hogy amennyiben $A_n, B_m \in \mathcal{P}$ olyan halmazok, amelyekre $\cup_{n=1}^{\infty} A_n = \cup_{m=1}^{\infty} B_m$ diszjunkt egyesítések, akkor $\sum_{n=1}^{\infty} \mu(A_n) = \sum_{m=1}^{\infty} \mu(B_m)$. Ugyanis rögzített n mellett $A_n = \cup_{m=1}^{\infty} (A_n \cap B_m)$ diszjunkt egyesítés, ezért minden egyes $n \in \mathbb{N}$ természetes szám mellett $\mu(A_n) = \sum_{m=1}^{\infty} \mu(A_n \cap B_m)$. Hasonlóan minden $m \in \mathbb{N}$-re $B_m = \cup_{n=1}^{\infty} (B_m \cap A_n)$ így $\mu(B_m) = \sum_{n=1}^{\infty} \mu(B_m \cap A_n)$. Összefoglalva

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu(A_n \cap B_m) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mu(B_m \cap A_n) = \sum_{m=1}^{\infty} \mu(B_m),$$

hiszen nem negatív tagú dupla indexű sor összegzési sorrendje felcserélhető.

Tekintsük most a generált $\mathrm{r}(\mathcal{P})$ gyűrű egy tetszőleges elemét, azaz legyen $E = \cup_{n=1}^{N} A_n$ diszjunkt egyesítés, ahol $A_n \in \mathcal{P}$, amint azt a generált gyűrű belső reprezentációját megadó 1.1.12. állításban meggondoltuk. Ahhoz, hogy az $\mathrm{r}(\mathcal{P})$-re kiterjesztett halmazfüggvényt

$$\hat{\mu}(E) = \cup_{n=1}^{N} \mu(A_n)$$

módon definiálhassuk, először is azt kell meggondolnunk, hogy amennyiben E előáll esetleg más módon $E = \cup_{m=1}^{M} B_m$ alakban, akkor $\hat{\mu}$ értéke az előállítástól független, azaz $\sum_{n=1}^{N} \mu(A_n) = \sum_{m=1}^{M} \mu(B_m)$. No de az előzőek szerint ez nyilvánvaló, hiszen válasszuk meg az N, illetve M indexeknél nagyobb indexű halmazokat az üres halmaznak és egyszerűen alkalmazhatjuk az imént már meggondolt $\sum_{n=1}^{\infty} \mu(A_n) = \sum_{m=1}^{\infty} \mu(B_m)$ egyenlőséget.

A $\hat{\mu} : \mathrm{r}(\mathcal{P}) \to \overline{\mathbb{R}}_+$ nem negatív halmazfüggvény tehát a fent kiemelt módon jól definiált. Nyilvánvaló, hogy $\hat{\mu}$ kiterjesztése μ-nek, sőt $\hat{\mu}$ még a végesen additív kiterjesztések közt is az egyetlen.

Most azt lássuk be, hogy $\hat{\mu}$ egy σ-additív halmazfüggvény, tehát amennyiben egy $A \in \mathrm{r}(\mathcal{P})$ halmaz előáll $A = \cup_{n=1}^{\infty} A_n$ diszjunkt egyesítés alakban, ahol $A_n \in \mathrm{r}(\mathcal{P})$, akkor $\hat{\mu}(A) = \sum_{n=1}^{\infty} \hat{\mu}(A_n)$. Legyenek tehát az A_n halmazok $A_n = \cup_{i=1}^{r_n} C_i^n$ diszjunkt egyesítés alakúak, ahol minden $C_i^n \in \mathcal{P}$, valamint $A = \cup_{k=1}^{N} B_k$ szintén diszjunkt halmazok egyesítése, ahol $B_k \in \mathcal{P}$. Ekkor $\cup$ $\cup_{k=1}^{N} B_k = \cup_{n=1}^{\infty} \cup_{i=1}^{r_n} C_i^n$ egyenlőségben mind a bal, mind a jobb oldalon egymástól diszjunkt, $\mathcal{P}$-beli halmazok legfeljebb megszámlálható egyesítése látható. Újra alkalmazva a bizonyítás első gondolatát azt kapjuk, hogy

$$\hat{\mu}(A) = \sum_{k=1}^{N} \mu(B_k) = \sum_{n=1}^{\infty} \sum_{i=1}^{r_n} \mu(C_i^n) = \sum_{n=1}^{\infty} \hat{\mu}(A_n). \qquad \square$$

1.2.10.
A fenti bizonyítás csekély módosításával az is látható, hogy amennyiben μ egy a $\mathcal{P}$ félgyűrűn értelmezett végesen additív nem negatív halmazfüggvény, akkor ilyen módon egyértelműen kiterjeszthető az $\mathrm{r}(\mathcal{P})$ generált gyűrűre a végesen additivitás megtartásával.

Most egy nagyon fontos σ-additív halmazfüggvényt definiálunk. Az értelmezési tartomány a valós egyenes korlátos, balról zárt, jobbról nyílt intervallumainak félgyűrűje. Ezt a megszámlálhatóan additív halmazfüggvényt fogjuk később kiterjeszteni a hatványhalmaz lehetőség szerint minél nagyobb részhalmazára. Ezt nevezik majd *Lebesgue–Stieltjes-mértéknek* vagy az identitásfüggvény speciális esetében *Lebesgue-mértéknek*.

1.2.11. definíció (λ_f)**.** Legyen $f : \mathbb{R} \to \mathbb{R}$ egy monoton növő, balról folytonos függvény. Egy $[a, b)$ intervallumon definiálja

$$\lambda_f([a, b)) = f(b) - f(a).$$

és $\lambda_f(\varnothing) = 0$.

Persze az $f = \mathrm{id}$ függvény speciális esetében λ_{id} a közönséges hosszúság függvény, ami a legtermészetesebb mértékhez vezet. Első látásra didaktikusabb az alábbiakat először csak az identitásfüggvényre végiggondolni, viszont a teljes általánosság esete technikailag alig különbözik az identitásfüggvény esetétől. Az alábbi regularitások igazolása az egyetlen komplikáció, ami a balról folytonosság szerepét emeli ki.

1.2.12. állítás (λ_f regularitása)**.** *Az összes korlátos, balról zárt, jobbról nyílt intervallumok $\mathcal{P}$ félgyűrűjén értelmezett $\lambda_f : \mathcal{P} \to \mathbb{R}_+$ halmazfüggvényre tel-*

jesülnek az alábbi regularitások:

$$\lambda_f(I) = \inf\{f(b) - f(a) : I \subseteq (a,b)\},$$
$$\lambda_f(I) = \sup\{f(b) - f(a) : [a,b] \subseteq I\}$$

minden $I \in \mathcal{P}$ mellett.

Bizonyítás. Az f monoton növekedése szerint λ_f valóban egy nem negatív halmazfüggvény. Legyen $I = [\alpha, \beta)$ és $I \subseteq (a,b)$. Ekkor $a < \alpha < \beta \leq b$, ezért az f monoton növekedése szerint $\lambda_f(I) = f(\beta) - f(\alpha) \leq f(b) - f(a)$.

Az f függvény α pontbeli balról folytonossága szerint, $\varepsilon > 0$-hoz létezik $a < \alpha$, melyre $f(\alpha) - f(a) < \varepsilon$. Ebből

$$f(\beta) - f(a) < f(\beta) - f(\alpha) + \varepsilon = \lambda_f(I) + \varepsilon,$$

amivel igazoltuk az infimumra vonatkozó regularitást, hiszen $I = [\alpha,\beta) \subseteq \subseteq (a,\beta)$.

Hasonlóan, ha $[a,b] \subseteq I = [\alpha,\beta)$, akkor $\alpha \leq a < b < \beta$, ezért f növekedése miatt $\lambda_f(I) = f(\beta) - f(\alpha) \geq f(b) - f(a)$.

Az f függvény β pontbeli balról folytonossága szerint minden $\varepsilon > 0$-hoz létezik $b < \beta$, amelyre $f(\beta) - f(b) < \varepsilon$. Így $[\alpha, b] \subseteq [\alpha,\beta) = I$ és

$$\lambda_f(I) - \varepsilon = f(\beta) - f(\alpha) - \varepsilon < f(b) - f(\alpha),$$

ami a szuprémumra vonatkozó egyenlőséget igazolja. □

A fenti λ_f függvény végesen additivitása egészen nyilvánvaló. Az alábbiakban azt mutatjuk meg, hogy egy végesen additív halmazfüggvény, amely kielégíti a fenti regularitásokat (1.2.12) már σ-additív is.

1.2.13. állítás (λ_f σ-additív)**.** *Jelölje $\mathcal{P}$ a valós egyenes korlátos, balról zárt, jobbról nyílt intervallumainak félgyűrűjét. Legyen $f : \mathbb{R} \to \mathbb{R}$ monoton növekedő, balról folytonos függvény. Definiálja $\lambda_f : \mathcal{P} \to \mathbb{R}_+$ az f megváltozását, azaz $[a,b) \in \mathcal{P}$ esetén*

$$\lambda_f([a,b)) = f(b) - f(a).$$

Az így definiált $\lambda_f : \mathcal{P} \to \mathbb{R}_+$ egy σ-additív halmazfüggvény a $\mathcal{P}$ félgyűrűn.

Bizonyítás. Ha $[a,b) = [a,u) \cup [u,b)$, akkor $f(b) - f(a) = f(b) - f(u) + + f(u) - f(a)$, ami igazolja λ_f végesen additivitását.

Jelölje $\hat{\lambda}_f : \mathrm{r}(\mathcal{P}) \to \overline{\mathbb{R}}_+$ a generált gyűrűre való végesen additív kiterjesztést (1.2.10).

Tegyük fel, hogy az $I \in \mathcal{P}$ balról zárt, jobbról nyílt intervallum előáll mint ugyanilyen $I_n \in \mathcal{P}$ intervallumok diszjunkt egyesítése, tehát $I = \cup_{n=1}^{\infty} I_n$. Először a σ-szuperadditivitást, majd σ-szubadditivitást igazoljuk.

1. Rögzített N természetes számra legyen $E_N = \cup_{n=1}^N I_n$. Ekkor $E_N, I \in \mathrm{r}(\mathcal{P})$ és $I = (I \setminus E_N) \cup E_N$ egy az $\mathrm{r}(\mathcal{P})$ gyűrűbeli diszjunkt előállítás. Így $\hat{\lambda}_f$ véges additivitása és definíciója szerint

$$\lambda_f(I) = \hat{\lambda}_f(I) = \hat{\lambda}_f(I \setminus E_n) + \hat{\lambda}_f(E_N) \geq \hat{\lambda}_f(E_N) = \sum_{k=1}^N \lambda_f(I_k).$$

Ez utóbbi egyenlőtlenség persze minden rögzített $N \in \mathbb{N}$ mellett is fennáll, ezért $\lambda_f(I) \geq \sum_{n=1}^\infty \lambda_f(I_n)$.

2. A σ-szubadditivitás igazolásához, legyen ε pozitív szám rögzítve. A λ_f halmazfüggvény regularitási tulajdonságait szem előtt tartva (1.2.12), válasszuk meg az alábbi intervallumokat:

$$[a,b] \subseteq I, \qquad \lambda_f(I) - \frac{\varepsilon}{2} < f(b) - f(a);$$
$$I_n \subseteq (a_n, b_n), \qquad \lambda_f(I_n) + \frac{\varepsilon}{2^{n+1}} > f(b_n) - f(a_n).$$

Világos, hogy ekkor $[a,b] \subseteq \cup_{n=1}^\infty (a_n, b_n)$, ezért a kompaktság definíciója, azaz a Heine–Borel-tétel miatt az (a_n, b_n) intervallumok közül véges sok is lefedi az $[a,b]$ korlátos és zárt intervallumot, ergo van olyan N természetes szám, amelyre

$$[a,b) \subseteq [a,b] \subseteq \cup_{k=1}^N (a_k, b_k) \subseteq \cup_{k=1}^N [a_k, b_k).$$

A $\hat{\lambda}_f$ monotonitása, szubadditivitása (1.2.7), és az $a, b, a_k, b_k \in \mathbb{R}$ számok megválasztása szerint

$$\begin{aligned}\lambda_f(I) - \frac{\varepsilon}{2} &< f(b) - f(a) = \\ &= \lambda_f([a,b)) = \hat{\lambda}_f([a,b)) \leq \hat{\lambda}_f\left(\cup_{k=1}^N [a_k, b_k)\right) \leq \\ &\leq \sum_{k=1}^N \hat{\lambda}_f([a_k, b_k)) = \sum_{k=1}^N \lambda_f([a_k, b_k)) = \\ &= \sum_{k=1}^N (f(b_k) - f(a_k)) < \sum_{k=1}^N \left(\lambda_f(I_k) + \frac{\varepsilon}{2^{k+1}}\right) \leq \sum_{k=1}^\infty \lambda_f(I_k) + \frac{\varepsilon}{2}.\end{aligned}$$

Mivel ez minden pozitív ε mellett fennáll, ezért a $\lambda_f(I) \leq \sum_{n=1}^\infty \lambda_f(I_n)$ egyenlőtlenség is igaz. $\square$

A λ_f definíciójában tehát az f balról folytonossága a regularitásokon keresztül (1.2.12) a σ-szubadditivitáshoz szükséges. Ellenpéldaként nézzük az

$$f(x) = \begin{cases} 1, & x \geq 1, \\ 0, & x < 1 \end{cases}$$

monoton növő, de az 1 pontban nem balról folytonos függvényt. Válasszunk tetszőleges olyan sorozatot, melyre $t_1 = 0$, és $t_n \to 1$ szigorúan monoton növekedően. Ekkor $[0,1) = \cup_{n=1}^{\infty} [t_n, t_{n+1})$. Mivel minden $n \in \mathbb{N}$ mellett $t_n < < 1$, ezért $\sum_{n=1}^{\infty} \lambda_f ([t_n, t_{n+1})) = 0$, de $\lambda_f ([0,1)) = 1$, így λ_f végesen additív de nem σ-additív halmazfüggvény. Ha f jobbról folytonosságát tesszük fel, akkor a balról nyílt jobbról zárt véges intervallumokon kell λ_f-et definiálnunk és ugyanígy megkapjuk a σ-szubadditivitást.

Feltűnő, hogy még a közönséges hosszúság függvény ($f = \mathrm{id}$ speciális eset) σ-additivitásának indoklásához is a valós egyenes viszonylag finom topológiai tulajdonságát használtuk (Heine–Borel-tétel).

1.3. Mérhető függvények

Látni fogjuk, hogy annál hatékonyabban tudunk bánni egy σ-algebrával minél több – lehetőség szerint minél szűkebb – generátor rendszerét tudjuk megadni.

Mivel egy σ-algebra elemeit mérhető halmazoknak tituláljuk, így a mérhetőség alábbi definíciója egyszerűen úgy fogalmazható, hogy minden mérhető halmaz ősképének is mérhetőnek kell maradnia.

1.3.1. definíció (mérhető függvény)**.** Legyen $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető tér, valamint $f : X \to Y$ függvény. Azt mondjuk, hogy f egy *mérhető* függvény e két mérhető tér közt, ha minden $E \in \mathcal{N}$ esetén $f^{-1}(E) \in \mathcal{M}$.

Legyen $A \in \mathcal{M}$ egy mérhető részhalmaz. Azt mondjuk, hogy az $f : X \to Y$ mérhető az A halmazon, ha $f^{-1}(E) \cap A \in \mathcal{M}$ teljesül minden $E \in \mathcal{N}$ halmaz esetén.

Legyen most $f : X_0 \to Y$ függvény értelmezve az $X_0 \in \mathcal{M}$ halmazon. Azt mondjuk, hogy f mérhető, ha f mérhető az X_0 értelmezési tartománya felett.

Az $X_0 \cap f^{-1}(E) = f^{-1}(Y) \cap f^{-1}(E) = f^{-1}(Y \cap E) = f^{-1}(E)$ azonosság azt mutatja, hogy fenti definíció harmadik esetében is elég arra emlékeznünk, hogy Y-beli mérhető halmaz ősképének is X-ben mérhetőnek kell lennie.

1.3.2.
Látható, hogy ha $A \cup B = X$, $A, B \in \mathcal{M}$ mérhető halmazok, akkor egy $f : X \to Y$ függvény mérhetősége ekvivalens azzal, hogy f mérhető az A és a B halmazok felett. Ha tehát mondjuk, $g_1 : A \to Y$ és $g_2 : A^c \to Y$ mérhető függvények, akkor az ezekből összerakott

$$g(x) = \begin{cases} g_1(x), & \text{ha } x \in A, \\ g_2(x), & \text{ha } x \notin A \end{cases}$$

függvény is mérhető. Ez utóbbira úgy hivatkozunk, hogy ha egy mérhető függvényt az értelmezési tartománya egy mérhető halmazán egy másik mérhető függvénnyel cserélünk ki, akkor a kapott függvény is mérhető marad.

Az első valódi észrevételünk az, hogy a mérhetőség definícióját elegendő az értékeket tartalmazó halmaz σ-algebrájának egy generátorára ellenőrizni.

1.3.3. állítás. *Legyen $f : X \to Y$ függvény az $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető terek közt. Tegyük fel, hogy $\mathcal{H} \subseteq \mathcal{P}(Y)$ olyan halmazrendszer, amely generálja $\mathcal{N}$-et, azaz $\sigma(\mathcal{H}) = \mathcal{N}$. Az f függvény pontosan akkor mérhető, ha minden $E \in \mathcal{H}$ generátorhalmaz esetén $f^{-1}(E) \in \mathcal{M}$.*

Bizonyítás. Definiáljuk az Y tér alábbi halmazrendszerét:

$$\mathcal{F} = \left\{E \subseteq Y : f^{-1}(E) \in \mathcal{M}\right\}. \tag{1.6}$$

Látható, hogy $\mathcal{F}$ egy σ-algebra, hiszen
– $\varnothing \in \mathcal{F}$, hiszen $f^{-1}(\varnothing) = \varnothing \in \mathcal{M}$.
– Ha $E \in \mathcal{F}$, akkor $f^{-1}(E^c) = \left(f^{-1}(E)\right)^c \in \mathcal{M}$, így $E^c \in \mathcal{F}$.
– Ha $E_n \in \mathcal{F}$, akkor $f^{-1}(\cup_{n=1}^{\infty} E_n) = \cup_{n=1}^{\infty} f^{-1}(E_n) \in \mathcal{M}$, azaz $\cup_{n=1}^{\infty} E_n \in \mathcal{F}$.
Ha tehát a generátorrendszer minden $H \in \mathcal{H}$ elemére $f^{-1}(H) \in \mathcal{M}$ teljesül, az azt jelenti, hogy $\mathcal{H} \subseteq \mathcal{F}$, amiből

$$\mathcal{N} = \sigma(\mathcal{H}) \subseteq \sigma(\mathcal{F}) = \mathcal{F}.$$

következik. No de az $\mathcal{N} \subseteq \mathcal{F}$ tartalmazás éppen a mérhetőséget jelenti. □

A mérhetőség ekvivalens az $\mathcal{N} \subseteq \mathcal{F}$ tartalmazással. Ezért azonnal látszik, hogy az Y halmaz felett a legbővebb olyan $\mathcal{N}$ σ-algebra, amelyre nézve az f függvény $\mathcal{M}, \mathcal{N}$ mérhető, éppen $\mathcal{N} = \mathcal{F}$, ahol $\mathcal{F}$ az (1.6) szerint értelmezett.

1.3.4. (Borel-mérhető függvény)
Speciálisan, ha az (Y, τ) egy topologikus tér, és $\mathcal{B} = \sigma(\tau)$ az Y Borel-halmazai, akkor egy $f : X \to Y$ függvény *Borel-mérhetőségén* az $(X, \mathcal{M})$ és $(Y, \mathcal{B})$ mérhető terek közti mérhetőséget értjük. Így ha $\mathcal{H} \subseteq \mathcal{B}$ akármelyik generátora a Borel-halmazoknak, akkor f Borel-mérhetősége

$$f^{-1}(H) \in \mathcal{M}, \forall H \in \mathcal{H}$$

feltétel teljesülését jelenti.

1.3.5.
Az egyik legfontosabb Borel-mérhető függvényosztály a folytonos függvények összessége. Ehhez fel kell tennünk, hogy az X téren és az Y téren is adott egy-egy τ_X és τ_Y topológia, és legyen az X-en a σ-algebra az X-beli topológia által generált Borel-halmazok halmaza $\mathcal{B}_X = \sigma(\tau_X)$, hasonlóan az Y-on is az ottani nyílt halmazok generálta legszűkebb σ-algebra, tehát $\mathcal{B}_Y = \sigma(\tau_Y)$. Ilyen módon tekintsünk a mérhető terek közti $f : (X, \mathcal{B}_X) \to (Y, \mathcal{B}_Y)$ függvényt. E függvény pontosan akkor *folytonos*, ha minden $V \in \tau_Y$ nyílt halmaz $f^{-1}(V) \in \tau_X$, azaz minden nyílt halmaz ősképe is nyílt. Ekkor persze az

is igaz, hogy minden nyílt halmaz ősképe is Borel halmaz, hiszen a Borel-halmazok éppen a nyílt halmazok generálta σ-algebra. Meggondoltuk tehát, hogy minden *folytonos függvény* Borel-mérhető.

Két nevezetes generátora a Borel-halmazoknak mindenképpen van: A nyílt halmazok halmaza, és a zárt halmazok halmaza. Látjuk tehát, hogy mind a *minden nyílt halmaz ősképe mérhető* feltétel, mind a *minden zárt halmaz ősképe mérhető* feltétel ugyanazt jelenti, amit *Borel-mérhetőségnek* neveztünk. Ha speciális topologikus tereket nézünk, akkor a Borel-halmazoknak szűkebb és szűkebb generátorait érdemes keresnünk, hogy a Borel-mérhetőségre is egyszerűbben és egyszerűbben ellenőrizhető feltételeket kapjunk.

1.3.6.
Lássuk el az $\overline{\mathbb{R}}$ számegyenes pontjait a szokásos euklideszi topológiával. Ha $(a, b) \subseteq \overline{\mathbb{R}}$ egy nyílt intervallum, akkor válasszunk $a_n \in \mathbb{Q}$ racionális számokat, melyekre $a_n \in (a, b)$ és $a_n \to a$. Így

$$(a, b) = \cup_{n=1}^{\infty} (a_n, b) .$$

Hasonló konstrukcióval látható, hogy minden nyílt intervallum előáll mint megszámlálhatóan sok, de racionális végpontú nyílt intervallum egyesítése. Mivel minden nyílt halmaz is előáll mint megszámlálhatóan sok nyílt intervallum uniója, ezért mondhatjuk, hogy minden nyílt halmaz tekinthető mint legfeljebb megszámlálhatóan sok racionális végpontú nyílt intervallum egyesítése. Ha tehát

$$\mathcal{H}_1 = \{(a, b) : a, b \in \mathbb{Q}, a < b\},$$

jelöli a racionális végpontú korlátos nyílt intervallumok halmazát, akkor minden nyílt halmaz eleme minden olyan σ-algebrának, amely $\mathcal{H}_1$-t tartalmazza, speciálisan

$$\tau \subseteq \sigma (\mathcal{H}_1) . \tag{1.7}$$

Jelölje most

$$\mathcal{H}_2 = \{[a, b) : a < b, a, b \in \mathbb{Q}\} .$$

Látható, hogy

$$(a, b) = \cup_{n=1}^{\infty} [a_n, b)$$

fennáll tetszőlegesen választott olyan $a_n \in (a, b) \cap \mathbb{Q}$ sorozatra, melyre $a_n \to a$. Ebből azonnal következik, hogy

$$\mathcal{H}_1 \subseteq \sigma (\mathcal{H}_2) \tag{1.8}$$

Most definiálja

$$\mathcal{H}_3 = \{(-\infty, a) : a \in \mathbb{Q}\} .$$

Persze minden $[a, b) \in \mathcal{H}_2$ esetén

$$[a, b) = (-\infty, b) \cap [a, \infty) = (-\infty, b) \cap (-\infty, a)^c ,$$

így $\mathcal{H}_2$ részhalmaza minden $\mathcal{H}_3$-at tartalmazó gyűrűnek, pláne

$$\mathcal{H}_2 \subseteq \sigma(\mathcal{H}_3). \tag{1.9}$$

Így a fenti kiemelt (1.7), (1.8), (1.9) tartalmazásokra alkalmazva a σ-operátort azt kapjuk, hogy

$$\mathcal{B} = \sigma(\tau) \subseteq \sigma(\sigma(\mathcal{H}_1)) = \sigma(\mathcal{H}_1) \subseteq$$
$$\subseteq \sigma(\sigma(\mathcal{H}_2)) = \sigma(\mathcal{H}_2) \subseteq \sigma(\sigma(\mathcal{H}_3)) = \sigma(\mathcal{H}_3) \subseteq \mathcal{B}.$$

Azt gondoltuk meg tehát, hogy $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ külön-külön generátorai az $\overline{\mathbb{R}}$ kiterjesztett számegyenes Borel-halmazai alkotta σ-algebrának.

Az eddigihez hasonló okoskodással látható, az alábbi halmazrendszerek is külön-külön generátor rendszerei a számegyenes Borel-halmazainak:

$$\{(a,b] : a,b \in \mathbb{Q}, a < b\},\ \{(-\infty, a] : a \in \mathbb{Q}\},$$
$$\{[a,+\infty) : a \in \mathbb{Q}\},\ \{(a,+\infty) : a \in \mathbb{Q}\}.$$

A Borel-halmazok fent megadott generátorrendszerei segítségével fogalmazzunk meg ekvivalens feltételeket egy valós függvény Borel-mérhetőségére! Pl. $\mathcal{H}_2$ halmazrendszert tekintve generátornak: egy $f : X \to \overline{\mathbb{R}}$ függvény pontosan akkor Borel mérető, ha minden $a, b \in \mathbb{Q}, a < b$ mellett $X(a \le f < b) = f^{-1}([a,b))$ halmaz mérhető. Az alábbi feltétel ezek közül olyan fontos és olyan sokszor használatos, hogy külön is kiemeljük:

1.3.7. állítás. *Legyen $(X, \mathcal{M})$ egy mérhető tér, $f : X \to \overline{\mathbb{R}}$ egy függvény. Az f pontosan akkor mérhető, ha minden $\alpha \in \mathbb{Q}$ esetén az $X(f > \alpha)$ nívóhalmazok mérhetők.*

Bizonyítás. Láttuk, hogy az $\{(-\infty, \alpha) : \alpha \in \mathbb{R}\}$ halmazrendszer a számegyenes Borel-halmazainak generátora. Mivel a mérhetőséget elegendő a generátorrendszer halmazaira ellenőrizni (1.3.3), ezért f mérhetősége ekvivalens az $X(f > \alpha) = f^{-1}((\alpha, \infty]) \in \mathcal{M}, \forall \alpha \in \mathbb{R}$ feltétellel. □

1.3.8.
Ha tekintünk egy $f : X \to \overline{\mathbb{R}}$ mérhető függvényt, akkor

$$X(f = \infty) = \cap_{n=1}^{\infty} X(f > n)$$

előállítás szerint $X(f = \infty) \in \mathcal{M}$. Hasonlóan látszik, hogy az $X(f = -\infty)$ halmaz is mérhető. Az X értelmezési tartomány tehát három diszjunkt mérhető halmaz egyesítése: $X = X(-\infty < f < +\infty) \cup X(f = +\infty) \cup X(f = -\infty)$.

Most meggondoljuk, hogy mérhető függvények kompozíciója is mérhető marad.

1.3.9. állítás. *Tegyük fel, hogy adott három mérhető tér,* $(X, \mathcal{M})$, $(Y, \mathcal{N})$, $(Z, \mathcal{O})$. *Legyen* $f_1 : X \to Y$ *mérhető és* $f_2 : Y \to Z$ *mérhető függvény. Ekkor ezek* $f_2 \circ f_1$ *kompozíciója is mérhető.*

Bizonyítás. Világos, hogy minden $H \in \mathcal{O}$ mérhető halmazra

$$(f_2 \circ f_1)^{-1}(H) = f_1^{-1}\left(f_2^{-1}(H)\right) \in \mathcal{M},$$

hiszen $f_2^{-1}(H) \in \mathcal{N}$ az f_2 mérhetősége szerint és $f_1^{-1}(G) \in \mathcal{M}$ az f_1 mérhetősége szerint a $G = f_2^{-1}(H)$ halmazra alkalmazva. □

Láttuk korábban, hogy minden folytonos függvény Borel-mérhető. A most bevezetendő függvényosztály még ennél is fontosabb. Látni fogjuk, hogy a karakterisztikus függvényekből mint építőkockákból végül minden mérhető függvény előállítható.

1.3.10. definíció (karakterisztikus függvény)**.** Az $E \subseteq X$ halmaz mellett, legyen $\chi_E : X \to \mathbb{R}$

$$\chi_E(x) = \begin{cases} 1, \text{ha } x \in E; \\ 0, \text{ha } x \notin E \end{cases}$$

az E halmaz *karakterisztikus függvénye.* Szokás a karakterisztikus függvényt – főleg valószínűség-számításban – *indikátor függvénynek* is nevezni.

1.3.11.
Nézzük meg χ_E nívóhalmazait. Az α értékétől függően

$$X(\chi_E > \alpha) = \begin{cases} X, & \text{ha } \alpha < 0, \\ E, & \text{ha } 0 \le \alpha < 1, \\ \varnothing, & \text{ha } 1 \le \alpha. \end{cases}$$

Mivel $\varnothing, X \in \mathcal{M}$ minden σ-algebra esetén, ezért χ_E karakterisztikus függvény mérhetősége egybeesik az E halmaz mérhetőségével.

Nézzük, hogyan hatnak a függvények mérhetőségére az algebrai operációk.

1.3.12. állítás (mérhető függvények összege mérhető)**.** *Legyenek* $u, v : X \to \overline{\mathbb{R}}$ *Borel-mérhető függvények az* $(X, \mathcal{M})$ *mérhető téren úgy, hogy* $u + v$ *értelmes. Ekkor* $u + v$ *is Borel-mérhető.*

Bizonyítás. Először is jegyezzük meg, hogy minden $V \subseteq \mathbb{R}^2$ nyílt halmaz előáll mint megszámlálhatóan sok $I \times J$ alakú nyílt halmaz egyesítése, ahol I és J nyílt, $\mathbb{R}$-beli intervallumok. Legyen $X_u = X(-\infty < u < +\infty)$ és $X_v = X(-\infty < v < +\infty)$. Tudjuk, hogy $X_u, X_v \in \mathcal{M}$. Legyen $f : X_u \cap X_v \to \mathbb{R}^2$ függvényre $f(x) = (u(x), v(x))$. Tehát ha $I \times J$ egy fenti típusú halmaz, akkor

$$f^{-1}(I \times J) = \{x \in X : u(x) \in I \text{ és } v(x) \in J\} = u^{-1}(I) \cap v^{-1}(J).$$

Az u és a v függvények Borel-mérhetősége miatt $f^{-1}(I \times J) \in \mathcal{M}$. Ha a V nyílt halmaz $V = \cup_{n=1}^{\infty}(I_n \times J_n)$ alakú, akkor $f^{-1}(V) = \cup_{n=1}^{\infty} f^{-1}(I_n \times J_n) \in \mathcal{M}$, hiszen $\mathcal{M}$ zárt a megszámlálható egyesítésre, tehát f mérhető.

Definiáljuk az $u + v : X \to \overline{\mathbb{R}}$ összeg függvényt:

$$(u+v)(x) = \begin{cases} (+ \circ f)(x), & \text{ha } x \in X_u \cap X_v, \\ +\infty, & \text{ha } u(x) = +\infty \text{ vagy } v(x) = +\infty, \\ -\infty, & \text{ha } u(x) = -\infty \text{ vagy } v(x) = -\infty. \end{cases}$$

Mivel $u + v$ értelmes X-en, a fenti három ág közül pontosan az egyik teljesül minden $x \in X$-re. Az $+ : \mathbb{R}^2 \to \mathbb{R}$ függvény folytonos, ezért mérhető. Az f mérhetőségét az imént láttuk, így a fenti függvény egy mérhető részhalmazon értelmezett mérhető függvény, és annak komplementere egyik mérhető részén konstans $+\infty$, másik mérhető részén konstans $-\infty$. Így $u + v$ is mérhető. □

A fenti bizonyítás előnye, hogy könnyedén átvihető szorzatra és hányadosra is. Nézzünk egy másik bizonyítást is.

Bizonyítás. Ha u egy mérhető függvény, akkor minden $\alpha \in \mathbb{R}$ mellett $u + \alpha$ is mérhető, hiszen egy $\overline{\mathbb{R}}$-beli nyílt halmaz eltoltja is nyílt, és $(u+\alpha)^{-1}(V) = u^{-1}(V - \alpha)$. Hasonlóan, mivel egy valós nyílt halmaz -1-szerese is nyílt, ezért $-u$ is mérhető az $(-u)^{-1}(V) = u^{-1}(-V)$ azonosság szerint. Most vegyük észre, hogy az u, v mérhető függvényekre az archimédeszi axióma szerint

$$X(u > v) = \cup_{\alpha \in \mathbb{Q}} \left(X(v < \alpha) \cap X(\alpha < u) \right).$$

Előállítottuk tehát az $X(u > v)$ halmazt mérhető halmazok megszámlálható metszetével és egyesítésével, ezért $X(u > v) \in \mathcal{M}$. Alkalmazzuk ezt v helyett az $\alpha - v$ mérhető függvényre. Azt kapjuk tehát, hogy

$$X(u > \alpha - v) = X(u + v > \alpha) \in \mathcal{M}.$$

fennáll minden $\alpha \in \mathbb{R}$ mellett. Ezt kellett belátni. □

Ha ebből a bizonyításból indulunk ki, fg mérhetőségéhez először f^2 mérhetőségét kell meggondolnunk az $X\left(f^2 < r\right) = X\left(f \in (-\sqrt{r}, \sqrt{r})\right)$ azonosság segítségével, majd az $fg = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right)$ azonossággal jutunk fg mérhetőségéhez.

1.3.13.

Az összeadás mérhetőségére vonatkozó állításban feltételként szerepelt, hogy $u(x) + v(x)$ értelmes minden $x \in X$ mellett. Erre csak azért volt szükség, hogy a fenti bizonyítások utolsó kiemelt képlete helyes legyen. E feltétel nélkül is igaz az állítás abban az értelemben, hogy $u + v$ értelmezési tartománya nem a teljes X halmaz. Ilyenkor azt lehet mondani, hogy $u + v$ függvényt értelmezzük azon a mérhető halmazon ahol az összeg értelmes, e mérhető halmaz komplementerén pedig tetszőleges konstanst beállítva a kiterjesztett összeg mérhető függvény marad.

Most azt nézzük meg, hogy a mérhetőség hogyan öröklődik nem algebrai hanem analízisbeli operációk hatására.

1.3.14. állítás (mérhető függvények szuprémuma, határértéke mérhető)**.** *Tegyük fel, hogy egy mérhető téren adva van megszámlálhatóan sok valós értékű mérhető függvény. Ezek pontonkénti szuprémuma, infimuma, limsup-ja, liminf-je és esetleges határértéke is mérhető függvény.*

Bizonyítás. Legyen $(X, \mathcal{M})$ a mérhető tér és $f_n : X \to \overline{\mathbb{R}}$ mérhető függvények. Nézzük először a szuprémumra vonatkozó állítást. Legyen minden $x \in X$ mellett $f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ a pontonkénti szuprémum. Vegyük észre, hogy a szuprémum definíciója szerint minden $\alpha \in \mathbb{R}$ valós szám mellett

$$X(f > \alpha) = \cup_{n=1}^{\infty} X(f_n > \alpha).$$

Na most f_n mérhetősége szerint minden $n \in \mathbb{N}$ mellett $X(f_n > \alpha) \in \mathcal{M}$, így a σ-algebra megszámlálható egyesítésre való zártsága miatt ezen halmazok egyesítése is mérhető.

Az infumum-ra vonatkozó állítás ennek és a -1 szeres mérhetőségének közvetlen folyománya, hiszen

$$\inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}} (-f_n).$$

Ennek megfelelően a limsup és liminf függvények mérhetősége a definícióból

$$\limsup_{n\in\mathbb{N}} f_n = \inf_{n\in\mathbb{N}} \left(\sup_{k\geq n} f_k\right), \quad \liminf_{n\in\mathbb{N}} f_n = \sup_{n\in\mathbb{N}} \left(\inf_{k\geq n} f_k\right)$$

következik. Ha a függvénysorozat pontonként konvergens, akkor a határérték megegyezik például a limsup-pal. Ezt kellett belátni. □

A fenti állításban fontos, hogy csak megszámlálhatóan sok függvényről van szó. Az állítás megszámlálhatónál nagyobb számosságú függvényosztály esetén csak nagyon speciális σ-algebra mellett maradhat igaz. Például, ha minden egyelemű halmaz mérhető, és van $E \notin \mathcal{M}$ nem mérhető halmaz, akkor a

$$\chi_E = \sup\left\{\chi_{\{x\}} : x \in E\right\}$$

egyenlőség szerint, χ_E nem mérhető függvény, ámbár mérhető függvények szuprémuma.

1.3.15. (mérhető függvény pozitív és negatív része is mérhető)
Emlékezzünk arra, hogy $a \in \mathbb{R}$ valós szám esetén $a^+ = \max\{a,0\}$ és $a^- = \max\{-a,0\}$. Világos, hogy $a^+, a^- \geq 0$, $a = a^+ - a^-$ és $|a| = a^+ + a^-$. Az a^+ és a^- az a valós szám *pozitív része*, illetve *negatív része*. Ezzel analóg módon egy $f : X \to \overline{\mathbb{R}}$ függvény pozitív és negatív része $f^+ = \max\{f,0\}$ és $f^- = \max\{-f,0\}$ módon definiált. Az előző állítás egyszerű következménye, hogy ha f mérhető függvény valamely $\mathcal{M}$ σ-algebrára nézve, akkor f^+ és f^- nem negatív függvények is $\mathcal{M}$-mérhetők.

1.3.16. definíció (lépcsős vagy egyszerű függvény)**.** Legyen $(X, \mathcal{M})$ egy mérhető tér. *Lépcsős függvénynek*, vagy *egyszerű függvénynek* nevezünk egy $s : X \to \mathbb{R}_+$ mérhető, véges értékkészletű, nem negatív függvényt.

Legyen az s egyszerű függvény értékkészlete $\mathcal{R}(f) = \{\alpha_1, \ldots, \alpha_n\}$. Ha bevezetjük minden $\alpha_i \in \mathcal{R}(f)$ mellett az $E_i = X(f = \alpha_i)$ jelölést, akkor az egyszerű függvény

$$s = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$$

alakban áll elő, ahol az $E_1, \ldots, E_n$ egymástól páronként diszjunkt, mérhető halmazok. Ezt az alakot nevezzük az egyszerű függvény *kanonikus alakjának*.

Fontos látni, hogy ugyanez az egyszerű függvény még nagyon sokféleképpen állítható elő a fenti $\sum_{i=1}^{m} \beta_i \chi_{F_i}$ alakban, ha nem kötjük meg azt, hogy a β_i nem negatív valós számok egymástól különbözzenek, vagy nem követeljük meg, hogy a β_i számokat az értékkészletből válasszuk, vagy ha nem írjuk elő az F_i halmazok diszjunktságát!

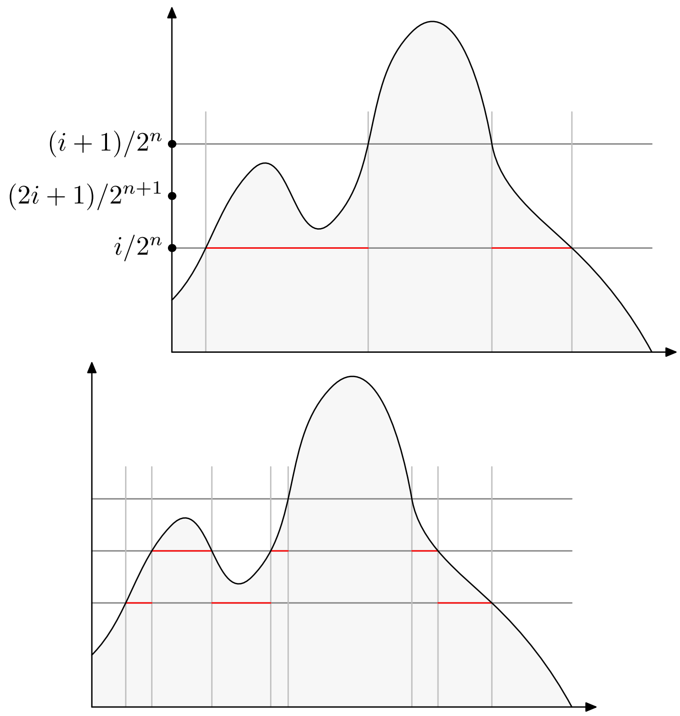
Például a $(0,1)$ felett 1 másutt zérus egyszerű függvény kanonikus alakja $0\chi_{(0,1)^c} + 1\chi_{(0,1)}$. Ugyanez a függvény még előáll például $\frac{1}{2}\chi_{(0,1)} + \frac{1}{2}\chi_{[1/2,1)} + \frac{1}{4}\chi_{(0,1/2)} + \frac{1}{4}\chi_{(0,1/2)}$ alakban is. Nagyon fontos tény, hogy az egyszerű függvény kanonikus alakjában való előállítása definíció szerint egyértelmű! Ilyenkor az $E_i = X(f_i = \alpha_i)$ halmazok az X alaphalmaz egy véges mérhető partícióját adják.

1.3.17. tétel (mérhető függvények alaptétele)**.** *Legyen $f : X \to \overline{\mathbb{R}}_+$ egy nem negatív valós értékű mérhető függvény az $(X, \mathcal{M})$ mérhető téren. Ekkor létezik s_n egyszerű függvények sorozata, melyre $0 \leq s_1 \leq s_2 \leq \ldots s_n \leq \ldots \leq f$ és $s_n \to f$ pontonként. Ha f korlátos, akkor még az is igaz, hogy s_n egyenletesen tart f-hez.*

A bizonyítás konstrukciójának egy szemléltetését látjuk az 1.1. ábrán.

Bizonyítás. Definiáljuk tetszőleges $n \in \mathbb{N}$ és $i = 0,1,2,\ldots,n2^n - 1$ mellett az alábbi halmazokat:

$$E_i^{(n)} = f^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right), \text{ továbbá } F^{(n)} = X(f \geq n).$$



1.1. ábra. Mérhető függvények alaptétele: Az s_n és s_{n+1} függvény az $E_i^{(n)}$ halmaz felett

Az f mérhetősége miatt $E_i^{(n)} \in \mathcal{M}$ és $F^{(n)} \in \mathcal{M}$ minden $n \in \mathbb{N}$ és $i = 0,1,2,\ldots,n2^n - 1$ esetén. Nyilvánvaló, hogy $X = \cup_{i=0}^{n2^n-1} E_i^{(n)} \cup F^{(n)}$ diszjunkt egyesítés fennáll minden $n \in \mathbb{N}$ mellett. Definiálja rögzített $n \in \mathbb{N}$ mellett

$$s_n = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} \chi_{E_i^{(n)}} + n\chi_{F^{(n)}}.$$

Világos, hogy s_n egyszerű függvény, hiszen véges az értékkészlete és mérhető függvények összegeként maga is mérhető. Tetszőleges $x \in X$ mellett, ha $x \in E_i^{(n)}$, akkor $s_n(x) = \frac{i}{2^n} \leq f(x)$, és ha $x \in F^{(n)}$, akkor $s_n(x) = n \leq f(x)$. Az is világos, hogy amennyiben $f(x)$ véges, úgy olyan n-hez amelyre $f(x) < n$ létezik i, hogy $x \in E_i^{(n)}$, ezért $f(x) - s_n(x) < \frac{1}{2^n}$. Ha $f(x) = +\infty$, akkor viszont $x \in F^{(n)}$ minden $n \in \mathbb{N}$ mellett, így $s_n(x) \to f(x)$ $\forall x \in X$ mellett pontonként. Ebből már látszik, hogy amennyiben f korlátos, úgy tetszőleges a korlátnál nagyobb n index mellett $f - s_n < \frac{1}{2^n}$ az egész X-en, tehát s_n valóban egyenletesen konvergál f-hez.

Már csak azt kell megmutatnunk, hogy a fenti módon definiált $(s_n)_{n\in\mathbb{N}}$ függvénysorozat pontonként monoton növő. Legyen tehát $x \in X$ rögzítve. Világos, hogy

$$E_i^{(n)} = E_{2i}^{(n+1)} \cup E_{2i+1}^{(n+1)},$$

hiszen $\frac{i}{2^n} = \frac{2i}{2^{n+1}}$ és $\frac{i+1}{2^n} = \frac{2i+2}{2^{n+1}}$. Ez azt jelenti, hogy amennyiben $x \in E_i^{(n)}$, úgy $s_n(x) = \frac{i}{2^n}$, de $s_{n+1}(x) = \frac{2i}{2^{n+1}}$ vagy $s_{n+1}(x) = \frac{2i+1}{2^{n+1}}$. Az első esetben $s_n(x) = s_{n+1}(x)$ míg a második esetben $s_n(x) < s_{n+1}(x)$. Ha $x \in F^{(n)}$, de $f(x) < n+1$, akkor $s_n(x) = n$ és $s_{n+1}(x) = \frac{i}{2^{n+1}}$ olyan i-re melyre $\frac{i}{2^{n+1}} \geq n$. Ha $x \in F^{(n)}$, és $f(x) \geq n+1$ is fennáll, akkor $s_n(x) = n$ és $s_{n+1}(x) = n+1$. Azt láttuk tehát, hogy tetszőleges $x \in X$ mellett $s_n(x) \leq s_{n+1}(x)$ valóban fennáll. Ezt kellett belátni. □

A mérhető alaptételt persze sorokra is megfogalmazhatjuk.

1.3.18. következmény. *Adott egy $(X, \mathcal{M})$ mérhető tér, és egy $f : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvény. Ekkor léteznek $\alpha_n \in \overline{\mathbb{R}}_+$ nem negatív valós számok és léteznek $E_n \in \mathcal{M}$ mérhető halmazok, amelyekkel az*

$$f = \sum_{n=1}^{\infty} \alpha_n \chi_{E_n}$$

pontonkénti sorösszeg-előállítás fennáll.

Bizonyítás. A mérhető függvények 1.3.17. alaptételét alkalmazzuk. Léteznek tehát $s_n \leq s_{n+1}$ egyszerű függvények, melyekre $s_n \to f$ az X halmazon pontonként. Az indexek esetleges elcsúsztatása után feltehetjük, hogy $s_1 = 0$. Jelölje

$$q_n = s_{n+1} - s_n.$$

Az s_n függvénysorozat monoton növekedése szerint $q_n \geq 0$, így q_n egyszerű függvények különbségeként maga is egyszerű függvény minden $n \in \mathbb{N}$-re. Az is látható, hogy a teleszkopikus összeget számolva

$$\sum_{k=1}^{n} q_k = s_{n+1} - s_1 = s_{n+1} \to f,$$

ha $n \to \infty$, azaz $\sum_{k=1}^{\infty} q_k = f$. No de minden egyes k mellett q_k is egyszerű függvény. Felírva a kanonikus alakját

$$q_k = \sum_{j=1}^{r_k} \alpha_j^{(k)} \chi_{E_j^{(k)}},$$

ahol $r_k \in \mathbb{N}$, $\alpha_j^{(k)} \in \mathbb{R}_+$ és $E_j^{(k)} \in \mathcal{M}$ mérhető halmazok. Ebből azonnal következik, hogy f a pontonkénti konvergencia értelmében

$$f = \sum_{k=1}^{\infty} q_k = \sum_{k=1}^{\infty} \sum_{j=1}^{r_k} \alpha_j^{(k)} \chi_{E_j^{(k)}}$$

alakú. Tudjuk, hogy nem negatív tagú sor konvergenciáját nem befolyásolja a tagok egy permutációja, így a fenti sort tetszőlegesen átindexelve kapjuk az

$$f = \sum_{n=1}^{\infty} \alpha_n \chi_{E_n}$$

előállítást. Ezt kellett belátni. $\square$

Láttuk tehát, hogy a legegyszerűbb mérhető függvény egy mérhető halmaz karakterisztikus függvénye. Az ilyenek nem negatív együtthatókkal vett lineáris kombinációját neveztük egyszerű vagy lépcsős függvénynek. Világos, hogy mérhető függvények határértéke is mérhető, emiatt a

$$\sum_{n=1}^{\infty} \alpha_n \chi_{E_n}$$

sor nem negatív α_n együtthatókkal és mérhető $E_n \in \mathcal{M}$ halmazokkal mindenképpen egy nem negatív mérhető függvényt ad. A fenti eredményünk azt jelenti, hogy ezzel az eljárással már minden nem negatív mérhető függvényt elő is állíthatunk.

2. fejezet

Integrál és konvergenciatételek

Az absztrakt integrált először egyszerű függvényekre definiáljuk. A mérhető függvények alaptétele szerint minden nem negatív mérhető függvény alulról közelíthető egyszerű függvénnyel. Ez lehetőséget ad arra, hogy a nem negatív mérhető függvény integrálját mint a nála nem nagyobb egyszerű függvények integráljának szuprémumát definiáljuk. Általánosabban, egy mérhető függvény integrálját úgy kapjuk, hogy a függvényt előállítjuk mint két nem negatív mérhető függvény különbsége, majd az integrál ezen függvények integráljai különbségeként adódik.

Az integrál bevezetése mellett a legfontosabb gondolat a konvergenciatételekbe van csomagolva. A kérdés az, hogy ha egy függvényt közelítünk egy függvénysorozattal, akkor a függvénysorozat elemeinek integrálja közelíti-e a függvény integrálját. Hasonlóan egy függvénysor összegfüggvényének integrálja előáll-e mint az integrálok sor-összege. Formálisan nézve, ez a limes-integrál jel cserélhetőségét és a szumma-integrál cserélhetőségét jelenti. Az itt definiált integrálfogalom legfőbb erénye, hogy általánossága mellett könnyen megjegyezhető, jól használható konvergenciatételeket kapunk, mind a limes-integrál, mind a szumma-integrál felcserélhetőségre vonatkozóan.

Legyen a fejezetben végig az $(X, \mathcal{M}, \mu)$ mértéktér rögzítve.

2.1. Egyszerű függvény integrálja

A szakaszban körüljárjuk az egyszerű függvények integráljának fogalmát. Megmutatjuk, hogy az integrál pozitív homogén, monoton és additív hozzárendelés. A legfontosabb gondolat, hogy egy egyszerű függvény integrálásával egy új mértéket konstruálhatunk.

2.1.1. definíció (egyszerű függvény integrálja)**.** Legyen $f : X \to \overline{\mathbb{R}}_+$ egy egyszerű függvény az $(X, \mathcal{M}, \mu)$ mértéktéren, amelynek kanonikus alakja

$$f = \sum_{i=1}^{n} \alpha_i \chi_{E_i},$$

és $E \in \mathcal{M}$ egy mérhető halmaz. Az f függvény E halmaz fölötti μ mérték szerinti absztrakt *integrálja* az

$$\int_E f \, d\mu = \sum_{i=1}^{n} \alpha_i \mu \left(E_i \cap E\right)$$

szám.

A fenti definícióban fontos a korábban bevezetett konvenció, miszerint $0 \cdot (+\infty) = 0$ és $\alpha \cdot (+\infty) = +\infty$ ha $\alpha > 0$. Azt látjuk, hogy a definícióban szereplő véges összegnek azért van értelme, mert $(+\infty) + (+\infty) = +\infty$ és az $\alpha + (+\infty) = +\infty$ minden $\alpha \geq 0$ esetén. Azért fontos, hogy csak a nem negatív függvényekre szorítkozzunk, hogy a fenti összegben a $\infty + (-\infty)$ művelet ne forduljon elő.

A definíció mögötti intuíció világos. Az integrálon a függvény görbéje alatti „területet” szeretnénk érteni. Egy lépcsős függvény esetén ezt véges sok téglalap területének gondoljuk. Persze nincs szó arról, hogy az $E_i \cap E$ mérhető halmaz bármilyen értelemben is összefüggő lenne, de mivel mérhető halmazként eleme a μ mérték értelmezési tartományának, ezért van valamekkora $\mu \left(E_i \cap E\right)$ mértéke. Abból az érzésünkből származik a definícióbeli $\alpha_i \mu \left(E_i \cap E\right)$ szorzat, hogy egy téglalap területét az alap és a magasság szorzataként értelmezzük.

Itt értjük meg, hogy a $0 \cdot (+\infty) = 0$ definícióra valóban rászorulunk, hiszen mondjuk az $[1, +\infty]$ alapú, de 0 magasságú téglalap területét is 0-nak gondoljuk.

Fontos speciális eset az egész X halmaz feletti integrálás. Definíció szerint ilyenkor

$$\int_X f \, d\mu = \sum_{i=1}^{n} \alpha_i \mu \left(E_i\right).$$

2.1.2.
Világos, hogy amennyiben f egy nem negatív lépcsős függvény, úgy $\int_E f \, d\mu \geq$ ≥ 0, és $\int_E f \, d\mu = +\infty$ is lehetséges. Viszont, ha az $E \in \mathcal{M}$ halmaz *nullmértékű*, azaz $\mu(E) = 0$, akkor $\int_E f \, d\mu = 0$ tekintet nélkül arra, hogy az f egyszerű függvény hogyan is van definiálva.

A következő kis megjegyzésünk nagyon egyszerű, de az integrál felépítése szempontjából fontos, ezért ki is emeljük.

2.1.3.
Amennyiben f egyszerű függvény az $E \in \mathcal{M}$ halmazon a nem negatív α konstans, akkor létezik i melyre $E \subseteq E_i$, ezért $E_j \cap E = \varnothing$ minden $j \neq i$-re, így

$$\int_E f\, d\mu = \alpha_i \mu\left(E_i \cap E\right) = \alpha \mu\left(E\right).$$

2.1.4. állítás (integrál-mérték)**.** *Legyen f egy lépcsős függvény. Ekkor a*

$$\varphi\left(E\right) = \int_E f\, d\mu$$

halmazfüggvény egy nem negatív mértéket definiál az $\mathcal{M}$ σ-algebrán. A φ mértéket az f lépcsős függvény integrál-mértékének *nevezzük.*

Bizonyítás. Az $\varnothing$ null-mértékű, tehát $\varphi\left(\varnothing\right) = 0$. Legyen $f = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$, az egyszerű függvény kanonikus alakja, továbbá tegyük fel, hogy F mérhető halmaz előáll $F = \cup_{j=1}^{\infty} F_j$ mérhető, diszjunkt halmazok egyesítéseként. Ekkor

$$\begin{aligned}
\varphi\left(F\right) = \varphi\left(\cup_{j=1}^{\infty} F_j\right) &= \int_{\cup_{j=1}^{\infty} F_j} f\, d\mu = \sum_{i=1}^{n} \alpha_i \mu\left(E_i \cap \left(\cup_{j=1}^{\infty} F_j\right)\right) = \\
&= \sum_{i=1}^{n} \alpha_i \mu\left(\cup_{j=1}^{\infty}\left(F_j \cap E_i\right)\right) = \sum_{i=1}^{n} \sum_{j=1}^{\infty} \alpha_i \mu\left(F_j \cap E_i\right) = \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^{n} \alpha_i \mu\left(E_i \cap F_j\right) = \\
&= \sum_{j=1}^{\infty} \int_{F_j} f\, d\mu = \sum_{j=1}^{\infty} \varphi\left(F_j\right).
\end{aligned}$$

Csupán μ mérték σ-additivitását használtuk és azt, hogy véges sok sor összegét a sorok tagonkénti összegeként alkotott sor is megadja. $\square$

Az integrálmérték additivitása következményeként kapjuk a tagonkénti integrálhatóság tételét.

2.1.5. állítás (az integrál additiv)**.** *Legyenek az s és a t lépcsős függvények az $(X, \mathcal{M}, \mu)$ mértéktéren. Ekkor*

$$\int_X \left(s + t\right)\, d\mu = \int_X s\, d\mu + \int_X t\, d\mu.$$

Bizonyítás. Legyen $s = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$ és $t = \sum_{j=1}^{m} \beta_j \chi_{F_j}$ a kanonikus előállítás. Ekkor $\cup_{j=1}^{m} F_j = X = \cup_{i=1}^{n} E_i$ partíciók. Világos, hogy

$$X = \cup_{i,j}^{n,m} (E_i \cap F_j)$$

is diszjunkt egyesítés. Az $E_i \cap F_j$ persze lehet üres, de még ekkor is igaz, hogy az $E_i \cap F_j$ halmazon az $s + t$ egyszerű függvény a konstans $\alpha_i + \beta_j$ értéket veszi fel. Hasonlóan az s és a t is konstans ezen a metszeten. Az előző 2.1.3. megjegyzésünk szerint ilyen halmazon az $s + t$ integrálja valamint az s és a t integrálja is könnyen számolható.

$$\int_{E_i \cap F_j} s + t \, d\mu = (\alpha_i + \beta_j) \, \mu (E_i \cap F_j) =$$
$$= \alpha_i \mu (E_i \cap F_j) + \beta_j \mu (E_i \cap F_j) = \int_{E_i \cap F_j} s \, d\mu + \int_{E_i \cap F_j} t \, d\mu.$$

Az $s + t$ függvény, majd az s és a t függvény integrál-mértékének végesen additivitása szerint

$$\int_X s + t \, d\mu = \sum_{i,j}^{n,m} \int_{E_i \cap F_j} s + t \, d\mu = \sum_{i,j}^{n,m} \left(\int_{E_i \cap F_j} s \, d\mu + \int_{E_i \cap F_j} t \, d\mu \right) =$$
$$= \sum_{i,j}^{n,m} \int_{E_i \cap F_j} s \, d\mu + \sum_{i,j}^{n,m} \int_{E_i \cap F_j} t \, d\mu = \int_X s \, d\mu + \int_X t \, d\mu. \qquad \square$$

Indukcióval azonnal következik, hogy véges sok lépcsős függvény összege tagonként integrálható.

Most azt gondoljuk meg, hogy az egyszerű függvény integráljának definíciója nem változik akkor sem, ha nem annak kanonikus alakjában írjuk fel a függvényt, hanem valahogy máshogyan, de karakterisztikus függvények nem negatív együtthatókkal vett lineáris kombinációjaként.

2.1.6. állítás (nem kanonikus alakú egyszerű függvény integrálja)**.** *Amennyiben az f egyszerű függvény*

$$f = \sum_{j=1}^{m} \beta_j \chi_{F_j}$$

indikátor függvények nem negatív együtthatós lineáris kombinációja, úgy

$$\int_E f \, d\mu = \sum_{j=1}^{m} \beta_j \mu (F_j \cap E) = \int_X f \chi_E \, d\mu$$

tetszőleges $E \in \mathcal{M}$ mérhető halmazra.

Bizonyítás. Most legyen $f = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$ a kanonikus előállítás. Rögzített i mellett tekintsük az $\alpha_i \chi_{E_i \cap E}$ függvényt. Világos, hogy ez maga is egy egyszerű függvény, melynek integráljára:

$$\int_X \alpha_i \chi_{E_i \cap E} = \alpha_i \mu \left(E_i \cap E\right).$$

Tudjuk már, hogy az egyszerű függvények összegét tagonként integrálhatjuk. Felírva tehát először a kanonikus alakra vonatkozó definíciót, majd használva a 2.1.5. állítást, azt kapjuk, hogy:

$$\int_E f \, d\mu = \sum_{i=1}^{n} \alpha_i \mu \left(E_i \cap E\right) = \sum_{i=1}^{n} \int_X \alpha_i \chi_{E_i \cap E} \, d\mu =$$
$$= \int_X \sum_{i=1}^{n} \alpha_i \chi_{E_i \cap E} \, d\mu = \int_X \left(\sum_{i=1}^{n} \alpha_i \chi_{E_i} \right) \chi_E \, d\mu = \int_X f \chi_E \, d\mu.$$

Most áttérve az állítás kimondásában felírt lineáris kombinációra, tagonkénti integrálással készen is vagyunk.

$$\int_X f \chi_E \, d\mu = \int_X \left(\sum_{j=1}^{m} \beta_j \chi_{F_j} \right) \chi_E \, d\mu =$$
$$= \int_X \sum_{j=1}^{m} \beta_j \chi_{F_j \cap E} \, d\mu = \sum_{j=1}^{m} \beta_j \mu \left(F_j \cap E\right). \qquad \square$$

Most vizsgáljuk meg az integrálnak a rendezéssel való kapcsolatát.

2.1.7. állítás (az integrál monoton és pozitív homogén)**.** *Legyen $\alpha \geq 0$ nem negatív valós szám és s egyszerű függvény. Ekkor*

$$\int_X \alpha s \, d\mu = \alpha \int_X s \, d\mu.$$

Ha t is egyszerű függvény, amelyre $s \leq t$, akkor

$$\int_X s \, d\mu \leq \int_X t \, d\mu.$$

Bizonyítás. Az első állítás a definíció egyenes következménye. Világos, hogy $t - s \geq 0$ miatt $t - s$ is egyszerű függvény. Persze az integrálja nem negatív. Alkalmazva az additivitást a $t = s + (t - s)$ összegre azt kapjuk, hogy

$$\int_X t \, d\mu = \int_X s \, d\mu + \int_X t - s \, d\mu \geq \int_X s \, d\mu. \qquad \square$$

2.2. Nem negatív mérhető függvények integrálja

Legyen továbbra is rögzítve az $(X, \mathcal{M}, \mu)$ mértéktér. Mérhető, $f : X \to \overline{\mathbb{R}}_+$ függvények integráljának fogalmát szeretnénk bevezetni úgy, hogy azt olyan egyszerű függvények integráljával közelíthessük, amelyek pontonként közelítik az integrálandó függvényt.

2.2.1. definíció (nem negatív mérhető függvény integrálja)**.** Legyen $f : X \to \overline{\mathbb{R}}_+$ egy nem negatív mérhető függvény az $(X, \mathcal{M}, \mu)$ mértéktéren.

$$\int_X f\, d\mu = \sup \left\{ \int_X s\, d\mu : 0 \leq s \leq f \text{ és } s \text{ egyszerű függvény} \right\}$$

az f függvény teljes X halmaz feletti absztrakt *integrálja* a μ mérték szerint. Ha integrandus argumentumát is meg akarjuk jelölni, akkor szokásos még a sokkal klasszikusabbnak számító,

$$\int_X f\, d\mu = \int_X f(x)\ \mu(dx)$$

jelölés is.

A szuprémum mögötti halmaznak a konstans zérus egyszerű függvény mindig eleme, ezért valóban nem üres halmaz szuprémumát kell vennünk. Emlékezzünk arra a megállapodásra, hogy a felsőhatár axiómát $\overline{\mathbb{R}}$-ban használjuk, ezért a fenti egyenlőség jobb oldala egy jól definiált nem negatív valós szám, vagy $+\infty$. Még az is lehetséges, hogy a szuprémum mögötti halmaznak eleme is a $+\infty$ érték.

2.2.2.
Mivel lépcsős függvényre már láttuk, hogy az integrál monoton, ezért a fenti definíció összhangban van azzal az esettel, ha a nem negatív mérhető függvény speciálisan maga is egy lépcsős függvény. Ebben az esetben sem definiáltuk felül az egyszerű függvények integráljának definícióját.

2.2.3.
Világos, hogy amennyiben f egy nem negatív mérhető függvény, úgy $\int_X f\, d\mu \geq 0$, és $\int_X f\, d\mu = +\infty$ is lehetséges. Szokás egy nem negatív mérhető függvényt *integrálhatónak* nevezni, ha integrálja véges valós szám.

Az integrál homogenitása könnyen következik az egyszerű függvények integráljának homogenitásából, a monotonitás a szuprémum operáció tulajdonsága, viszont az additivitáshoz már az egyik legfontosabb és legtöbbet használt konvergenciatétel kell.

2.2.4. állítás (az integrál monoton és pozitív homogén)**.** *Legyen $f : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvény és $\alpha \geq 0$ valós szám. Ekkor*

$$\int_X \alpha f\, d\mu = \alpha \int_X f\, d\mu.$$

Ha $g : X \to \overline{\mathbb{R}}_+$ egy másik nem negatív mérhető függvény, amelyre $f \le g$, akkor

$$\int_X f\, d\mu \le \int_X g\, d\mu.$$

Bizonyítás. Az $0 \cdot (+\infty) = 0$ azonosság szerint az $\alpha = 0$ esetén az állítás nyilvánvaló. Minden pozitív $\alpha > 0$ mellett az egyszerű függvények pozitív homogenitása (2.1.7) szerint

$$\int_X \alpha f\, d\mu =$$

$$= \sup\left\{\int_X s\, d\mu \colon s \le \alpha f,\ s \text{ egyszerű}\right\} = \sup\left\{\int_X s\, d\mu : \frac{1}{\alpha} s \le f,\ s \text{ egyszerű}\right\} =$$

$$= \sup\left\{\int_X \alpha t\, d\mu \colon t \le f,\ t \text{ egyszerű}\right\} = \sup\left\{\alpha \int_X t\, d\mu : t \le f,\ t \text{ egyszerű}\right\} =$$

$$= \alpha \cdot \sup\left\{\int_X t\, d\mu : t \le f,\ t \text{ egyszerű}\right\} = \alpha \int_X f\, d\mu.$$

Az integrál monotonitása annak következménye, hogy bővebb halmaz szuprémuma nem kisebb. ☐

Az integrál definíciójában a nem negatív mérhető függvény integrálját csak az egész X alaphalmaz felett definiáltuk. Mint egyszerű függvények esetén, most is szükségünk van egy mérhető részhalmazon vett integrálra is.

2.2.5. definíció. Az $f : X \to \overline{\mathbb{R}}_+$ mérhető függvénynek az $E \in \mathcal{M}$ mérhető részhalmazon az integrálját

$$\int_E f\, d\mu = \int_X f\chi_E\, d\mu$$

módon definiáljuk.

Mivel nem negatív egyszerű függvényekre a fenti egyenlőséget beláttuk (2.1.6), ezért ez a definíció is összhangban van azzal az esettel, amikor a fenti f egy egyszerű függvény.

2.2.6.
Tetszőleges f nem negatív mérhető függvényre és $E \in \mathcal{M}$ mérhető halmazra $f\chi_E \le f$, így az integrál monotonitását alkalmazva

$$\int_E f\, d\mu = \int_X f\chi_E\, d\mu \le \int_X f\, d\mu.$$

Egy mérhető halmazon konstans függvény e halmaz feletti integráljának kiszámítási módja is öröklődik egyszerű függvényről nem negatív mérhető függvényre. Ezt rögzítjük a következő észrevételben.

2.2.7.
Ha egy f nem negatív mérhető függvény, amely egy $E \in \mathcal{M}$ mérhető halmazon az α konstans értéket veszi fel, akkor az $f\chi_E$ függvény is egyszerű függvény, így

$$\int_E f\, d\mu = \int_X f\chi_E\, d\mu = \int_E f\chi_E\, d\mu = \alpha\mu\,(E)$$

továbbra is fennáll.

2.2.1. A monoton konvergenciatétel

A nem negatív mérhető függvények integráljának tulajdonságaiból a fentiekben annyit és csak annyit gondoltunk meg, amely a három legfontosabb konvergenciatétel közül az elsőhöz éppen szükséges. Definíció szerint a nem negatív mérhető függvény integrálja egy nem megszámlálható halmaz szuprémumaként definiált. Ezen könnyít a monoton konvergenciatétel azzal, hogy a mérhető függvények alaptétele (1.3.17) szerint megszámlálható halmaz szuprémumára vezeti vissza az integrált. Ennek a könnyítésnek segítségével tudjuk majd a tagonkénti integrálhatóságot biztosítani.

2.2.8. tétel (Beppo Levi-, vagy monoton konvergenciatétel)**.** *Tegyük fel, hogy az $(f_n)_{n\in\mathbb{N}}$ nem negatív mérhető függvények monoton növekedő sorozata, és jelölje f ezen sorozat pontonkénti határértékét. Láttuk, hogy f is nem negatív mérhető függvény. Ekkor*

$$\int_X f_n\, d\mu \to \int_X f\, d\mu,$$

midőn $n \to \infty$.

Bizonyítás. Mivel monoton növő $\overline{\mathbb{R}}$-beli számoknak van $\overline{\mathbb{R}}$-beli határértéke, ezért legyen $\alpha = \lim \int_X f_n\, d\mu \in \overline{\mathbb{R}}$. Mivel $\forall n \in \mathbb{N}$ mellett $\int_X f_n\, d\mu \le \int_X f\, d\mu$, ezért nyilvánvalóan $\alpha \le \int_X f\, d\mu$.

Most megmutatjuk, hogy $\int_X f\, d\mu \le \alpha$ is teljesül. Legyen g egy tetszőleges lépcsős függvény melyre $g \le f$, valamint $c \in (0{,}1)$ tetszőleges valós szám. Azt fogjuk megmutatni, hogy $c\int_X g\, d\mu \le \alpha$. Ennek kulcsa abban rejlik, hogy tetszőleges lépcsős függvény integrál-mértéke mérték (2.1.4) és minden mérték monoton folytonos (1.2.8). Tekintsük az

$$E_n = X\,(f_n \ge cg)$$

mérhető halmazt. Az f_n függvénysorozat pontonkénti monoton növekedése miatt minden $n \in \mathbb{N}$ mellett $E_n \subseteq E_{n+1}$. Az is látható, hogy $X = \cup_{n=1}^{\infty} E_n$, hiszen ha valamely $x \in X$ mellett $g\,(x) = 0$, akkor az f_n függvények nem negatívitása miatt $x \in E_n$ minden n mellett fennáll, ha viszont $g\,(x) > 0$, akkor mivel $g(x) \neq +\infty$, ezért $cg(x) < g(x)$, így létezik $n \in \mathbb{N}$ melyre $cg(x) < f_n(x)$, tehát $x \in E_n$ valóban fennáll. Alkalmazhatjuk most a $\varphi(E) = \int_E cg\, d\mu$ mérték monoton folytonosságát az E_n monoton növő halmazsorozatra. Így azt kapjuk, hogy

$$\varphi(E_n) \to \varphi(X) = \int_X cg\, d\mu.$$

No de $cg(x)\chi_{E_n}(x) \leq f_n(x)$, minden $x \in X$ mellett, ezért

$$\varphi(E_n) = \int_{E_n} cg\, d\mu = \int_X cg\chi_{E_n}\, d\mu \leq \int_X f_n\, d\mu \leq \alpha.$$

Ez épp azt jelenti, hogy $\int_X cg\, d\mu = c\int_X g\, d\mu \leq \alpha$ valóban fennáll. □

Végesen additív halmazfüggvénnyel nem lehet hasonló konvergenciatételt megfogalmazni. A bizonyítás nagyon erősen támaszkodik a μ halmazfüggvény σ-additivitására, hiszen az integrál-mérték monoton folytonossága (1.2.8) ennek következménye. Megjegyezzük, hogy a tényt, miszerint egy lépcsős függvénnyel való integrálás egy megszámlálhatóan additív halmaz függvény (2.1.4), most használtuk először annak teljes erejével, hiszen a lépcsős függvény tagonkénti integrálhatóságának igazolásához az integrál-mértéknek csak a végesen additivitását használtuk. A mérhető függvények tagonkénti integrálhatósága sem adódik a mérték σ-additivitásának felhasználása nélkül.

2.2.2. A monoton konvergenciatétel közvetlen következményei

A tagonkénti integrálhatóság az első fontos következmény. Ezzel együtt tehát tudjuk már, hogy a nem negatív mérhető függvények integrálása is pozitív homogén, additív függvény-operáció.

2.2.9. állítás (az integrál additivitása)**.** *Legyenek $f, g : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvények. Ekkor*

$$\int_X f + g\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu.$$

Bizonyítás. Láttuk, hogy az állítás teljesül egyszerű függvényekre (2.1.5). Legyen most $f = \lim_{n\to\infty} s_n$ és $g = \lim_{n\to\infty} t_n$, ahol s_n és t_n egyszerű függvények monoton növő sorozata. A mérhető függvények alaptétele (1.3.17) miatt ilyen s_n és t_n függvényeket valóban találunk. Nyilván $s_n + t_n \to f + g$ monoton növő módon. A monoton konvergenciatételt (2.2.8) háromszor alkalmazva azt kapjuk, hogy

$$\int_X f+g\,d\mu=\lim_{n\to\infty}\int_X s_n+t_n\,d\mu=\lim_{n\to\infty}\left(\int_X s_n\,d\mu+\int_X t_n\,d\mu\right)=$$
$$=\lim_{n\to\infty}\int_X s_n\,d\mu+\lim_{n\to\infty}\int_X t_n\,d\mu=\int_X f\,d\mu+\int_X g\,d\mu. \quad \square$$

A végtelen összeg tagonkénti integrálhatóságára vonatkozó tételt is szokás monoton konvergenciatételnek nevezni. Úgy is fogalmazhatjuk, hogy nem negatív mérhető függvények esetén a szumma és az integrál jel felcserélhető.

2.2.10. tétel (monoton konvergenciatétel sorokra)**.** *Legyenek f_n nem negatív mérhető függvények. Ekkor az $f=\sum_{n=1}^{\infty} f_n$ nem negatív mérhető függvényre*

$$\int_X f\,d\mu=\sum_{n=1}^{\infty}\int f_n\,d\mu.$$

Bizonyítás. Jelölje $g_n=\sum_{k=1}^{n} f_k$. Ekkor a g_n függvénysorozat monoton növekedőleg tart f-hez. Így a monoton konvergenciatétel és a tagonkénti integrálhatóság szerint

$$\sum_{k=1}^{n}\int_X f_k\,d\mu=\int_X\sum_{k=1}^{n} f_k\,d\mu=\int_X g_n\,d\mu\to\int_X f\,d\mu.$$

midőn $n\to\infty$. Ezt kellett belátni. $\square$

A következő állítás nem konvergens függvénysorozatokra általánosítja a Beppo Levi-tételt. Szokás ezt Fatou-lemmának is nevezni.

2.2.11. lemma (Fatou)**.** *Legyenek f_n nem negatív mérhető függvények. Ekkor*

$$\int_X \liminf_{n\to\infty} f_n\,d\mu\le\liminf_{n\to\infty}\int_X f_n\,d\mu.$$

Bizonyítás. Tudjuk, hogy $f=\liminf_{n\to\infty} f_n$ jelölés mellett $g_n\to f$ monoton növő módon, ahol $g_n=\inf_{k\ge n} f_k$. Egyrészt a monoton konvergenciatétel szerint

$$\int_X f\,d\mu=\lim_{n\to\infty}\int_X g_n\,d\mu,$$

másrészt $g_n\le f_n$ és az integrál monotonitása miatt

$$\lim_{n\to\infty}\int_X g_n=\liminf_{n\to\infty}\int_X g_n\le\liminf_{n\to\infty}\int_X f_n. \quad \square$$

Emlékezzünk, hogy egy egyszerű függvény integrál-mértéke (2.1.4) kulcs szerepet töltött be a monoton konvergenciatétel igazolásában. Fontos következménye a monoton konvergenciatételnek, hogy nem negatív mérhető függvény integrálásával kapott halmazfüggvény is mérték marad.

2.2.12. állítás (integrál-mérték)**.** *Legyen f nem negatív mérhető függvény. Ekkor a*

$$\varphi(E) = \int_E f \, d\mu$$

halmazfüggvény is mérték az $\mathcal{M}$ σ-algebrán. E φ mértéket az f által generált integrál-mértéknek *mondjuk.*

Bizonyítás. Ha $E = \cup_{n=1}^{\infty} E_n$ diszjunkt egyesítés, akkor $\chi_E = \sum_{n=1}^{\infty} \chi_{E_n}$. Ezért a monoton konvergenciatétel sorokra vonatkozó állítása szerint

$$\varphi(E) = \int_X f\chi_E \, d\mu = \int_X \sum_{n=1}^{\infty} f\chi_{E_n} \, d\mu = \sum_{n=1}^{\infty} \int_X f\chi_{E_n} = \sum_{n=1}^{\infty} \varphi(E_n).$$

Be kell még látnunk, hogy φ nem a konstans $+\infty$ halmazfüggvény. Világos, hogy $f\chi_\varnothing = 0$, ezért $\varphi(\varnothing) = \int_\varnothing f \, d\mu = \int_X f\chi_\varnothing \, d\mu = 0$. □

2.2.13. állítás (integrál-mérték szerinti integrál)**.** *Legyen $g : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvény az $(X, \mathcal{M}, \mu)$ mértéktéren, és*

$$\varphi(E) = \int_E g \, d\mu$$

a g integrál-mértéke. Ekkor tetszőleges $f : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvény integrálja az $(X, \mathcal{M}, \varphi)$ mértéktéren

$$\int_X f \, d\varphi = \int_X f \cdot g \, d\mu.$$

Bizonyítás. Az, hogy az állítás igaz abban a speciális esetben, amikor f egy mérhető halmaz karakterisztikus függvényének nem negatív konstans szorosa, a φ integrál-mérték definíciója, és az integrál homogenitásának következménye, hiszen $f = \alpha\chi_E$ jelöléssel:

$$\begin{aligned}\int_X f \, d\varphi &= \int_X \alpha\chi_E \, d\varphi = \alpha\varphi(E) = \\ &= \alpha \int_E g \, d\mu = \alpha \int_X \chi_E g \, d\mu = \int_X \alpha\chi_E g \, d\mu = \int_X fg \, d\mu.\end{aligned}$$

Tudjuk azt is, hogy minden f nem negatív mérhető függvény előáll mint megszámlálhatóan sok fenti típusú függvény sor-összege (1.3.18), azaz alkalmas $\alpha_n \geq 0$ számokkal és $E_n \in \mathcal{M}$ mérhető halmazokkal $f_n = \alpha_n \chi_{E_n}$ jelölés mellett
$$f = \sum_{n=1}^{\infty} f_n.$$
Az f_n függvényekre már bizonyítottuk az állítást, így a sorokra vonatkozó monoton konvergenciatételt (2.2.10) először az $(X, \mathcal{M}, \varphi)$ mértéktéren, másodszor az $(X, \mathcal{M}, \mu)$ mértéktéren alkalmazva
$$\int_X f\, d\varphi = \int_X \sum_{n=1}^{\infty} f_n\, d\varphi = \sum_{n=1}^{\infty} \int_X f_n\, d\varphi = \sum_{n=1}^{\infty} \int_X f_n g\, d\mu =$$
$$= \int_X \sum_{n=1}^{\infty} f_n g\, d\mu = \int_X \left(\sum_{n=1}^{\infty} f_n \right) g\, d\mu = \int_X f g\, d\mu. \quad \square$$

2.2.14.
A fenti tétel az oka a következő jelöléseknek. Azt, hogy a φ mérték a g függvény μ szerinti integrál-mértéke, azaz, hogy $\varphi : \mathcal{M} \to \overline{\mathbb{R}}_+$ definíciója
$$\varphi(E) = \int_E g\, d\mu$$
minden $E \in \mathcal{M}$ mérhető halmazra, szokás egyszerűbben
$$d\varphi = g\, d\mu$$
módon jelölni. Mi több, néha ezzel ekvivalens módon
$$\frac{d\varphi}{d\mu} = g$$
módon fejezzük ki azt, hogy φ a g függvény μ mérték szerinti integrál-mértéke. Ilyen módon működik a következő formalizmus:
$$\int_X f\, d\varphi = \int_X f \cdot \frac{d\varphi}{d\mu}\, d\mu.$$

2.2.15. (mértékek szorzata a téglák félgyűrűjén)
Tegyük fel, hogy adott két mértéktér az $(X, \mathcal{M}, \mu)$ és az $(Y, \mathcal{N}, \nu)$. A μ mérték az $\mathcal{M}$ és a ν mérték az $\mathcal{N}$ σ-algebrán van értelmezve, így a $\mu \times \nu$ *szorzat-halmazfüggvényt* természetes értelmezni az $\mathcal{M} \times \mathcal{N}$ szorzat-félgyűrűn (1.1.3) a
$$(\mu \times \nu)(A \times B) = \mu(A)\, \nu(B)$$
definícióval. Az $\mathcal{M} \times \mathcal{N}$ szorzat-félgyűrű elemeit szokás *mérhető tégláknak* nevezni.

Most megmutatjuk, hogy $\mu \times \nu$ egy σ-additív halmazfüggvény ezen a félgyűrűn. A szorzat-mérték definíciója (4.1.8) ezen az állításon alapul.

2.2.16. állítás (mértékek szorzata a téglák félgyűrűjén megszámlálhatóan additív)**.** *Legyenek* $(X, \mathcal{M}, \mu)$ *és* $(Y, \mathcal{N}, \nu)$ *mértékterek. Legyen* $Q = A \times B$, $Q_n = A_n \times B_n$ *és* $Q = \cup_{n=1}^{\infty} Q_n$ *diszjunkt egyesítés, ahol* $A, A_n \in \mathcal{M}$ *és* $B, B_n \in \mathcal{N}$ *minden* $n \in \mathcal{N}$ *mellett. Ekkor*

$$\mu(A)\,\nu(B) = \sum_{n=1}^{\infty} \mu(A_n)\,\nu(B_n),$$

azaz $\mu \times \nu : \mathcal{M} \times \mathcal{N} \to \overline{\mathbb{R}}_+$ *egy* σ*-additív halmazfüggvény az* $\mathcal{M} \times \mathcal{N}$ *félgyűrűn.*

Bizonyítás. A $Q = \cup_n Q_n$ diszjunkt előállítás miatt $\chi_Q = \sum_{n=1}^{\infty} \chi_{Q_n}$, így

$$\chi_Q(x, \cdot) = \sum_{n=1}^{\infty} \chi_{Q_n}(x, \cdot)$$

minden rögzített $x \in X$ mellett. Mivel a Q és a Q_n halmazok mérhető téglák, ezért rögzített $x \in X$ mellett

$$\chi_Q(x, \cdot) = \chi_A(x)\,\chi_B, \quad \chi_{Q_n}(x, \cdot) = \chi_{A_n}(x)\,\chi_{B_n}.$$

Így a $\chi_Q(x, \cdot)$ és a $\chi_{Q_n}(x, \cdot)$ függvények mérhetők az $\mathcal{N}$ σ-algebrára nézve. Értelmes tehát felírni a fenti függvények integrálját az $(Y, \mathcal{N}, \nu)$ mértéktéren. Bevezetve az

$$\varphi(x) = \int_Y \chi_Q(x, \cdot)\,d\nu = \int_Y \chi_A(x)\,\chi_B\,d\nu = \chi_A(x) \int_Y \chi_B\,d\nu = \chi_A(x)\nu(B),$$

$$\varphi_n(x) = \int_Y \chi_{Q_n}(x, \cdot)\,d\nu = \int_Y \chi_{A_n}(x)\,\chi_{B_n}\,d\nu =$$
$$= \chi_{A_n}(x) \int_Y \chi_{B_n}\,d\nu = \chi_{A_n}(x)\,\nu(B_n)$$

jelöléseket, majd a sorokra vonatkozó monoton konvergenciatételt (2.2.10) alkalmazva továbbra is rögzített $x \in X$ mellett kapjuk a

$$\varphi(x) = \int_Y \chi_Q(x, \cdot)\,d\nu = \sum_{n=1}^{\infty} \int_Y \chi_{Q_n}(x, \cdot)\,d\nu = \sum_{n=1}^{\infty} \varphi_n(x)$$

azonosságot. No de φ és φ_n nem negatív mérhető függvények az $\mathcal{M}$-re nézve, ezért a monoton konvergenciatétel sorokra vonatkozó alakját most a μ mértékre alkalmazva

$$\mu(A)\nu(B) = \int_X \varphi\,d\mu = \sum_{n=1}^{\infty} \int_X \varphi_n\,d\mu = \sum_{n=1}^{\infty} \mu(A_n)\nu(B_n). \qquad \square$$

2.2.3. Null-mértékű halmazok

A következő részben a null-mértékű halmazok szerepét vizsgáljuk. Tekintsük az $(X, \mathcal{M}, \mu)$ mértéktéren a null-mértékű halmazok

$$\mathcal{M}_0 = \{A \in \mathcal{M} : \mu(A) = 0\}$$

halmazrendszerét. Világos, hogy ez zárt a különbségre a mérték monotonitása (1.2.5) miatt, és a mérték σ-szubadditivitása (1.2.6) szerint zárt a megszámlálható egyesítésre is. Ez azt jelenti, hogy $\mathcal{M}_0$ az $\mathcal{M}$ σ-algebra egy rész-σ-gyűrűje.

A valós analízisben és a valószínűség-számításban kiemelt szerep jut a null-mértékű halmazoknak. Ennek oka az itt bevezetendő *majdnem mindenütt* terminológia, ami a minden pontban teljesülés gyengítése. A fontos konvergenciatételeinket ebben a gyengített formában is átgondoljuk.

2.2.17.
Láttuk, hogy egyszerű függvényeknek egy null-mértékű halmaz felett az integrálja zérus. Ha $A \in \mathcal{M}$ egy null-mértékű halmaz, és $s \le f\chi_A$, akkor $s = s\chi_A$, emiatt $\int_X s\,d\mu = \int_A s\,d\mu = 0$. Ez azt jelenti, hogy az $f\chi_A$ integráljának definíciójában a sup mögötti halmaznak egyetlen eleme van, a zérus. Azt gondoltuk meg tehát, hogy tetszőleges f nem negatív mérhető függvény és $A \in \mathcal{M}, \mu(A) = 0$ mellett

$$\int_A f\,d\mu = 0.$$

2.2.18. definíció (mérték abszolút folytonossága)**.** Legyen λ és μ két mérték (akár előjeles vagy komplex mérték) az $\mathcal{M}$ σ-algebrán. Azt mondjuk, hogy λ *abszolút folytonos* a μ-re nézve, ha minden $A \in \mathcal{M}, \mu(A) = 0$ esetén $\lambda(A) = 0$ is fennáll. Ezt $\lambda \ll \mu$ módon jelöljük.

Példaként látjuk, hogy amennyiben $d\lambda = f\,d\mu$ valamely mérhető f függvényre, akkor $\lambda \ll \mu$. Meg fogjuk mutatni, hogy ez, bizonyos mértékek körében, máshogy nem is fordulhat elő. Ez a Radon–Nikodym-tétel (5.3.5), ami anyagunk egyik legfontosabb gondolata. Végül is elmondhatjuk majd, hogy $\lambda \ll \mu$ úgy és csak úgy teljesülhet, hogy valamely f nem negatív mérhető függvénnyel $d\lambda = f\,d\mu$ fennáll.

2.2.19. definíció (majdnem mindenütt szóhasználat)**.** Tekintsünk egy olyan tulajdonságot, amely az $(X, \mathcal{M}, \mu)$ mértéktér pontjain van értelmezve. Azt mondjuk, hogy ez a tulajdonság *majdnem mindenütt* teljesül, ha van olyan null-mértékű halmaz, amely komplementerének minden pontjában a tulajdonság teljesül.

Azt, hogy a T tulajdonság a μ mértékre nézve majdnem mindenütt teljesül, szokás úgy rövidíteni, hogy T igaz μ-m.m. $x \in X$ mellett, vagy ha formálisabb környezetben kell használni, akkor $T(x), \forall x \in X[\mu]$.

Érdemes formalizálni a fenti definíciót. A T tulajdonság akkor teljesül az X majdnem minden pontjában, ha létezik $N \in \mathcal{M}, \mu(N) = 0$, melyre[1]

$$\{x \in X : T(x) \text{ nem igaz}\} \subseteq N.$$

A Dirichlet-függvény példája mutatja, hogy egy Riemann-integrálható függvényt akár csak megszámlálhatóan sok pontban megváltoztatva olyan függvényt kaphatunk, amely nem is Riemann-integrálható. Ezzel ellentétben az absztrakt integrálfogalom sokkal szebben viselkedik még a null-mértékű halmazokon történt változtatásokkal szemben is. Egyszerűen az történik, hogy az integrál érzéketlen a null-mértékű halmazon való mérhető változtatásra.

2.2.20.
Tegyük fel, hogy az f, g nem negatív mérhető függvényekre $f \leq g$ majdnem mindenütt. Ez azt jelenti, hogy létezik olyan A halmaz, melynek A^c komplementere null-mértékű, és $f\chi_A \leq g\chi_A$ most már a teljes X alaphalmazon. Mivel egy null-mértékű halmazon az integrál zérus, így

$$\begin{aligned}\int_X f\,d\mu = \int_A f\,d\mu + \int_{A^c} f\,d\mu = \int_X f\chi_A\,d\mu + 0 \leq \\ \leq \int_X g\chi_A\,d\mu + 0 = \int_A g\,d\mu + \int_{A^c} g\,d\mu = \int_X g\,d\mu.\end{aligned}$$

Látjuk tehát, hogy a monotonitás az $f \leq g$ feltétel helyett a kicsit gyengébb $f \leq g\,[\mu]$ feltétellel is működik.

Így, ha f és g nem negatív mérhető függvények, amelyek μ-majdnem minden pontban azonosak, azaz $f = g\,[\mu]$, akkor

$$\int_X f\,d\mu = \int_X g\,d\mu.$$

2.2.21. állítás. *Legyen $(X, \mathcal{M}, \mu)$ mértéktér, és $f : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvény.*

$$\int_X f\,d\mu = 0 \textit{ pontosan akkor, ha } f = 0 \ \mu\text{-m.m.}$$

[1] Apró finomságnak tűnik, de fontos, hogy az irodalom nem teljesen egységes a majdnem mindenütt szóhasználatot illetően. Az alternatív definíció, hogy a T tulajdonság μ-majdnem minden pontban teljesül, ha az

$$\{x \in X : T(x) \text{ nem igaz}\}$$

halmaz maga is mérhető és mértéke zérus. [Rudin (1987)] és [Folland (1999)] definíciója a mi definíciónk (2.2.19), de például [Browder (1996)] a fentit használja.

Bizonyítás. Ha $f = 0\,[\mu]$, akkor az f integrálja és a konstans zéró integrálja azonos.

Tegyük most fel, hogy az f függvény integrálja zérus. Tekintsük az $A_n = X\left(f \geq \frac{1}{n}\right)$ inverzképet. Világos, hogy $A_n \in \mathcal{M}$, és $f \geq \frac{1}{n}\chi_{A_n}$ mindenütt. Ezért

$$0 = \int_X f\, d\mu \geq \int_X \frac{1}{n}\chi_{A_n}\, d\mu = \frac{1}{n}\mu(A_n).$$

Tehát minden $n \in \mathbb{N}$ mellett $\mu(A_n) = 0$, ezért

$$\mu\left(X\left(f > 0\right)\right) = \mu\left(\cup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu\left(A_n\right) = 0. \qquad \square$$

A majdnem mindenütt pontonkénti konvergencia fogalmának fontosságát kiemelendő írjuk fel a 2.2.19. definíciót ebben a speciális esetben.

2.2.22. definíció (majdnem mindenütt pontonkénti konvergencia)**.** Legyenek adva az $f_n, f : X \to \overline{\mathbb{R}}_+$ nem negatív függvények ($n \in \mathbb{N}$). Azt mondjuk, hogy $\lim_{n\to\infty} f_n = f$ μ-majdnem mindenütt pontonként, ha létezik $E \in \mathcal{M}$ mérhető halmaz, melyre $\mu(E) = 0$, és minden $x \in E^c$ esetén $f_n(x) \to f(x)$. Jelölje ezt $f_n \to f\ [\mu]$ vagy $f = \lim_{n\to\infty} f_n\ [\mu]$.

A null-mértékű halmazok használatára példaként nézzük az eddigi két legfontosabb konvergenciatételünknek, a monoton konvergenciatételnek és a Fatou-lemmának, a majdnem mindenütt konvergenciára vonatkozó általánosítását. Mivel általában nem igaz, hogy mérhető függvények majdnem mindenütt pontonkénti határértéke is mérhető lenne, ezért a határfüggvény mérhetősége is a feltételek közé tartozik.

2.2.23. állítás (Beppo Levi-tétel, Fatou-lemma)**.** *Legyen $f_n, f : X \to \overline{\mathbb{R}}_+$ mérhető függvények $\forall n \in \mathbb{N}$, amelyre*

$$\lim_{n\to\infty} f_n = f$$

μ-majdnem mindenütt pontonként X-en. Ekkor a
Fatou-lemma:

$$\int_X f\, d\mu \leq \liminf_{n\to\infty} \int_X f_n\, d\mu,$$

monoton konvergenciatétel: *feltéve, hogy $f_n \leq f_{n+1}\ [\mu]$, minden $n \in \mathbb{N}$ mellett fennáll,*

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_X f_n\, d\mu.$$

Bizonyítás. Nézzük a monoton konvergenciatétel itteni alakját. Ha $f_n \leq f_{n+1}\ [\mu]$, az azt jelenti, hogy létezik $E_n \in \mathcal{M}$ halmaz melynek $x \in E_n$

pontjaiban $f_n(x) \leq f_{n+1}(x)$ és $\mu(E_n^c) = 0$. Mivel $f_n \to f\,[\mu]$ is fennáll, azért létezik $E \in \mathcal{M}$, hogy minden $x \in E$ mellett $f_n(x) \to f(x)$ és $\mu(E^c) = 0$. Világos, hogy

$$A = E \cap \left(\bigcap_{n=1}^{\infty} E_n\right)$$

halmazra $A \in \mathcal{M}$, $f_n \chi_A \to f \chi_A$ monoton növekedőleg már a teljes X halmazon, és $\mu(A^c) \leq \mu(E^c) + \sum_{n=1}^{\infty} \mu(E_n^c) = 0$, a mérték σ-szubadditivitása szerint. Alkalmazva a monoton konvergenciatételt (2.2.8) azt kapjuk, hogy

$$\int_X f_n \, d\mu = \int_X f_n \chi_A \, d\mu \to \int_X f \chi_A \, d\mu = \int_X f \, d\mu,$$

felhasználva, hogy az integrál értéke érzéketlen a null-mértékű halmazon való konstans megváltoztatásra (2.2.20).

A Fatou-lemma igazolása a monoton konvergencia itteni alakjából a pontonkénti Fatou-lemma indoklásával (2.2.11) analóg. □

Hasznos gyakorlatként próbáljuk meg igazolni a Fatou-lemma fenti alakjából a monoton konvergenciatételt! Itt szeretném megjegyezni, hogy az eddigi két legfontosabb konvergenciaeredményünk a Fatou-lemma és a monoton konvergenciatétel. Felépítésünkben a monoton konvergenciatétel az elsődleges, következményeként tárgyaltuk a Fatou-lemmát. Más felépítést választva a Fatou-lemmát lehet előbb igazolni, majd abból a monoton konvergenciatételt. A két konvergenciaeredmény ekvivalens egymással.

Az eddigiekben szándékosan nem beszéltünk két függvény különbségének integráljáról. Ennek elsősorban az az oka, hogy $f, g : X \to \overline{\mathbb{R}}_+$ esetén az $f - g$ nem feltétlen értelmes, hiszen elképzelhető, hogy ugyanabban a pontban mindkét függvény épp a $+\infty$ értéket veszi fel. Az alábbi gondolat fényében ez a helyzet csak null-mértékű halmazon fordulhat elő, feltéve a kivonandó függvény integrálhatóságát. Ez meg is oldja a problémát, hiszen null-mértékű halmazon tetszőleges mérhető függvényre, például valamilyen konstansra átdefiniálva a függvények értékét, a különbség integrálja ugyanaz marad.

2.2.24. állítás (integrálható függvény majdnem mindenütt véges)**.**
Legyen f egy integrálható nem negatív függvény. Ekkor f majdnem minden pontban véges, azaz $\mu(X(f = +\infty)) = 0$, valamint minden $\alpha > 0$ mellett $\mu(X(f > \alpha)) < \infty$.

Bizonyítás. Legyen $A = X(f = \infty)$. Világos, hogy minden $n \in \mathbb{N}$ természetes szám mellett $n \cdot \chi_A \leq f$. Ezért $n\mu(A) \leq \int_X f \, d\mu < \infty$ minden n mellett, amiből persze $\mu(A) = 0$ következik.

Legyen most rögzített $\alpha > 0$ mellett $A = X(f > \alpha) \in \mathcal{M}$. Világos, hogy $\alpha \chi_A \leq f$, emiatt $\alpha \mu(A) = \int_X \alpha \chi_A \, d\mu \leq \int_X f \, d\mu < \infty$. Ez csak $\mu(A) < +\infty$ esetben lehetséges. □

Ha tehát $f \geq 0$ egy integrálható függvény, akkor $X = X(f = \infty) \cup \cup X(f > 0) \cup X(f = 0)$, ahol az első halmaz null-mértékű, a második halmaz pedig előáll megszámlálhatóan sok véges mértékű halmaz egyesítéseként.

2.2.25.
Ha $f, g : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvények, melyekre $f \geq g$, és g integrálható, azaz $\int_X g\,d\mu < \infty$, akkor $f - g$ majdnem minden pontban értelmes, ahol nem értelmes, ott tetszőlegesen választott konstansra állítva az értékét $f - g$ mérhető, és integráljára

$$\int_X f - g\,d\mu = \int_X f\,d\mu - \int_X g\,d\mu.$$

Ugyanis fennáll az $f = (f - g) + g$ azonosság, ahol a két jobb oldali függvény nem negatív. Tagonként integrálva, és az átrendezéshez felhasználva g integrálhatóságát készen is vagyunk.

Ezzel a monoton konvergenciatétel fogyó függvényekre vonatkozó alakja már visszavezethető a növő alakra: *Legyen $f_n : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvények majdnem mindenütt monoton fogyó sorozata, azaz $f_n \geq \geq f_{n+1}\ [\mu], \forall n \in \mathbb{N}$. Tegyük fel, hogy f_1 integrálható, azaz*

$$\int_X f_1\,d\mu < \infty,$$

és tegyük fel, hogy az f majdnem mindenütt pontonkénti határfüggvény is mérhető. Ekkor

$$\int_X f_n\,d\mu \to \int_X f\,d\mu.$$

Igazolásként elég annyit megjegyezni, hogy térjünk át az $f_1 - f_n$ nem negatív mérhető függvények monoton növő sorozatára, alkalmazzuk a monoton konvergenciatételt (2.2.23), majd az átrendezéshez használjuk ki, hogy $\int_X f_n\,d\mu, \int_X f\,d\mu < \infty$.

A csökkenő eset új feltétele tehát, hogy van a sorozatnak integrálható majoránsa. Ki fog derülni, hogy ez a feltétel a Lebesgue konvergenciatétel feltétele is (2.3.21).

A következő kis állítás valamilyen mértékelméleti skatulya-elvet fejez ki és valószínűség-számítási környezetben különösen sokszor használatos. Ha adva van egy megszámlálható halmazrendszer, melyek mérték-összege véges, akkor a tér majdnem minden pontja ezen halmazok közül csak véges sokhoz tartozhat. Borel–Cantelli-lemma néven szokás erre a gondolatra hivatkozni.

2.2.26. definíció (halmaz-sorozat limesz szuperiorja)**.** Tekintsük az $A_n \subseteq X$ halmaz-sorozatot. E halmaz-sorozat *limes superior-ja* az a halmaz, amely azon $x \in X$ pontokat tartalmazza, amelyekre az $x \in A_n$ tartalmazás végtelen sok különböző n mellett teljesül. Magyarul:

$$\limsup_n A_n = \cap_{n=1}^{\infty} \cup_{k \geq n} A_k.$$

2.2.27. lemma (Borel–Cantelli)**.** *Tekintsünk egy* $(X, \mathcal{M}, \mu)$ *mértékteret. Legyenek* $A_n \in \mathcal{M}, n \in \mathbb{N}$ *mérhető halmazok. Ekkor* $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ *esetén* $\mu(\limsup_n A_n) = 0$.

Bizonyítás. Tekintsük az $f = \sum_{n=1}^{\infty} \chi_{A_n}$ nem negatív függvényt. A monoton konvergenciatétel (2.2.10) szerint a sor tagonként integrálható, így

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} \mu(A_n) < \infty.$$

Ekkor f majdnem mindenütt véges (2.2.24), azaz $\mu(X(f = +\infty)) = 0$. No de világos, hogy egy x mellett $f(x)$ pontosan akkor véges, ha $x \in A_n$ csak véges sok n mellett teljesül. Így $X(f = +\infty) = \limsup_n A_n$. Ezt kellett belátni. □

Mértékterek teljessége

Láttuk korábban, hogy ha úgy változtatunk meg egy nullmértékű halmazon egy nem negatív mérhető függvényt, hogy annak mérhetősége azért megmaradjon, akkor e változtatás nincs hatással a függvény integráljának értékére. Az az ezzel kapcsolatban kialakult érzésünk, hogy egy null-mértékű halmaz olyan kicsi, hogy az e feletti viselkedés érdektelen az integrálás szempontjából, ezért akár figyelmen kívül is hagyható, ignorálható. Ez egy fontos gondolat, amelyet rendszeresen használunk. Éppen emiatt vezettük be a *majdnem minden* szóhasználatot is. Egyetlen apróságra viszont fel kell hívni a figyelmet, aminek pontatlan kezelése nagy veszedelmet tartogat. Ha egy halmazt elhanyagolunk annak kicsisége miatt, akkor rejtve azzal a feltevéssel élünk, hogy az elhanyagolható halmazunk egy részhalmaza is elhanyagolható, ami nem feltétlenül van így.

Nézzünk például egy $A \in \mathcal{M}$ mérhető halmazt, ami még null-mértékű is, tehát $\mu(A) = 0$, de van nem mérhető $E \notin \mathcal{M}, E \subseteq A$ részhalmaza. Ekkor $\chi_A = \chi_E\ [\mu]$, hiszen A^c halmazon mindkét függvény azonosan zérus. A χ_E függvény integrálját viszont nincs jogunk felírni, hiszen az integrál csak mérhető függvények esetén definiált. Az történt, hogy a χ_A függvényt az értelmezési tartománya egy mérhető részhalmazán, az A halmazon úgy definiáltuk felül, hogy a függvény elvesztette a mérhetőségét. Nem feltétlenül igaz tehát, hogy egy ignorálható halmaz része még inkább ignorálható lenne.

Az ignorálható halmazok szerepét a null-mértékű halmazok játsszák. Összesen arról van tehát szó, hogy vajon minden null-mértékű halmaz minden részhalmaza null-mértékű marad-e? A válasz, hogy attól függ. Ez a mértéktér egy tulajdonsága. Van olyan mértéktér, ami ilyen, és van olyan, ami nem ilyen. A legjobb lenne persze, ha minden mérték jó lenne ebből a célból. Látni fogjuk később, hogy például a hosszúságfogalom általánosítását jelentő Lebesgue-mérték jó is. Viszont sokszor fordul elő, hogy ezzel a tulajdonsággal nem rendelkező mértékekkel kell bánnunk, ami sok-sok komplikáció forrása.

Kicsit más oldaláról nézve ugyanezt a jelenséget, a fenti kis példa arról szól, hogy ha két függvény majdnem minden pontban megegyezik, akkor az egyik mérhetősége nem elegendő a másik mérhetőségéhez. Eddig a pontig a szakasz állításaira úgy is tekinthettünk, hogy korábban igazolt gondolatokat kiterjesztettük minden pontban teljesülő tulajdonság helyett csak a majdnem minden pontban teljesülés esetére. Fontos kivétel a mérhető függvények azonosítására, például a határértékükre vonatkozó állítás. Ha az előbbi kis példában $f_n = \chi_A$ minden $n \in \mathbb{N}$ mellett, és $f = \chi_E$, akkor az f_n mérhető függvények sorozata majdnem mindenütt pontonként konvergál az f határfüggvényhez, és f mégsem mérhető! Ez a jelenség az oka annak, hogy a monoton konvergenciatétel és a Fatou-lemma majdnem mindenütt konvergens alakjánál a feltételek közé kellett bevennünk a pontonként majdnem minden pontban létező határfüggvény mérhetőségét is.

2.2.28. definíció (teljes mértéktér)**.** Egy $(X, \mathcal{M}, \mu)$ mértékteret *teljesnek* nevezünk, ha minden null-mértékű halmaz tetszőleges részhalmaza mérhető. Magyarul: $E \subseteq F \in \mathcal{M}$ és $\mu(F) = 0 \Rightarrow E \in \mathcal{M}$.

Persze, ha egy null-mértékű halmaz egy részhalmaza mérhető, akkor az a mérték monotonitása (1.2.5) miatt már null-mértékű is lesz.

2.2.29. állítás (teljes mértéktér karakterizációja függvények mérhetőségével)**.** *Az alábbi feltételek ekvivalensek:*

1. *Az $(X, \mathcal{M}, \mu)$ mértéktér teljes;*

2. *Tetszőleges X-en értelmezett f, g függvény esetén, ha $f = g\ [\mu]$ és f mérhető, akkor g is mérhető;*

3. *Tetszőleges X-en értelmezett f_n függvénysorozat és f függvény esetén, ha $f_n \to f\ [\mu]$ és f_n mérhető minden $n \in \mathbb{N}$ mellett, akkor f is mérhető.*

Bizonyítás. Körbe bizonyítunk, tehát a) $1 \Rightarrow 2$, b) $2 \Rightarrow 3$ és c) $3 \Rightarrow 1$.
a) Legyen $A \in \mathcal{M}$ olyan halmaz melynek pontjaiban $f = g$ és $\mu(A^c) = 0$. Ekkor minden az értékkészletben mérhető E halmazra

$$g^{-1}(E) = \big(g^{-1}(E) \cap A\big) \cup \big(g^{-1}(E) \cap A^c\big) = \big(f^{-1}(E) \cap A\big) \cup \big(g^{-1}(E) \cap A^c\big)\,.$$

A jobb oldali első halmaz az f függvény mérhetősége miatt mérhető, míg a második halmaz egy nullmértékű halmaz részhalmazaként mérhető.

b) Most legyen $A \in \mathcal{M}$ olyan halmaz melynek pontjaiban $f_n \to f$ és $\mu(A^c) = 0$. Ekkor $f_n \chi_A \to f \chi_A$ most már az X minden pontjában, ezért $f \chi_A$ mérhető függvények pontonkénti határértékeként mérhető. Persze $f \chi_A = f\ [\mu]$ így a feltevés szerint f valóban mérhető.
c) Legyen $E \subseteq A$, ahol $A \in \mathcal{M}$ melyre $\mu(A) = 0$. Definiálja minden $n \in \mathbb{N}$ mellett $f_n = \chi_A$ és $f = \chi_E$. Ekkor $f_n \to f\ [\mu]$, emiatt a χ_E függvény mérhető, tehát $E \in \mathcal{M}$. □

A mértéktér teljessége tehát azt a többletet fejezi ki, hogy egy null-mértékű halmazon tetszőlegesen megváltoztatva egy mérhető függvényt az mérhető marad, míg általában ez csak a mérhető változtatásokra marad igaz. Emiatt teljes mértéktér esetében a majdnem mindenütt konvergens monoton konvergenciatétel és a Fatou-lemma feltétételei közé nem kell bevenni a határfüggvény mérhetőségét.

Most azt mutatjuk meg, hogy egy nem teljes mértéktér beágyazható egy teljes mértéktérbe. Ennek eredménye az lesz, hogy amennyiben van egy nem teljes mértékterünk, melyen két függvény majdnem mindenütt azonos, akkor – ha ugyan nem is következik az egyik függvény mérhetőségéből a másik mérhetősége – azért valamilyen gyengébb értelmű implikációról mégis beszélhetünk. Konkrétan a Borel-mérhetőség és a Lebesgue-mérhetőség kapcsolatának megértéséhez lesz fontos ez az állítás.

2.2.30. állítás (mérték teljessé tétele)**.** *Minden $(X, \mathcal{M}, \mu)$ mértéktérhez létezik egyetlen őt tartalmazó legszűkebb $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$ teljes mértéktér, amelyre $\mathcal{M} \subseteq \bar{\mathcal{M}}$, és $\mu(E) = \bar{\mu}(E)$ minden $E \in \mathcal{M}$ esetén. Ezt nevezzük a* mértéktér teljessé tételének.

Bizonyítás. Adott $(X, \mathcal{M}, \mu)$ mértéktérből kiindulva legyen

$$\bar{\mathcal{M}} = \{E \subseteq X : \exists A, B \in \mathcal{M},\ A \subseteq E \subseteq B \text{ és } \mu(B \setminus A) = 0\}. \tag{2.1}$$

Könnyen látható, hogy $\mathcal{M} \subseteq \bar{\mathcal{M}}$ és $\bar{\mathcal{M}}$ egy σ-algebra. Ugyanis, ha $A_n \subseteq E_n \subseteq B_n$, akkor $A = \cup_n A_n, E = \cup_n E_n, B = \cup_n B_n$ jelöléssel $A \subseteq E \subseteq B$ és $B \setminus A \subseteq \cup_n (B_n \setminus A_n)$, amiből következik a megszámlálható egyesítésre való zártság. Ha $A \subseteq E \subseteq B$, akkor $B^c \subseteq E^c \subseteq A^c$ és $A^c \setminus B^c = B \setminus A$ mutatja a komplementer zártságot.

Amennyiben $E \in \bar{\mathcal{M}}$, valamint $A, B \in \mathcal{M}$, $A \subseteq E \subseteq B$, $\mu(B \setminus A) = 0$ és van egy másik $A_1, B_1 \in \mathcal{M}$, $A_1 \subseteq E \subseteq B_1$, $\mu(B_1 \setminus A_1) = 0$ előállítás is, akkor $A_1 \setminus A \subseteq E \setminus A \subseteq B \setminus A$, ezért $\mu(A_1 \setminus A) = 0$. Hasonlóan kapjuk, hogy $\mu(A \setminus A_1) = 0$ is fennáll, amiből az $A = (A \cap A_1) \cup (A \setminus A_1)$ és $A_1 = (A_1 \cap A) \cup (A_1 \setminus A)$ egyenlőségek miatt

$$\mu(A) = \mu(A \cap A_1) = \mu(A_1 \cap A) = \mu(A_1)$$

is teljesül. Azt kaptuk tehát, hogy $E \in \bar{\mathcal{M}}$ esetén egy olyan $A \in \mathcal{M}$ halmazra, melyhez létezik $B \in \mathcal{M}, A \subseteq E \subseteq B, \mu(B \setminus A) = 0$ a

$$\bar{\mu}(E) = \mu(A)$$

definíció jól definiálja a $\bar{\mu}$ halmazfüggvényt az $\bar{\mathcal{M}}$ σ-algebrán.

Látható, hogy $\bar{\mu}$ egy σ-additív kiterjesztése μ-nek $\mathcal{M}$-ről $\bar{\mathcal{M}}$-ra, hiszen, ha az E_n diszjunkt halmazokra $A_n \subseteq E_n \subseteq B_n$, akkor az A_n halmazok is páronként diszjunktak, így

$$\bar{\mu}(E) = \mu(\cup_n A_n) = \sum_n \mu(A_n) = \sum_n \bar{\mu}(E_n).$$

Világos, hogy ha $F \subseteq E$ és $\bar{\mu}(E) = 0$, akkor létezik $A, B \in \mathcal{M}$, amelyekre $A \subseteq E \subseteq B$ valamint $\mu(B \setminus A) = 0$, és $\mu(A) = 0$. Így $B = A \cup (B \setminus A)$ miatt $\mu(B) = 0$ is teljesül, ami azt jelenti, hogy találtunk $\varnothing, B \in \mathcal{M}$ halmazokat, amelyekre $\varnothing \subseteq F \subseteq B$ és $\mu(B \setminus \varnothing) = 0$, azaz $F \in \bar{\mathcal{M}}$. Ez éppen azt jelenti, hogy a definiált $\bar{\mu}$ mérték teljes.

Tegyük most fel, hogy van egy $(X, \mathcal{N}, \nu)$ teljes mértéktér is, amely szintén kiterjesztése az $(X, \mathcal{M}, \mu)$ mértéktérnek. Ha $E \in \bar{\mathcal{M}}$, akkor $E = A \cup (E \setminus A)$ előállításban $A \in \mathcal{M} \subseteq \mathcal{N}$ és $E \setminus A \subseteq B \setminus A$ miatt az $E \setminus A$ halmaz részhalmaza a ν-null mértékű $B \setminus A$ halmaznak, amiből az $(X, \mathcal{N}, \nu)$ teljessége miatt $E \setminus A \in \mathcal{N}$ is következik. Így tehát $\bar{\mathcal{M}} \subseteq \mathcal{N}$, ami azt jelenti, hogy valóban a fent konstruált $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$ a legszűkebb teljes kiterjesztése az $(X, \mathcal{M}, \mu)$ mértéktérnek. $\square$

A fenti tételben nagyon fontos a teljessé tett mértéktér konstrukciója is. Emiatt rögzítjük a konstrukció egy variánsát.

2.2.31. állítás (teljessé tett σ-algebra)**.** *Legyen $(X, \mathcal{M}, \mu)$ mértéktér és $\left(X, \bar{\mathcal{M}}, \bar{\mu},\right)$ annak teljessé tétele. Az $E \subseteq X$ halmazra $E \in \bar{\mathcal{M}}$ pontosan akkor teljesül, ha minden $\varepsilon > 0$-hoz létezik $A, B \in \mathcal{M}$, $A \subseteq E \subseteq B$, melyre $\mu(B \setminus A) < \varepsilon$.*

Bizonyítás. Láttuk a teljessé tett σ-algebra definiálása során (2.1), hogy ha $E \in \bar{\mathcal{M}}$, akkor még $\mu(B \setminus A) = 0$, $A \subseteq E \subseteq B$, $A, B \in \mathcal{M}$ is található.

Válasszunk minden $n \in \mathbb{N}$ mellett $A_n, B_n \in \mathcal{M}$ halmazokat, melyekre $A_n \subseteq E \subseteq B_n$ és $\mu(B_n \setminus A_n) < \frac{1}{n}$. Legyen $A = \cup_{n=1}^{\infty} A_n$ és $B = \cap_{n=1}^{\infty} B_n$. Persze $A \subseteq E \subseteq B$, $A, B \in \mathcal{M}$, és $B \setminus A \subseteq B_n \setminus A_n$ minden $n \in \mathbb{N}$ mellett. Ebből $\mu(B \setminus A) \leq \mu(B_n \setminus A_n) < \frac{1}{n}$ miatt csak $\mu(B \setminus A) = 0$ lehetséges, tehát $E \in \bar{\mathcal{M}}$ valóban teljesül. $\square$

2.2.32. állítás (mérhetőség a teljessé tett σ-algebrára)**.** *Legyen $(X, \mathcal{M}, \mu)$ mértéktér és $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$ annak teljessé tétele.*

1. *Az $f : X \to \overline{\mathbb{R}}_+$ nem negatív függvény pontosan akkor mérhető az $\bar{\mathcal{M}}$ σ-algebrára nézve, ha létezik $h_1, h_2 : X \to \overline{\mathbb{R}}_+$ nem negatív $\mathcal{M}$-mérhető függvény, amelyekre fennáll a $h_1 \le f \le h_2$ egyenlőtlenség és $h_1 = = h_2\,[\mu]$.*

2. *Az $f : X \to \overline{\mathbb{R}}$ függvény pontosan akkor mérhető az $\bar{\mathcal{M}}$ σ-algebrára nézve, ha létezik $h : X \to \overline{\mathbb{R}}$ az $\mathcal{M}$ σ-algebrára nézve mérhető függvény, hogy $f = h\ [\mu]$.*

Bizonyítás. Először a nem negatív függvényekre vonatkozó állítást, majd az általános esetet indokoljuk.

1) A mérhető függvények alaptétele szerint (1.3.18) a nem negatív, $\bar{\mathcal{M}}$-ra nézve mérhető függvény

$$f = \sum_{n=1}^{\infty} \alpha_n \chi_{E_n}$$

alakú, ahol $\infty > \alpha_n \ge 0$ és $E_n \in \bar{\mathcal{M}}$. A teljesség konstrukciója szerint (l. (2.1)) léteznek $A_n \subseteq E_n \subseteq B_n$, $A_n, B_n \in \mathcal{M}$ halmazok, melyekre $\mu\,(B_n \setminus A_n) = 0$. Jelölje

$$h_1 = \sum_{n=1}^{\infty} \alpha_n \chi_{A_n}, \qquad h_2 = \sum_{n=1}^{\infty} \alpha_n \chi_{B_n}.$$

Világos, hogy h_1, h_2 az $\mathcal{M}$ σ-algebrára mérhető függvények határértékeként maga is $\mathcal{M}$-mérhető, és $\chi_{A_n} \le \chi_{E_n} \le \chi_{B_n}$ miatt $h_1 \le f \le h_2$. Na most, ha egy $x \in X$ pontra $h_2(x) > h_1(x)$, akkor legalább egy $n \in \mathbb{N}$ mellett $1 = \chi_{B_n}(x) > \chi_{A_n}(x) = 0$, ami azt jelenti, hogy $x \in B_n \setminus A_n$. Ez azt jelenti, hogy

$$X\,(h_1 \ne h_2) \subseteq \bigcup_{n=1}^{\infty} B_n \setminus A_n,$$

amiből persze $\mu\,(X\,(h_1 \ne h_2)) \le \sum_{n=1}^{\infty} \mu\,(B_n \setminus A_n) = 0$ következik.

Megfordítva tegyük fel, hogy adottak a h_1, h_2 függvények, amelyek mérhetők az $\mathcal{M}$-re nézve, $h_1 \le f \le h_2$ és $h_1 = h_2\ [\mu]$. Létezik tehát $A \in \mathcal{M}$ halmaz, hogy minden $x \in A$ mellett $h_1\,(x) = f\,(x)$ és $\mu(A^c) = 0$. Az $\mathcal{M} \subseteq \bar{\mathcal{M}}$ miatt h_1 $\bar{\mathcal{M}}$-mérhető is, és $f = h_1\ [\bar{\mu}]$. Az $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$ mértéktér teljessége szerint f valóban $\bar{\mathcal{M}}$-mérhető (2.2.29).

2) Most legyen f egy $\bar{\mathcal{M}}$-mérhető függvény. Ez előáll a pozitív része és a negatív része $f = f^+ - f^-$ különbségeként (1.3.15), ahol $f^+, f^- : X \to \overline{\mathbb{R}}_+$ nem negatív $\bar{\mathcal{M}}$-mérhető függvények. Alkalmazzuk a már igazolt állítást. Létezik $h_1 : X \to \overline{\mathbb{R}}_+$ nem negatív $\mathcal{M}$-mérhető függvény, melyre $h_1\,(x) \le \le f^+\,(x)$ minden $x \in X$ mellett, de $f^+ = h_1\ [\mu]$. Ugyanígy, van olyan $g_1 : X \to \overline{\mathbb{R}}_+$ nem negatív $\mathcal{M}$-mérhető függvény, melyre $g_1\,(x) \le f^- - (x)$ minden $x \in X$-re, de $f^- = g_1\ [\mu]$. Az $f^+\,(x) - f^-\,(x)$ különbség

minden x mellett értelmes, emiatt a $h_1(x) - g_1(x)$ is értelmes a tér minden pontjában. Világos, hogy $h_1 - g_1$ egy $\mathcal{M}$-mérhető függvény, amelyre $f = h_1 - g_1\ [\mu]$.

Megfordítva, ha $f = h\ [\mu]$, ahol $h : X \to \overline{\mathbb{R}}$ $\mathcal{M}$-mérhető függvény, akkor h $\bar{\mathcal{M}}$-mérhető is és $f = h\ [\bar{\mu}]$. Így az $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$ mértéktér teljessége szerint f $\bar{\mathcal{M}}$-mérhető (2.2.29). □

A következő állításban nem engedjük meg, hogy a szereplő függvények a $+\infty$ vagy a $-\infty$ értékeket is felvegyék. Az előző állításnál könnyebben ellenőrizhető, de csak elegendő feltételt kapunk a teljessé tett σ-algebrára vonatkozó mérhetőségre.

2.2.33. állítás (elegendő fetétel a teljessé tett mérhetőségre)**.** *Jelölje az* $(X, \mathcal{M}, \mu)$ *mértéktér teljessé tételét* $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$*, és legyen* $f : X \to \mathbb{R}$ *egy függvény. Tegyük fel, hogy minden* $\varepsilon > 0$ *számhoz létezik* $h_1, h_2 : X \to \mathbb{R}$ *az* $\mathcal{M}$ σ*-algebrára nézve mérhető függvény, amelyre*

$$h_1 \leq f \leq h_2 \quad \textit{és} \quad \int_X h_2 - h_1\, d\mu < \varepsilon.$$

Ekkor f *mérhető a teljessé tett* $\bar{\mathcal{M}}$ σ*-algebrára nézve.*

Bizonyítás. Válasszunk minden $n \in \mathbb{N}$ természetes számhoz olyan $\varphi_n \leq f \leq \psi_n$ függvényeket, melyekre φ_n, ψ_n $\mathcal{M}$-mérhetők és $\int_X \psi_n - \varphi_n\, d\mu < \frac{1}{n}$. Legyen

$$h_1 = \limsup_{n\to\infty} \varphi_n, \quad h_2 = \liminf_{n\to\infty} \psi_n.$$

Világos, hogy h_1, h_2 valós értékű, $\mathcal{M}$-mérhető függvények, $h_1 \leq f \leq h_2$, és

$$0 \leq h_2 - h_1 = \liminf_{n\to\infty} \psi_n + \liminf_{n\to\infty} (-\varphi_n) \leq \liminf_{n\to\infty} (\psi_n - \varphi_n).$$

No de a Fatou-lemma szerint (2.2.11)

$$\int_X \liminf_{n\to\infty} (\psi_n - \varphi_n)\, d\mu \leq \liminf_{n\to\infty} \int_X \psi_n - \varphi_n\, d\mu = 0$$

Ebből következik, hogy $0 \leq h_2 - h_1 \leq \liminf_{n\to\infty} (\psi_n - \varphi_n) = 0\ [\mu]$. Ezért $h_2 = h_1\ [\mu]$. Összevetve ezt $h_1 \leq f \leq h_2$ egyenlőtlenségekkel, azt kapjuk, hogy $f = h_1\ [\mu]$, ergo f valóban mérhető $\bar{\mathcal{M}}$-re nézve (2.2.32). □

2.2.34.
Vegyük észre, hogy valamely T tulajdonságra μ-majdnem mindenütt és a $\bar{\mu}$-majdnem mindenütt teljesülés azonos fogalmak. Ha ugyanis $T(x)$ igaz μ-majdnem minden pontban, akkor létezik $A \in \mathcal{M}$, $\mu(A) = 0$, melyre

$$\{x \in X : T(x) \text{ nem igaz}\} \subseteq A.$$

Mivel $\left(X, \bar{\mathcal{M}}, \bar{\mu}\right)$ kiterjesztése az $(X, \mathcal{M}, \mu)$ mértéktérnek, ezért $A \in \bar{\mathcal{M}}$ és $\bar{\mu}(A) = 0$ is teljesül, tehát $T(x)$ igaz $\bar{\mu}$-majdnem minden $x \in X$ mellett is.

A fordított irányú következtetés azon az észrevételen múlik, hogy amennyiben $E \in \bar{\mathcal{M}}$, $\bar{\mu}(E) = 0$, akkor a teljessé tétel konstrukciója (l.(2.1)) szerint létezik $E \subseteq B$, $B \in \mathcal{M}$, $\mu(B) = 0$ halmaz. Így ha $\bar{\mu}$ majdnem minden pontban teljesül T, azaz

$$\{x \in X : T(x) \text{ nem igaz}\} \subseteq E,$$

valamely $E \in \bar{\mathcal{M}}$, $\bar{\mu}(E) = 0$ halmazra, akkor létezik $B \in \mathcal{M}$, $\mu(B) = 0$ halmaz is, melyre

$$\{x \in X : T(x) \text{ nem igaz}\} \subseteq B,$$

azaz $T(x)$ igaz μ-majdnem minden $x \in X$ pontban.

2.3. Mérhető függvény integrálja

Az integrál bevezetésének utolsó lépéseként előjelmegkötés nélküli mérhető függvények integrálját vezetjük be.

Az $(X, \mathcal{M}, \mu)$ mértéktéren legyen $f : X \to \overline{\mathbb{R}}$ mérhető függvény. Világos, hogy f előáll

$$f = f^+ - f^-$$

alakban, ahol f^+ az f pozitív része és f^- az f negatív része. Azt is tudjuk (1.3.15), hogy f^+ és f^- nem negatív mérhető függvények, amelyekre $|f| = f^+ + f^-$.

2.3.1. definíció (integrál, integrálható függvény)**.** Legyen $(X, \mathcal{M}, \mu)$ egy mértéktér. Az $f : X \to \overline{\mathbb{R}}$ mérhető függvény *absztrakt integrálján* az

$$\int_X f\, d\mu = \int_X f^+\, d\mu - \int_X f^-\, d\mu$$

kiterjesztett valós számot értjük feltéve, hogy a fenti összeg értelmes, azaz jobb oldali két integrál közül legalább az egyik véges. Azt mondjuk, hogy az $f : X \to \overline{\mathbb{R}}$ mérhető függvény *integrálható* az $(X, \mathcal{M}, \mu)$ mértéktéren, ha $\int_X f^+\, d\mu$ és $\int_X f^-\, d\mu$ integrálok mindegyike véges valós szám.

Figyeljünk a terminológiára! Ha egy függvény integrálható, az azt jelenti, hogy az integrálja véges. Ez összhangban van a korábbi szóhasználattal és jelöléssel, hiszen előfordulhat, hogy egy függvény integrálja $\pm\infty$. Lehet egy f függvényre $\int_X f\, d\mu$ értelmes anélkül, hogy a függvényt integrálhatónak neveznénk. A továbbiakban az $\int_X f\, d\mu$ kifejezésen azt értjük, hogy a pozitív rész és a negatív rész integráljainak egyike véges és e két $\overline{\mathbb{R}}_+$-beli szám különbsége az integrál értéke. Egyéb esetben a fenti integrál nem is definiált.

Érdemes szem előtt tartanunk a következő észrevételt.

2.3.2.
Az $f : X \to \overline{\mathbb{R}}$ mérhető függvény μ mérték szerinti absztrakt integrálja pontosan akkor definiált, ha létezik $h : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvény, amelyre $\int_X h\, d\mu < \infty$, és az $f \geq -h$ vagy az $f \leq h$ egyenlőtlenségek egyike teljesül. A függvény pontosan akkor integrálható, ha mindkét iménti egyenlőtlenség teljesül.

A mérhető részhalmazon vett integrált ugyanúgy definiáljuk, mint azt nem negatív mérhető függvények esetén tettük.

2.3.3. definíció. Legyen $f : X \to \overline{\mathbb{R}}$ egy mérhető függvény és $E \in \mathcal{M}$ egy mérhető halmaz. Az f függvény E halmazon vett μ mérték szerinti integrálját az

$$\int_E f\, d\mu = \int_X f\chi_E\, d\mu$$

egyenlőség definiálja, feltéve, hogy jobb oldali integrál értelmes.

Világos, hogy ha az $\int_X f\, d\mu$ értelmes, akkor minden $E \in \mathcal{M}$ mérhető halmaz felett $\int_E f\, d\mu$ is értelmes. Hasonlóan, ha az egész halmaz felett egy függvény integrálható, akkor ez minden mérhető részhalmaza felett is igaz marad.

A rendezéssel kapcsolatban kényelmes tulajdonságai vannak a most bevezetett integrálfogalomnak.

2.3.4. állítás (az integrál monoton és homogén)**.** *Legyenek $f, g : X \to \overline{\mathbb{R}}$ olyan mérhető függvények, amelyek integrálja értelmes.*

1. *Ha $f \leq g$ $[\mu]$, akkor*
$$\int_X f\, d\mu \leq \int_X g\, d\mu.$$
2. *Minden $\alpha \in \mathbb{R}$ mellett az αf integrálja is értelmes, továbbá*
$$\alpha \int_X f\, d\mu = \int_X \alpha f\, d\mu.$$
3. *Az integrálok háromszög-egyenlőtlensége:*
$$\left| \int_X f\, d\mu \right| \leq \int_X |f|\, d\mu.$$

Bizonyítás. Az állítás kimondásának sorrendjében és jelöléseivel:
1. Ha $f \leq g$ $[\mu]$, akkor $f^+ \leq g^+$ $[\mu]$ és $g^- \leq f^-$ $[\mu]$. Ebből a nem negatív függvények integráljának monotonitása (2.2.20) szerint
$$\int_X f\, d\mu = \int_X f^+\, d\mu - \int_X f^-\, d\mu \leq \int_X g^+\, d\mu - \int_X g^-\, d\mu = \int_X g\, d\mu.$$

2. A $0 \cdot (\pm\infty) = 0$ konvenciónk szerint az $\alpha = 0$ esetben az állítás nyilvánvaló.

Ha $\alpha > 0$, akkor $(\alpha f)^+ = \alpha f^+$ és $(\alpha f)^- = \alpha f^-$. Így az f és az αf pozitív részeinek integráljai egyszerre valósak a nem negatív függvények integráljának homogenitása szerint (2.2.4). Ugyanígy a negatív részek integráljaira is, emiatt αf integrálja valóban értelmes és

$$\int_X \alpha f \, d\mu = \int_X \alpha f^+ \, d\mu - \int_X \alpha f^- \, d\mu = \alpha \int_X f^+ \, d\mu - \alpha \int_X f^- \, d\mu = \alpha \int_X f \, d\mu.$$

Ha $\alpha < 0$, akkor $(\alpha f)^+ = -\alpha f^-$ valamint $(\alpha f)^- = -\alpha f^+$. Így αf integrálja értelmes és

$$\begin{aligned}\int_X \alpha f \, d\mu = \int_X -\alpha f^- \, d\mu - \int_X -\alpha f^+ \, d\mu = \\ = -\alpha \left(\int_X f^- \, d\mu - \int_X f^+ \, d\mu \right) = \alpha \int_X f \, d\mu.\end{aligned}$$

3. A háromszög-egyenlőtlenség:

$$\begin{aligned}\left| \int_X f \, d\mu \right| = \left| \int_X f^+ \, d\mu - \int_X f^- \, d\mu \right| \leq \\ \leq \int_X f^+ \, d\mu + \int_X f^- \, d\mu = \int_X f^+ + f^- \, d\mu = \int_X |f| \, d\mu\end{aligned}$$

a nem negatív függvényekre igazolt tagonkénti integrálhatóság (2.2.9) szerint. □

2.3.5. állítás (integrál invariáns a null-mértékű halmazon történt változtatásra)**.** *Legyen $f, g : X \to \overline{\mathbb{R}}$ két mérhető függvény, amelyekre $f = g\,[\mu]$. Ekkor*

$$\int_X f \, d\mu = \int_X g \, d\mu,$$

abban az értelemben, hogy ha az egyik integrál értelmes, akkor a másik is az, és az értékek megegyeznek. Ebből következően a két függvény integrálhatósága is ekvivalens.

Bizonyítás. Az f és g mérhetősége miatt f^+ és g^+ is mérhető, ezért $X(f^+ \neq g^+) \in \mathcal{M}$ is mérhető. Világos, hogy $X(f^+ \neq g^+) \subseteq X(f \neq g)$. Mivel null-mértékű halmaz mérhető részhalmaza null-mértékű, ezért $f^+ = g^+\,[\mu]$. Analóg módon kapjuk, hogy $f^- = g^-\,[\mu]$. Mivel nem negatív függvények körében az integrál érzéketlen a null-mértékű halmazon való megváltoztatásra (2.2.20), ezért a két pozitív résznek és a két negatív résznek azonos az integrálja. □

2.3.6. állítás (integrál-mérték szerinti integrál)**.** *Legyen* $(X, \mathcal{M}, \mu)$ *mértéktér és* $g : X \to \overline{\mathbb{R}}_+$ *nem negatív mérhető függvény. Jelölje* φ *a* g *generálta integrálmértéket, azaz* $d\varphi = g\, d\mu$*, tehát minden* $E \in \mathcal{M}$ *mérhető halmazra.*

$$\varphi(E) = \int_E g\, d\mu.$$

Az $f : X \to \overline{\mathbb{R}}$ *mérhető függvény* φ *mérték szerinti integrálja akkor és csak akkor értelmes, ha az* fg *függvény* μ *szerinti integrálja az. Ebben az esetben*

$$\int_X f\, d\varphi = \int_X fg\, d\mu.$$

Bizonyítás. Mivel $g \geq 0$, ezért $(fg)^+ = f^+g$ és $(fg)^- = f^-g$. Láttuk már az állítást abban a speciális esetben, mikor az f integrandus egy nem negatív mérhető függvény (2.2.13). Ezért

$$\int_X f^+\, d\varphi = \int_X f^+g\, d\mu = \int_X (fg)^+\, d\mu.$$

Ugyanígy a negatív részre is

$$\int_X f^-\, d\varphi = \int_X f^-g\, d\mu = \int_X (fg)^-\, d\mu.$$

A bal oldali számok tehát pontosan akkor végesek, ha a jobb oldaliak is azok, ami éppen a kívánt ekvivalenciát igazolja. Az integrál definíciója szerint ekkor

$$\int_X f\, d\varphi = \int_X f^+\, d\varphi - \int_X f^-\, d\varphi = \int_X (fg)^+\, d\mu - \int_X (fg)^-\, d\mu = \int_X fg\, d\mu.$$

□

2.3.1. Integrálható függvények

A szakaszban az integrálható függvények néhány később fontossá váló tulajdonsága mellett elsősorban a strukturális alaptulajdonságokat vizsgáljuk. Ezek elsősorban a vektortér műveletekkel kapcsolatosak, azon belül is leginkább az összeadás művelettel. A nehézséget az jelenti, hogy két integrálható függvénynek még az összege sincs függvényként definiálva. Végül látni fogjuk, hogy az összeadás műveletet a majdnem mindenütt való egyenlőséggel vegyítve hatékony módszert kapunk arra, hogy az integrálható függvényeket mint egy Banach-tér elemeit is tekinthessük.

Ez nem elsősorban a következő szakasz konvergenciatételei miatt fontos, hanem a későbbi Radon–Nikodym-kör vizsgálatánál látjuk, hogy milyen szépen illeszkedik ide és működik ez a strukturális szemlélet.

Az integrál eddigi felépítését a hatékony konvergenciatételek motiválták. A nem negatív mérhető függvények integráljának bevezetésekor is arra törekedtünk, hogy a megfelelő konvergenciatételek, a Beppo Levi-tétel és a Fatou-lemma, minél hamarabb rendelkezésre álljanak. Az összeadás művelettel való kapcsolatot is éppen a konvergenciatételből származtattuk. Ebből a szempontból ez az alfejezet kitérő és megszakítja a felépítés eddigi egyenes vonalát. Az ok a fenti strukturális fogalom alkotás fontosságának hangsúlyozása. Ha az olvasó ragaszkodik a tárgyalás eddigi sorrendjéhez, akkor ezen a ponton szabadsága van abban, hogy mindent kihagyva a 92. lapra, a 2.4. szakaszra ugorjon. Ennek oka, hogy az összeadás művelettel való kompatibilitás, azaz a tagonkénti integrálhatóság, független az előjelmegkötés nélküli mérhető függvények monoton konvergenciatételétől, a Fatou-lemmától, és az új konvergenciatételtől, amit majd Lebesgue-tételnek, vagy dominált konvergenciatételnek mondunk. Ezeknek a tételeknek egyszerűen nincs kapcsolata az integrálható függvények összeadás műveletével. A függvénysorozatok konvergenciatételeinek gyors eléréséhez folytassuk a tárgyalást a 2.4. szakasszal minden addigit kihagyva, majd annak megértése után térjünk a jelen szakaszhoz vissza.

Azok számára, akik nem térnek el a tárgyalás menetétől, a fejezetet záró 2.4. szakasz a konvergenciatételek összefoglalását adja.

2.3.7. állítás. *Egy $f : X \to \overline{\mathbb{R}}$ mérhető függvény pontosan akkor integrálható, ha*

$$\int_X |f| \, d\mu < \infty.$$

Bizonyítás. Az $|f| : X \to \overline{\mathbb{R}}_+$ nem negatív mérhető függvényre

$$\int_X |f| \, d\mu = \int_X f^+ \, d\mu + \int_X f^- \, d\mu$$

a tagonkénti integrálhatóság szerint (2.2.9). Világos, hogy két nem negatív valós szám összege pontosan akkor véges, ha mindkét szám véges. □

2.3.8.

A következő állítás előtt érdemes megjegyezni, hogy ha f egy integrálható függvény, akkor $|f|$, vagy ami ezzel ekvivalens f is majdnem mindenütt véges (2.2.24). Így ha f és g két integrálható függvény, akkor $f(x) + g(x)$ null-mértékű halmaztól eltekintve véges szám, hiszen két null-mértékű halmaz egyesítése null-mértékű marad. Az $f + g$ összeg-függvény, tehát majdnem minden pontban értelmesen definiált. Tudjuk azt is, hogy szóban forgó null-mértékű halmazon elképzelhető, hogy $f(x) + g(x)$ értelmetlen, de itt akármilyen konstansnak is definiáljuk az összeget, az $f + g$ már az egész X alaphalmazon értelmezett mérhető függvény lesz. Felvethető kérdés tehát az összeg integrálhatósága.

2.3.9. állítás (integrál additív és homogén)**.** *Ha f és g integrálható, akkor $f+g$ is integrálható, valamint*

$$\int_X f+g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu.$$

Amennyiben $\alpha \in \mathbb{R}$ valós, akkor f integrálhatóságából αf integrálhatósága is következik, és

$$\int_X \alpha f\,d\mu = \alpha \int_X f\,d\mu.$$

Bizonyítás. Először is, ha f és g integrálható, akkor $f+g$ majdnem mindenütt véges, így majdnem mindenütt értelmes mérhető függvény. A háromszög-egyenlőtlenség és a nem negatív függvények integráljának additivitása (2.2.9) és monotonitása (2.2.4) szerint

$$\int_X |f+g|\,d\mu \le \int_X |f|+|g|\,d\mu = \int_X |f|\,d\mu + \int_X |g|\,d\mu < \infty.$$

Így az $f+g$ függvény integrálhatósága valóban következik az f és a g függvények integrálhatóságából.

Legyen most $h = f+g$. Azon a halmazon, ahol f és g is véges az alábbi egyenlőségben csak valós számok szerepelnek, ezért a kijelölt műveletek értelmesek. No de a szóban forgó halmaz komplementere null-mértékű, tehát

$$h^+ - h^- = f^+ - f^- + g^+ - g^- \; [\mu],$$

vagy ami ezzel ekvivalens

$$h^+ + f^- + g^- = g^+ + h^- + f^+ \; [\mu].$$

Két majdnem mindenütt egyenlő nem negatív függvény integrálja azonos (2.2.20), és a nem negatív függvény integráljának additivitása szerint (2.2.9)

$$\int_X h^+\,d\mu + \int_X f^-\,d\mu + \int_X g^-\,d\mu = \int_X g^+\,d\mu + \int_X h^-\,d\mu + \int_X f^+\,d\mu.$$

A fenti sorban f, g és h integrálhatósága miatt minden integrál véges, ezért visszarendezve azt kapjuk, hogy

$$\begin{aligned}\int_X h\,d\mu &= \int_X h^+\,d\mu - \int_X h^-\,d\mu = \\ &= \int_X f^+\,d\mu - \int_X f^-\,d\mu + \int_X g^+\,d\mu - \int_X g^-\,d\mu = \\ &= \int_X f\,d\mu + \int_X g\,d\mu.\end{aligned}$$

A homogenitási tulajdonságot már láttuk (2.3.4). $\square$

Az világos, hogy ha egy g integrálható függvény felülről korlátos, akkor az $\frac{1}{\mu(E)}\int_E g\,d\mu$ alakú integrálközepek a valós egyenes korlátos részhalmazát alkotják, méghozzá a g korlátjával. A következő állítás szerint ennek a megfordítása is igaz, tehát az összes értelmes integrálközép felső korlátja – ha van ilyen –, a függvénynek is felső korlátja majdnem mindenütt értelemben.

2.3.10. állítás (integrálközepek)**.** *Legyen az $(X, \mathcal{M}, \mu)$ mértéktéren $g : X \to \overline{\mathbb{R}}$ olyan integrálható függvény, amelyre minden $E \in \mathcal{M}$, $0 < \mu(E) < \infty$ mérhető halmaz esetén*

$$\frac{1}{\mu(E)}\left|\int_E g\,d\mu\right| \le \alpha \quad \textit{valamely } \alpha > 0 \textit{ valós szám mellett.}$$

Ekkor $|g| \le \alpha\ [\mu]$ is fennáll, azaz $\mu(X(|g| > \alpha)) = 0$.

Bizonyítás. Az $r \in \mathbb{Q}, |r| > \alpha$ racionális számhoz válasszunk olyan $\varepsilon_r > 0$ számot, amelyre $|r| - 2\varepsilon_r > \alpha$ teljesül. Jelölje $E_r = X(r - \varepsilon_r < g < r + \varepsilon_r)$. Világos, hogy g mérhetősége szerint $E_r \in \mathcal{M}$. Megmutatjuk, hogy $\mu(E_r) = 0$. Mivel $E_r \subseteq X(|g| > \alpha)$, így g integrálhatósága és $\alpha > 0$ szerint (2.2.24)

$$\mu(E_r) \le \mu(X(|g| > \alpha)) < \infty.$$

Ha $\mu(E_r) > 0$ lenne, akkor az E_r halmazon számolt integrálközép abszolút értéke legfeljebb α lehetne, ezért

$$2\varepsilon_r < \left|\frac{1}{\mu(E_r)}\int_{E_r} g\,d\mu - r\right| = \left|\frac{1}{\mu(E_r)}\left(\int_{E_r} g\,d\mu - \int_{E_r} r\,d\mu\right)\right| \le$$
$$\le \frac{1}{\mu(E_r)}\int_{E_r} |g - r|\,d\mu \le \varepsilon_r,$$

ami ellentmondás, így minden ilyen r mellett $\mu(E_r) = 0$. Persze

$$\mathbb{R} \setminus [-\alpha, \alpha] \subseteq \bigcup_{r\in\mathbb{Q}, |r|>\alpha} (r - \varepsilon_r, r + \varepsilon_r),$$

ezért $X(g > \alpha) = X(|g| = \infty) \cup \left(\bigcup_{r\in\mathbb{Q},|r|>\alpha} E_r\right)$, ami megszámlálhatóan sok null-mértékű halmaz egyesítéseként maga is null-mértékű. □

Egyszerű következménye az állításnak, hogy az integrálható függvényeket az összes véges mértékű halmazokon számolt integráljaik már egyértelműen meghatározzák.

2.3.11. állítás. *Legyen $f, g : X \to \overline{\mathbb{R}}$ két integrálható függvény az $(X, \mathcal{M}, \mu)$ mértéktéren, amelyekre minden $E \in \mathcal{M}$, $\mu(E) < \infty$ véges mértékű halmazra*

$$\int_E f\,d\mu = \int_E g\,d\mu.$$

Ekkor $f = g\ [\mu]$.

Bizonyítás. Legyen $h = f - g$ integrálható függvény. Világos, hogy minden mérhető, véges mértékű halmazon számolt integrálközépre

$$\frac{1}{\mu(E)} \int_E h \, d\mu = 0.$$

Ez azt jelenti (2.3.10), hogy minden $n \in \mathbb{N}$ mellett $\mu\left(X\left(|h| > \frac{1}{n}\right)\right) = 0$, amiből, figyelembe véve, hogy $X(f \neq g) \subseteq \cup_{n=1}^{\infty} X\left(|h| > \frac{1}{n}\right)$ kapjuk, hogy $\mu(X(f \neq g)) = 0$. □

Az integrálható függvények mint normált tér

Sokszor nagyon kényelmes az, ha nem teszünk különbséget két integrálható függvény közt feltéve, hogy azok csak null-mértékű halmazban különböznek. Ha például olyan tulajdonságokat nézünk, ami a függvények integráljával kapcsolatos, akkor ez valóban érthető, hiszen egy null-mértékű halmazon történő változás az integrálokat nem érinti (2.3.5). Eddig is zavaró lehetett például, hogy mikor szükségünk van két integrálható függvény összegére, akkor mindig ugyanazon fárasztó köröket kell meggondolni arról, hogy elképzelhető ugyan, hogy a két függvény összege bizonyos pontokban nem értelmezett, de ez csak null-mértékű halmaz pontjaiban lehetséges, majd itt valahogyan átdefiniálva az összeget, olyan függvényt kapok, amely mérhető, integrálható, ráadásul az integrál értéke független az átdefiniálástól.

Megoldásként mondhatnánk, hogy ne engedjünk meg, csak valós függvényeket. Ez a fenti példát meg is oldja, de attól még igaz, hogy nullmértékű halmazra az integrál érzéketlen marad. Nagyon kényelmes lenne például az is, ha egy f nem negatív mérhető függvényre $\int_X f \, d\mu = 0$ kizárólag a konstans zérus f függvényre teljesülne. Ez biztosan nem igaz, amennyiben ragaszkodunk a klasszikus függvényfogalomhoz, mely szerint egy függvényt az értékei határoznak meg, tehát ha két függvény akár csak egyetlen pontban is különbözik, akkor mint függvények különböző objektumok.

Csábító viszont, hogy a többi normált tér axióma milyen könnyedén teljesül az integrálható függvényekre az alábbi normával. Ha mégis valahogy normált teret tudnánk így faragni, akkor az integrálható függvények vektorterén bevezethetnénk az

$$\|f\| = \int_X |f| \, d\mu$$

leképezést mint normát, amely új megvilágításba helyezné az integrálható függvényeket. Például a normakonvergencia ebben a térben implikálná az integrálok konvergenciáját az

$$\left| \int_X f \, d\mu - \int_X f_n \, d\mu \right| \leq \int_X |f - f_n| \, d\mu = \|f - f_n\| \to 0$$

egyenlőtlenség szerint. Egy normált térben, ha az teljes, a Cauchy-tulajdonságot könnyebb ellenőrizni, mint a konvergenciát, hiszen a Cauchy-tulajdonsághoz magára a határértékre nincs is szükségünk. Ilyen módon esélyt láthatunk arra, hogy az integrálható függvényeket mint megfelelő normált tér elemeket tekintve új konvergenciaeredményekhez jussunk. Ez a fő motivációja az L_1 vektortér alábbi bevezetésének. Egyszerűen arról van szó, hogy úgy érezzük, a világ egyszerűbb lenne, ha a nullmértékű halmazokat nem kéne figyelembe venni.

A következő definíció előtt emlékezzünk vissza a vektortér *faktorterének* fogalmára. Ha adott egy V vektortér és annak egy N altere, akkor bevezettünk egy relációt V vektorai közt. A (v_1, v_2) pár relációban van, ha $v_1 - v_2 \in N$ teljesül. Látható, hogy ez ekvivalenciareláció, az ekvivalenciaosztályok halmaza a V/N faktortér alaphalmaza. Értelmeztük a vektortér műveleteket az ekvivalenciaosztályok közt. Ha v_1 az egyik ekvivalenciaosztály egy eleme és v_2 a másiké, akkor a két ekvivalenciaosztály összege az az ekvivalenciaosztály, amely a v_1+v_2 elemet tartalmazza. A skalárral való szorzás definíciója is ezzel analóg. Ellenőriztük, hogy ilyen módon a bevezetett műveletek jól definiáltak – tehát függetlenek az ekvivalenciaosztály egy-egy elemének megválasztásától –, és V/N a bevezetett műveletekkel valóban teljesíti a vektortér axiómákat.

Legyen $v \in V$ egy rögzített vektor. Látható, hogy a v-t tartalmazó ekvivalenciaosztály a $v+N = \{v + u : u \in N\}$ komplexus összeg. Ha a bevezetjük a $[v] = v+N$ jelölést, akkor a faktortér elemei $V/N = \{[v] : v \in V\}$ alakúak, és a V/N vektortér műveletei: $[v_1]+[v_2] = [v_1 + v_2]$, valamint $\alpha\,[v] = [\alpha v]$, ahol a bal oldalon a definiálandó műveletek vannak, és a jobb oldalon az eredeti V vektortér műveletei szerepelnek.

2.3.12. definíció ($L_1(X, \mathcal{M}, \mu)$)**.** Legyen $(X, \mathcal{M}, \mu)$ mértéktér. Tekintsük a valós értékű, integrálható függvények vektorterét

$$\mathcal{L}_1(X, \mathcal{M}, \mu) = \left\{ f : X \to \mathbb{R} | \quad f \text{ mérhető, } \int_X |f| \, d\mu < +\infty \right\}.$$

Legyen

$$\mathcal{L}_0(X, \mathcal{M}, \mu) = \{ f : X \to \mathbb{R} | f \ \text{ mérhető, } f = 0\,[\mu] \}.$$

Látható, hogy $\mathcal{L}_0$ altere $\mathcal{L}_1$-nek. Jelölje L_1 az $\mathcal{L}_1$ térnek ezen $\mathcal{L}_0$ altere szerinti faktorterét, azaz

$$L_1(X, \mathcal{M}, \mu) = \mathcal{L}_1(X, \mathcal{M}, \mu) / \mathcal{L}^0(X, \mathcal{M}, \mu).$$

Az L_1 vektortér elemeit nevezzük integrálható függvényosztálynak, vagy egyszerűbben integrálható függvényeknek.

Az L_1 vektortér elemei tehát nem függvények, hanem ekvivalenciaosztályok! Az $f \in \mathcal{L}_1$ függvényt tartalmazó ekvivalenciaosztály az

$$\begin{aligned}[f] &= f + \mathcal{L}_0 = \\ &= \{f + h|\, h : X \to \mathbb{R},\ \text{mérhető } h = 0\,[\mu]\} = \\ &= \left\{f + h|\, h \in \mathcal{L}_1, \int_X |h|\, d\mu = 0\right\}\end{aligned}$$

komplexus összeg, azaz függvényeknek halmaza. Az L_1 vektortéren a faktortérbeli összeadás és szorzás a műveletek. Tehát $[f], [g] \in L_1$ mellett $[f]+[g] = [f+g]$ és $\alpha \in \mathbb{R}$ esetén $\alpha\,[f] = [\alpha f]$. Tudjuk, hogy mint minden faktortér, ez is teljesíti a vektortér axiómákat.

2.3.13.
Legyen $[f] \in L_1$ és $g \in [f]$. Ez azt jelenti, hogy $f, g \in \mathcal{L}_1$ két olyan integrálható függvény, amelyre $f = g\,[\mu]$. Tudjuk, hogy ekkor $\int_X f\, d\mu = \int_X g\, d\mu$. Ez azt jelenti, hogy az f függvényt tartalmazó ekvivalenciaosztályban valamennyi függvénynek azonos az integrálja. Bevezethetjük tehát egy ekvivalenciaosztály integráljának a fogalmát úgy, hogy annak egy tetszőleges elemét integráljuk, hiszen a fogalom valóban nem a reprezentánst, hanem az ekvivalenciaosztályt jellemzi. Formálisabban, $[f] \in L_1$ mellett definiálja

$$\int_X [f]\, d\mu = \int_X f\, d\mu.$$

Most bevezetünk egy normát is az L_1 vektortéren.

2.3.14. definíció (norma L_1-en)**.** Tetszőleges $[f] \in L_1(X, \mathcal{M}, \mu)$ mellett jelölje

$$\|[f]\|_{L_1(X,\mathcal{M},\mu)} = \int_X |f|\, d\mu.$$

Ha a szövegkörnyezetben világos, hogy melyik mértéktérről van szó, akkor a bevezetett normát $\|\cdot\|_{L_1}$ vagy még egyszerűbben $\|\cdot\|_1$ jelöli.

2.3.15. állítás. *Az* $(L_1(X, \mathcal{M}, \mu), \|\cdot\|_{L_1})$ *tér normált tér.*

Bizonyítás. A normaaxiómákat ellenőrizzük:

1. Ha $[f] \in L_1(X, \mathcal{M}, \mu)$, akkor $\|[f]\|_{L_1} = 0$ pontosan akkor, ha $\int_X |f|\, d\mu = 0$, ami pontosan akkor áll fenn, ha $f = 0\,[\mu]$. No de az L_1 tér mint vektortér zérus eleme éppen a konstans nulla függvény reprezentálta ekvivalenciaosztály, tehát $\|[f]\|_{L_1} = 0$ éppen akkor áll fenn, ha $f \in \mathcal{L}_1$ az L_1 zérus elemét reprezentálja, azaz $[f] = [0]$.
2. Tetszőleges $\alpha \in \mathbb{R}$ mellett $\|\alpha[f]\|_{L_1} = |\alpha|\, \|[f]\|_{L_1}$ nyilvánvaló, az integrál homogenitása miatt.

3. Hasonlóan: $\|[f]+[g]\|_{L_1} = \|[f+g]\|_{L_1} = \int_X |f+g|\,d\mu \le \int_X |f|+|g|\,d\mu = = \int_X |f|\,d\mu + \int_X |g|\,d\mu = \|[f]\|_{L_1} + \|[g]\|_{L_1}$ egyenlőtlenség minden $[f],[g] \in \in L_1$ mellett fennáll az integrál additivitása szerint. □

Most egy részbenrendezést is bevezetünk az L_1 vektortéren.

2.3.16.
Legyen $f, g, f_1, g_1 \in \mathcal{L}_1$ függvények. Tegyük fel, hogy $f \ge g\,[\mu]$, $f_1 = f\,[\mu]$ és $g = g_1\,[\mu]$. Ekkor három nullmértékű halmaz egyesítésének mint nullmértékű halmaznak a komplementerén $f_1 \ge g_1$. Ez azt jelenti, hogy ha $f \ge g\,[\mu]$ fennáll valamely $f, g \in \mathcal{L}_1$ integrálható függvényre, akkor tetszőleges $f_1 \in [f]$ és $g_1 \in [g]$ mellett $f_1 \ge g_1\,[\mu]$ is fennáll.

Éppen azt gondoltuk meg, hogy definiálhatjuk L_1 ekvivalenciaosztályai közti részbenrendezést a reprezentánsok segítségével.

2.3.17. definíció ($\ge$ reláció L_1-en)**.** Tekintsük az $L_1(X, \mathcal{M}, \mu)$ vektorteret. Legyenek $[f],[g] \in L_1$ tetszőleges elemek. Definiáljuk a $\ge$ relációt L_1 elemei közt. Azt mondjuk, hogy

$$[f] \ge [g],$$

ha $f \ge g\,[\mu]$ teljesül. Világos, hogy a bevezetett $\ge$ reláció *részbenrendezés*, azaz reflexív, tranzitív, és antiszimmetrikus. A $\ge$ reláció a vektortér műveletekkel is kompatibilis. Ez azt jelenti, hogy $[h] \in L_1$ elemre $[f] \ge [g]$ esetén $[f]+[h] \ge [g]+[h]$ és ha $\alpha \ge 0$ valós szám és $[f] \ge [0]$, akkor $\alpha[f] \ge [0]$ is érvényben marad.

2.3.18. állítás (integrál pozitív, lineáris funkcionál L_1-en)**.** *Az $L_1(X, \mathcal{M}, \mu)$ vektortéren az integrálás pozitív lineáris funkcionál, azaz $[f],[g] \in L_1$ és $\alpha, \beta \in \mathbb{R}$ mellett*

$$\int_X [\alpha f + \beta g]\; d\mu = \alpha \int_X [f]\,d\mu + \beta \int_X [g]\,d\mu,$$

továbbá

$$[f] \ge [0] \quad \textit{esetén} \quad \int_X [f]\,d\mu \ge 0.$$

Bizonyítás. Mivel az integrált a reprezentánssal definiáltuk, így a linearitás az $\mathcal{L}_1$-beli függvényekre már igazolt állításból nyilvánvaló. Ha $[f] \ge [0]$, akkor $f \ge 0\,[\mu]$, emiatt $f^- = 0\,[\mu]$. Ebből persze $\int_X [f]\,d\mu = \int_X f^+\,d\mu \ge 0$. □

2.3.19.
Tetszőleges $[f],[g] \in L_1$ mellett, ha $[f] \ge [g]$, akkor $[f-g] \ge [0]$, így a pozitivitás és linearitás szerint $\int_X [f]\;d\mu - \int_X [g]\;d\mu = \int_X [f]-[g]\;d\mu = = \int_X [f-g]\;d\mu \ge 0$, ami azt jelenti, hogy az integrál monoton.

Mivel egyetlen valós számnál sem kisebb az abszolút értéke, ezért $[f] \leq \leq |[f]|$ és $[-f] \leq |[f]|$. A monotonitást kihasználva kapjuk, hogy $\int_X [f] \, d\mu \leq \leq \int_X |[f]| \, d\mu$ és $-\int_X [f] \, d\mu \leq \int_X |[f]| \, d\mu$. Ez azt jelenti, hogy az

$$\left| \int_X [f] \, d\mu \right| \leq \int_X |[f]| \, d\mu$$

egyenlőtlenség minden $[f] \in L_1$ mellett fennáll.

Sikerült tehát formalizálni azt az igényünket, hogy null-mértékű halmazban különböző függvényeket hogyan tarthatunk azonosnak.

Az ekvivalenciaosztályokra áttérve valóban fennáll, hogy $\int_X |[f]| \, d\mu = 0$ kizárólag $[f] = [0]$ mellett teljesülhet. Kisvártatva látni fogjuk, hogy a bevezetett $(L_1(X, \mathcal{M}, \mu), \|\cdot\|)$ normált tér még teljes is, ergo Banach-tér, amivel célunkat elérjük.

Az ár, amit a null-mértékű halmazok ignorálásáért fizettünk, jelentős. Eddig függvényekről beszéltünk, ezután, ha L_1 elemeiről beszélünk, akkor ekvivalenciaosztályokról kell beszélnünk. Egy függvény pontonként van értelmezve, az értelmezési tartományon felvett értékei határozzák meg. Amikor $[f] \in L_1$ „függvényre” hivatkozunk, akkor nem egy konkrét függvényt kell elképzelnünk, hanem egy függvényhalmazt, amelyek egymástól null-mértékű halmazon különböznek, és ezek közül csak az egyik f. Mivel mindegyik ilyennek az integrálja azonos, ezért $[f]$ halmaz bármelyik elemét vehetjük az integrál kiszámításához. Nincs viszont arról szó, hogy minden kérdésben azonosak lennének $[f]$ elemei. Például mennyi az értéke $[f](2)$-nek? A kérdés nem értelmes, így nem lehet megválaszolni sem. Amennyiben ugyanis a μ mértékünk olyan, hogy minden egyelemű halmaz mértéke zérus, akkor $[f]$ elemét tetszőlegesen változtatva olyan függvényt kapunk, amely továbbra is az $[f]$ osztályba tartozik. No de a 2 valós szám nem kitüntetett eleme egy $\mathbb{R}$-en értelmezett mérhető függvény értelmezési tartományának, és a fenti gondolat az értelmezési tartomány tetszőleges elemére elmondható. Így azt kapjuk, hogy $[f]$ ekvivalenciaosztályt a reprezentánsainak az értelmezési tartomány egyes pontjaiban felvett értékei nem határozzák meg! Hasonlóan: nincs értelme kiértékelni L_1 egy elemét mint „függvényt” egyetlen pontjában sem, mert a függvényértékek reprezentánsról reprezentánsra változnak, emiatt az L_1 tér $[f]$ eleméről a többi érték változatlan hagyása mellett nem hordoznak információt.[2]

Most kiszélesítjük az eddigi szóhasználatunkat és egyszerűsítjük a kialakított jelöléseket. Ha pusztán a pontosság igénye vezérelné jelöléseinket és szóhasználatunkat, akkor a könyv hátralévő részében is követni kéne az ebben az alfejezetben az L_1 vektortér egy elemére használt $[f]$ jelölést. Miután

[2] Hasonlóan ahhoz, ahogy egy véletlen esemény egyetlen kimenetele sem hordoz önmagában semmilyen információt várható értékével kapcsolatban.

megértettük az L_1 vektortér struktúráját, a gördülékenyebb fogalmazás és az irodalomban szokásos szóhasználathoz való igazodás kedvéért visszatérünk a függvény szó használatához, annak ellenére, hogy tudjuk, nem az eddigi értelemben használt függvényfogalomról van szó. Ennyit a függvény szó használatáról.

Most kiterjesztjük az L_1-beli függvény fogalmat $\overline{\mathbb{R}}$ értékű integrálható függvényekre is. Azért, hogy a fenti faktorizációt megtehessük, vektortérből kellett kiindulnunk, emiatt kellett kizárni az olyan függvényeket, amelyek $\pm\infty$-t mint értéket is felvehetik. Azzal, hogy $\mathcal{L}_1/\mathcal{L}_0$ elemei vektorteret alkotnak, a faktorizáció elérte célját, most $\mathcal{L}_1/\mathcal{L}_0$ ekvivalenciaosztályai elemei közé visszatehetjük az $f : X \to \overline{\mathbb{R}}$ integrálható függvényeket is. Ha tehát f egy ilyen függvény, akkor f majdnem mindenütt véges. Ahol nem véges, ott legyen mondjuk 0. Az így megváltoztatott függvény eleme az $\mathcal{L}_1/\mathcal{L}_0$ egyik és csak egyik ekvivalenciaosztályának, tegyük ebbe az ekvivalenciaosztályba a kiinduló f függvény is. Ez a változtatás nincs hatással az $\mathcal{L}_1/\mathcal{L}_0$ vektortér voltára, és $[f]$-re továbbra is mint az $f : X \to \overline{\mathbb{R}}$ integrálható függvény által reprezentált ekvivalenciaosztályra hivatkozunk.

2.3.20.
Legyen $f : X \to \overline{\mathbb{R}}$ egy integrálható függvény. A továbbiakban használni fogjuk az $f \in L_1$ – olvasd: L_1-beli függvény – terminológiát. Ez azt jelenti, hogy tekintsük egy olyan valós értékű, minden pontban definiált $\bar{f} : X \to \mathbb{R}$ integrálható függvényt, amely μ-majdnem minden pontban megegyezik f-fel, és f helyett gondoljunk az $\bar{f}$ függvény által reprezentált $\left[\bar{f}\right]$ ekvivalenciaosztályra. Jelöljük tehát a továbbiakban f-fel ezt az $[\bar{f}]$ ekvivalenciaosztályt. Időnként szükség van arra, hogy egy konkrét függvényről mint reprezentánsról beszéljünk, de azt mindig külön hangsúlyozzuk. A lényeg, hogy L_1-beli függvény mindig egy ekvivalenciaosztály, a függvény szó magában pedig ugyanazt jelenti, mint eddig.

Az L_1 vektortér műveleteit a következőképpen kell végrehajtani. Tegyük fel, hogy $f, g \in L_1$ mellett az $f+g$ összegről van szó. Ezt magunkban lefordítjuk a faktorizált vektortér $[f]+[g]$ összegére. Ezt kell definiálnunk: keresünk $\bar{f} \in [f]$ és $\bar{g} \in [g]$ függvényeket, amelyek minden pontban végesek. Pontonkénti összeadással definiáljuk az $\bar{f}+\bar{g}$ integrálható függvényt. Az eredmény az ezt tartalmazó ekvivalenciaosztály, tehát $[f]+[g] = \left[\bar{f}+\bar{g}\right]$. Hasonlóan, ha egy L_1-beli függvény α-szorosát, azaz $\alpha\,[f]$-et definiáljuk, akkor először $\bar{f} \in [f]$ minden pontban valós értékű függvényt keresünk, végrehajtjuk a pontonként definiált $\alpha\bar{f}$ szorzatot, majd az $\alpha\bar{f}$ szorzat ekvivalenciaosztálya definiálja a műveletet $\alpha\,[f] = \left[\alpha\bar{f}\right]$.

Csak olyan állításoknak van értelme tehát az $f \in L_1$ függvényekre, amelyek nem függnek null-mértékű halmaztól való megváltoztatástól, így L_1 egy ekvivalenciaosztályának reprezentánsaitól. A továbbiakban erre mindig ügyelnünk kell, és minden kimondott állításnál meg kell gondolnunk ezt.

2.3.2. A dominált konvergenciatétel

Talán a legfontosabb konvergenciatétel következik. A pontonkénti konvergencia természetes feltétele mellé, az új feltétel, hogy a függvénysorozatnak létezzék integrálható majoránsa.

2.3.21. tétel (Lebesgue-féle dominált konvergenciatétel)**.** *Legyen* $(X, \mathcal{M}, \mu)$ *mértéktér, és* $f_n : X \to \overline{\mathbb{R}}, n \in \mathbb{N}$ *mérhető függvények, amelyek* $(f_n)_{n\in\mathbb{N}}$ *sorozata* μ*-majdnem mindenütt pontonként konvergál az* $f : X \to \overline{\mathbb{R}}$ *mérhető függvényhez. Tegyük fel, hogy létezik* $h \in L_1(X, \mathcal{M}, \mu)$ *függvény, melyre* $|f_n| \leq h\ [\mu]$ *minden* $n \in \mathbb{N}$ *esetén. Ekkor*

1. f_n *és* $f \in L_1$ *;*
2. $f_n \to f$ *az* L_1 *normában, azaz* $\|f - f_n\|_{L_1} \to 0$ *;*
3. $\int_X f_n \, d\mu \to \int_X f \, d\mu$ *midőn* $n \to \infty$*.*

Bizonyítás. Világos, hogy $|f| \leq h\ [\mu]$ az L_1-beli h-t reprezentáló akármelyik függvényre. Ezért $\int_X |f| \, d\mu \leq \int_X |h| \, d\mu < \infty$ miatt f egy integrálható függvény, ezért eleme az L_1 egyik ekvivalenciaosztályának ergo $f \in L_1$ valóban fennáll. Ugyanígy kapjuk az $f_n \in L_1$ tartalmazást is.

Persze $|f_n - f| \leq |f_n| + |f| \leq 2h\ [\mu]$, ezért a $g_n = 2h - |f_n - f|$ egy nem negatív tagú majdnem mindenütt konvergens függvénysorozat, melyre

$$\lim_{n\to\infty} g_n = 2h\ [\mu] .$$

Alkalmazzuk a Fatou-lemma majdnem mindenütt variánsát (2.2.23):

$$\int_X 2h \, d\mu \leq \liminf_{n\to\infty} \int_X g_n \, d\mu = \int_X 2h \, d\mu + \liminf_{n\to\infty} \left(\int_X - |f_n - f| \, d\mu \right) =$$
$$= \int_X 2h \, d\mu - \limsup_{n\to\infty} \int_X |f_n - f| \, d\mu.$$

Mivel $\int_X 2h \, d\mu \in \mathbb{R}$ értelmes az egyszerűsítés, miszerint

$$\limsup_{n\to\infty} \int_X |f_n - f| \, d\mu = 0.$$

Ez nyilván azt jelenti, hogy $\|f - f_n\|_{L_1} \to 0$. Ebből

$$\left| \int_X f_n \, d\mu - \int_X f \, d\mu \right| \leq \int_X |f_n - f| \, d\mu = \|f - f_n\|_{L_1} \to 0$$

már könnyen következik. $\square$

2.3.3. A dominált konvergenciatétel közvetlen következményei

2.3.22. tétel (Lebesgue tétele sorokra)**.** *Tegyük fel, hogy $f_n : X \to \overline{\mathbb{R}}$ olyan mérhető függvények sorozata, az $(X, \mathcal{M}, \mu)$ mértéktéren, amelyekre*

$$\sum_{n=1}^{\infty} \int_X |f_n| \, d\mu < \infty.$$

Ekkor a $\sum_{n=1}^{\infty} f_n(x)$ függvénysor μ-majdnem mindenütt abszolút konvergens. Jelölje f a pontonkénti összegfüggvényt. Ekkor $f \in L_1(X, \mathcal{M}, \mu)$ és a $\sum_{n=1}^{\infty} f_n$ függvénysor L_1-ben is konvergál f-hez, azaz

$$\|f - \sum_{k=1}^{n} f_k\|_{L_1} \to 0$$

midőn $n \to \infty$, ezért

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

Bizonyítás. Legyen $\varphi^* = \sum_{n=1}^{\infty} |f_n|$. A monoton konvergenciatétel sorokra vonatkozó alakja (2.2.10) és a tétel feltétele szerint

$$\int_X \varphi^* \, d\mu = \sum_{n=1}^{\infty} \int_X |f_n| \, d\mu < \infty,$$

ami azt jelenti, hogy $\varphi^* \in L_1$, tehát φ^* majdnem mindenütt véges, ezért a $\sum_{n=1}^{\infty} f_n(x)$ függvénysorozat majdnem mindenütt abszolút konvergens. Tudjuk, hogy $\mathbb{R}$ teljessége szerint, egy abszolút konvergens sor konvergens is, ezért majdnem minden $x \in X$ mellett a $\sum_{n=1}^{\infty} f_n(x)$ sor konvergens. Jelölje f az ezek szerint majdnem minden pontban létező összegfüggvényt. Ha s_n a $\sum_n f_n$ sor n-edik részletösszege, akkor definíció szerint $s_n \to f\,[\mu]$ pontonként, és $|s_n| \leq \varphi^*\,[\mu]$. Alkalmazható tehát Lebesgue tétele (2.3.21), így

$$f \in L_1, \;\; \|s_n - f\|_1 \to 0, \;\; \sum_{k=1}^{n} \int_X f_k \, d\mu = \int_X s_n \, d\mu \to \int_X f \, d\mu, \text{ midőn } n \to \infty.$$

□

2.3.23. (L_1 Banach-tér)
A fenti tételt egyszerűbben úgy is fogalmazhatnánk, hogy ha $\sum_{n=1}^{\infty} f_n$ függvénysor L_1-ben abszolút konvergens, azaz $\sum_{n=1}^{\infty} \|f_n\|_1 < \infty$, akkor majdnem mindenütt pontonként abszolút konvergens, és a sor L_1-ben is konvergens. Ez éppen azt jelenti (0.1.7), hogy az L_1 normált tér minden Cauchy-sorozata konvergens is, azaz Banach-tér.

2.3.24. állítás (integrál-mérték)**.** *Legyen* $f \in L_1(X, \mathcal{M}, \mu)$*. Minden* $E \in \mathcal{M}$ *mérhető halmaz mellett definiálja* $\lambda : \mathcal{M} \to \mathbb{R}$ *a*

$$\lambda(E) = \int_E f \, d\mu$$

halmazfüggvényt. Ekkor λ *egy előjeles mérték.*

Bizonyítás. Legyen $E = \cup_n^\infty E_n$ mérhető halmazok diszjunkt egyesítése. Világos, hogy $f\chi_E = \sum_{n=1}^\infty f\chi_{E_n}$. Mivel nem negatív mérhető függvényre már igazoltuk az állítást (2.2.12), így

$$\sum_{n=1}^\infty \int_X |f\chi_{E_n}| \, d\mu = \sum_{n=1}^\infty \int_{E_n} |f| \, d\mu = \int_E |f| \, d\mu \le \int_X |f| \, d\mu < \infty.$$

Alkalmazható tehát az $f\chi_{E_n}$ függvényekre a sorokra vonatkozó Lebesgue-tétel (2.3.22)

$$\lambda(E) = \int_E f \, d\mu = \int_X f\chi_E \, d\mu = \sum_{n=1}^\infty \int_X f\chi_{E_n} \, d\mu = \sum_{n=1}^\infty \int_{E_n} f \, d\mu = \sum_{n=1}^\infty \lambda(E_n).$$

□

2.4. Újra a konvergenciatételek

A fejezet zárásaként néhány gondolat az igazolt konvergenciatételekről. A két legfontosabb tétel a monoton konvergenciatétel (2.2.8) és a dominált konvergenciatétel (2.3.21), valamint ezek sorokra vonatkozó (2.2.10 és 2.3.22) alakja. Meggondoltuk korábban, hogy a Fatou-lemma (2.2.11 és 2.2.23) ekvivalens a monoton konvergenciatétellel. Feltűnő, hogy a mérték monoton folytonossági tétele (1.2.8), amely kulcs szerepet játszott az integrál felépítésében, a monoton konvergenciatétel speciális esete monoton növekedő karakterisztikus függvények sorozatára. Az is azonnal látszik, hogy a monoton konvergenciatétel nem negatív, de fogyó függvények esetére (2.2.25) a dominált konvergenciatétel azonnali következményének is tekinthető, és a mérték monoton folytonosságának fogyó formája a monoton konvergenciatétel fogyó alakjának alkalmazása fogyó karakterisztikus függvények sorozatára.

Mi a monoton konvergenciatétel és a vele ekvivalens Fatou-lemma előjelmegkötés nélküli függvényekre vonatkozó alakja? A dominált konvergenciatételt úgy is fogalmazhatjuk, hogy $f_n : X \to \overline{\mathbb{R}}$ *mérhető függvények pontonként konvergens* $f_n \to f$ *sorozata esetén, ha létezik* $h \ge 0$ *integrálható függvény, melyre*

$$-h \le f_n \le h \qquad (\dagger)$$

fennáll minden $n \in \mathbb{N}$ mellett, akkor az f_n függvények és az f határfüggvény is integrálható, továbbá

$$\int_X f_n \, d\mu \to \int_X f \, d\mu.$$

Mennyiben gyengíthető (†) feltevés az f_n függvények monotonitása mellett?

2.4.1. tétel (Beppo Levi-tétel)**.** *Legyenek $f_n : X \to \overline{\mathbb{R}}$ mérhető függvények, és $f : X \to \overline{\mathbb{R}}$ ezek pontonkénti határértéke.*

1. *Ha $f_n \le f_{n+1}$ monoton növő és létezik $h \ge 0$ nem negatív integrálható függvény, amelyre*
$$-h \le f_1,$$
akkor
$$\int_X f_n \, d\mu \to \int_X f \, d\mu.$$

2. *Ha $f_n \ge f_{n+1}$ monoton fogyó és létezik $h \ge 0$ nem negatív integrálható függvény, amelyre*
$$f_1 \le h,$$
akkor
$$\int_X f_n \, d\mu \to \int_X f \, d\mu.$$

Tehát a monoton növekedés feltétele a (†)-beli felső becslést, míg a monoton csökkenés feltétele az alsó becslést teszi szükségtelenné!

Bizonyítás. Nézzük a monoton növekedés esetét. Világos, hogy $f_n^+ \to f^+$ monoton növekvően, és $f_n^- \to f^-$ monoton csökkenően. Az $f_1 \ge -h$ és h integrálható feltételből $f_1^- \le h$ következik, ami azt jelenti, hogy a negatív részekből álló függvénysorozatnak van integrálható majoránsa, ergo a fogyó monoton konvergenciatétel (2.2.25) szerint

$$\int_X f_n^- \, d\mu \to \int_X f^- \, d\mu,$$

és a fenti sorozat elemei és határértéke is véges számok. A pozitív részekre alkalmazhatjuk a monoton konvergenciatételt (2.2.8), így

$$\int_X f_n^+ \, d\mu \to \int_X f^+ \, d\mu.$$

Azt kaptuk, hogy a pozitív és negatív részek integráljai közül az egyik – a negatív részé – véges, az összeg tehát értelmes. Így

$$\int_X f_n \, d\mu \to \int_X f \, d\mu.$$

A fogyó eset a növő eset következménye a függvények (-1)-szeresére alkalmazva. $\square$

Azzal analóg módon, ahogyan a Fatou-lemmát is származtattuk a monoton konvergenciatételből, kapjuk a Fatou-lemmának az előjelmegkötés nélküli mérhető függvényekre vonatkozó alakját.

2.4.2. lemma (Fatou-lemma)**.** *Legyenek $f_n : X \to \overline{\mathbb{R}}$ mérhető függvények az $(X, \mathcal{M}, \mu)$ mértéktéren.*

1. *Ha létezik $h \geq 0$ nem negatív integrálható függvény, amelyre*

$$-h \leq f_n$$

minden $n \in \mathbb{N}$ mellett teljesül, akkor

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

2. *Ha létezik $h \geq 0$ nem negatív integrálható függvény, amelyre*

$$f_n \leq h$$

minden $n \in \mathbb{N}$ mellett teljesül, akkor

$$\limsup_{n\to\infty} \int_X f_n \, d\mu \leq \int_X \limsup_{n\to\infty} f_n \, d\mu.$$

Így, ha az f_n függvénysorozatnak van integrálható majoránsa, akkor

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu \leq \limsup_{n\to\infty} \int_X f_n \, d\mu \leq \int_X \limsup_{n\to\infty} f_n \, d\mu.$$

Bizonyítás. A $g_n = \inf_{k\geq n} f_k$ függvénysorozatra alkalmazzuk a mérhető függvények monoton konvergenciatételét (2.4.1). Így

$$\int_X \liminf_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_X g_n \, d\mu = \liminf_{n\to\infty} \int_X g_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

A fogyó esetre vonatkozó állítást a függvények ellentettjére alkalmazva kapjuk. $\square$

2.4.3. (Lebesgue)
Így ha az $f_n : X \to \overline{\mathbb{R}}$ mérhető függvények sorozata majdnem mindenütt pontonként konvergens, akkor $\liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n$ $[\mu]$, és ha az f_n sorozatnak van integrálható majoránsa, akkor a Fatou-lemma (2.4.2) következtetése szerint

$$\int_X \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu.$$

No de az $f = \lim_{n\infty} f_n$ határfüggvény is integrálható, ezért f majdnem mindenütt véges (2.2.24), ezért az $f_n \to f$ $[\mu]$ feltétel ekvivalens az $|f - f_n| \to 0$ $[\mu]$ feltétellel. A különbségek abszolút értékeiből álló $|f - f_n| \to 0$ $[\mu]$ sorozatnak is van integrálható majoránsa, tehát a már meggondolt limes-integrál felcserélhetőség szerint

$$\lim_{n\to\infty} \int_X |f - f_n| \, d\mu = \int_X \lim_{n\to\infty} |f - f_n| \, d\mu = \int_X 0 \, d\mu = 0$$

is fennáll, ami a dominált konvergenciatétel (2.3.21) újragondolását adja.

A három legfontosabb konvergenciatételünk a monoton konvergenciatétel, a Fatou-lemma és a dominált konvergenciatétel. Ezek közül az univerzális a Fatou-lemma, amely mindkettőt magában foglalja. A monoton konvergenciatétel és a dominált konvergenciatétel a Fatou-lemmának felhasználások szempontjából éppen kellemes alakja.

3. fejezet

Mérték konstrukció

Az előző fejezetekben láttuk, hogy egy mértéknek mint σ-additív halmazfüggvénynek segítségével hatékony, tehát jól használható konvergenciatételekkel rendelkező, integrál-elméletet építettünk fel. Ha nem is említettük, de a kialakított integrál intuitív fogalmába beletartozott, hogy olyan mértéket képzelünk, amelyre az intervallumok mértéke a hosszukkal egyezik meg. Eddig a pontig nem adtunk konkrét, nem triviális példát mértékre. Könnyen meghatározható triviálisnak mondható mértékek persze vannak. Ilyen például a számossági mérték[1], a Dirac-mérték[2], vagy a megszámlálható–co-megszámlálható[3] mérték. A probléma ezekkel, hogy amikor az integrálra gondolunk, akkor tipikusan nem ilyen triviális eseteket képzelünk mértéknek, hanem sokkal klasszikusabban: hosszúságot, a területet, a térfogatot, még inkább több dimenzióban valamiféle terjedelmet, vagy fizikai analógiát követve tömeget.

Olyan halmazfüggvényre van szükségünk, amely σ-additív és az intervallumok mértéke megegyezik azok hosszával. Ez utóbbi elvárásunkból természetesen adódik az *eltolás-invariancia* feltétele, mely szerint X vektortér esetében az $(X, \mathcal{M}, \mu)$ mértéktér eltolás-invariáns, ha minden $E \in \mathcal{M}$ mérhető halmazra és minden $x \in X$ vektorra az $x + E \in \mathcal{M}$ eltolt is mérhető és $\mu(E) = \mu(x + E)$. Jogos kérdés, hogy miért kell egyáltalán σ-algebra ehhez.

[1] A hatványhalmaz a szigma-algebra, és egy halmaz mértéke az elemeinek száma, vagy ∞ attól függően, hogy a halmaz véges számosságú vagy sem.

[2] A szigma-algebra újra a teljes hatványhalmaz, és egy rögzített x_0 pont mellet egy halmaz mértéke 1 vagy 0 attól függően, hogy x_0 hozzátartozik a halmazhoz vagy sem.

[3] A szigma-algebra azon halmazokat tartalmazza, amelyek vagy a komplementerük legfeljebb megszámlálható számosságú. Egy halmaz mértéke azonos a számossági mértékkel.

Miért nem azt várjuk el, hogy a hatványhalmaz legyen a mérték értelmezési tartománya?

A probléma megértéséhez nézzük a következő példát. A $[0,1)$ intervallum két eleme akkor legyen az R relációban, ha a különbségük racionális szám, azaz

$$R = \{(x,y) \in [0,1) \times [0,1) : x - y \in \mathbb{Q}\}.$$

Világos, hogy R *ekvivalenciareláció*, tehát

1) reflexív, ugyanis $x - x = 0 \in \mathbb{Q}$;
2) szimmetrikus, ugyanis ha $x - y \in \mathbb{Q}$, akkor $y - x \in \mathbb{Q}$;
3) tranzitív, ugyanis ha $x - y \in \mathbb{Q}$ és $y - z \in \mathbb{Q}$, akkor

$$x - z = (x - y) + (y - z) \in \mathbb{Q}.$$

Válasszunk ki minden egyes ekvivalenciaosztályból egyetlen elemet – kiválasztási axióma – és jelölje E az így kapott halmazt. Persze $E \subseteq [0,1)$. Cseréljük fel ennek a halmaznak a $[0,1-r)$ és az $[1-r,1)$ intervallumokba eső részeit, azaz tetszőleges $r \in [0,1)$ mellett jelölje

$$\begin{aligned}
E_r^1 &= (E + r) \cap [r,1) = E \cap [0,1-r) + r;\\
E_r^2 &= ((E + r) \cap [1,1+r)) - 1 = E \cap [1-r,1) + r - 1;\\
E_r &= E_r^1 \cup E_r^2.
\end{aligned}$$

Az E_r halmaz az E_r^1 és E_r^2 halmazok diszjunkt egyesítése ugyanis, $E_r^1 \subseteq [r,1)$ és $E_r^2 \subseteq [0,r)$. Ha $m : \mathcal{P}(\mathbb{R}) \to \overline{\mathbb{R}}_+$ a valós egyenes hatványhalmazán értelmezett nem negatív, végesen additív, eltolás-invariáns halmazfüggvény, akkor

$$\begin{aligned}
m(E_r) &= m\left(E_r^1\right) + m\left(E_r^2\right) =\\
&= m(E \cap [0,1-r) + r) + m(E \cap [1-r,1) + r - 1) =\\
&= m(E \cap [0,1-r)) + m(E \cap [1-r,1)) =\\
&= m(E),
\end{aligned}$$

tehát az összes E_r halmaznak azonos a mértéke, mégpedig $m(E_r) = m(E_0) = m(E)$.

Most azt gondoljuk meg, hogy a

$$[0,1) = \bigcup_{r \in [0,1) \cap \mathbb{Q}} E_r$$

diszjunkt egyesítés is teljesül. Világos, hogy $E_r \subseteq [r,1) \cup [0,r) \subseteq [0,1)$ az összes szóba jövő $r \in [0,1)$ mellett teljesül, ezért a jobb oldal része a bal oldalnak. Ha $y \in [0,1)$ akkor y eleme valamely R szerinti ekvivalenciaosztálynak, ezért van egyetlen $e \in E$, amelyre $y-e \in \mathbb{Q}$. Ha $e \leq y$, akkor az $r = y-e \in [0,1) \cap \mathbb{Q}$ jelöléssel

$$y = e + (y-e) = e + r \in (E+r) \cap [r,1) = E_r^1 \subseteq E_r.$$

Ha pedig $e > y$, akkor $r = 1-(e-y) \in [0,1) \cap \mathbb{Q}$ jelöléssel

$$y = e + 1 - (e-y) - 1 = e + r - 1 \in ((E+r) \cap [1,r)) - 1 = E_r^2 \subseteq E_r.$$

Ha $x \in E_r \cap E_s$, akkor $x = e_1 + r$ vagy $x = e_1 + r - 1$ valamely $e_1 \in E$; és $x = e_2 + s$ vagy $x = e_2 + s - 1$ valamely $e_2 \in E$ elem választása mellett. Az összes esetben $e_1 - e_2 \in \mathbb{Q}$ lenne, ami az E halmaz konstrukciója szerint csak úgy lenne lehetséges, ha $e_1 = e_2$. Ez csak az $r = s$ esetben fordulhat elő. Megmutattuk tehát, hogy $r \neq s$, $r, s \in [0,1) \cap \mathbb{Q}$ mellett $E_r \cap E_s = \varnothing$.

Ha tehát m az egész hatványhalmazon értelmezett σ-additív halmazfüggvény, melyre $m([0,1)) = 1$, akkor

$$1 = m([0,1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} m(E_r) = \begin{cases} +\infty, & \text{ha } m(E) > 0, \\ 0, & \text{ha } m(E) = 0, \end{cases}$$

ami ellentmondás.

Mivel az itt ismertetett példa több dimenzióra is átírható, azt gondoltuk meg, hogy *nem létezik az n-dimenziós tér teljes hatványhalmazán értelmezett* $m : \mathcal{P}(\mathbb{R}^n) \to \overline{\mathbb{R}}_+$ *halmazfüggvény, amely a következő tulajdonságokkal rendelkezne:*

1. $A_1, A_2, \ldots, A_k, \ldots$ *véges vagy megszámlálható végtelen diszjunkt halmazsorozat esetén*

   $$m(A_1 \cup A_2 \cup \ldots \cup A_k \cup \ldots) = m(A_1) + m(A_2) + \cdots + m(A_k) + \ldots;$$

2. *Ha* $E, F \subseteq \mathbb{R}^n$ *egybevágó (eltolással vagy forgatással vagy tükrözéssel egymásba vihető) halmazok, akkor*

   $$m(E) = m(F);$$

3. *Az egységkocka mértéke* 1, *azaz*

   $$m(\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : 0 \leq x_i < 1, \forall i = 1, \ldots, n\}) = 1.$$

Ez azt jelenti, hogy esélyünk sincs arra, amire első pillanatban vágytunk: a valós egyenes hosszúság függvényét nem lehet kiterjeszteni a σ-additivitás és az eltolás-invariancia megtartásával a hatványhalmazra.

Kétségtelen tény, hogy a fenti paradoxnak is érezhető példában a probléma magja a σ-additivitás megkövetelése. Jogosan merül fel ezért, hogy miért nem helyettesítjük a mérték definíciójában a σ-additivitást az annál sokkal természetesebb végesen additivitással? Két súlyos érv szól ez ellen:

1. Láttuk, hogy a konvergenciatételeinknél a σ-additivitás elkerülhetetlenül fontos volt. Hatékony konvergenciatételek nélkül a felépített integrál elmélet hasznossága megkérdőjelezhető.

2. Még a végesen additivitás feltétele mellett sem lehet a teljes hatványhalmazra kiterjeszteni az intervallumokon értelmezett hosszúság függvényt $\mathbb{R}^n$-en, az $n \geq 3$ esetben. Ugyanis Banach és Tarski 1924-ben megmutatták, hogy a következő állítás teljesül, amelyet az irodalom Banach–Tarski-paradoxonnak nevez: *Legyen U és V két korlátos nyílt halmaza $\mathbb{R}^n$-nek, ahol $n \geq 3$. Ekkor létezik $k \in \mathbb{N}$ szám, léteznek $U_1, \ldots, U_k$ diszjunkt halmazok és léteznek $V_1, \ldots, V_k$ diszjunkt halmazok, amelyekre*

$$U = \cup_{j=1}^{k} U_j; \quad V = \cup_{j=1}^{k} V_j,$$

továbbá minden $j = 1, \ldots, k$ esetén U_j egybevágó V_j-vel.

A tanulság tehát az, hogy nem érdemes a megszámlálható additivitást végesen additivitásra gyengíteni, ettől még nem lesz kiterjeszthető a teljes hatványhalmazra a mérték, és még az így felépülő integrál-elméletünk is kevésbé használható konvergenciatételeket fog eredményezni.

Úgy oldhatjuk fel tehát a paradox tulajdonságot, hogy nem várjuk el az intervallumokon értelmezett hosszúság függvénytől, hogy az a teljes hatványhalmazon legyen értelmezve. Persze jó lenne ha minél több halmaznak tudnánk mértéket tulajdonítani. A cél tehát az, hogy terjesszük ki az intervallumokon értelmezett hosszúságfüggvényt mint megszámlálhatóan additív halmazfüggvényt az intervallumokat tartalmazó minél bővebb σ-algebrára. A Banach–Tarski-paradoxon ekkor persze nem paradoxon, hiszen összesen csak annyi történik, hogy az $U_1, \ldots, U_k$ és $V_1, \ldots, V_k$ halmazok nem esnek bele ebbe a σ-algebrába, tehát nem mérhetők, az U mértékének nem is kell megegyeznie a V mértékével.

3.1. A Caratheodory-féle kiterjesztési eljárás

Az egész fejezetben legyen $\mathcal{H}$ olyan X-beli halmazrendszer, amelyre $\varnothing \in \mathcal{H}$.

A külső mérték fogalmának megértéséhez nézzük a σ-additivitás következő karakterizációját.

3.1.1. állítás. *Legyen $\mu : \mathcal{P} \to \overline{\mathbb{R}}_+$ félgyűrűn értelmezett halmazfüggvény. Ekkor a*

i. *μ nem a konstans $+\infty$ halmazfüggvény;*

ii. *μ σ-additív $\mathcal{P}$ félgyűrűn*

feltételek együtt ekvivalensek az

a. $\mu(\emptyset) = 0$;

b. $A, B \in \mathcal{P}, A \subseteq B$ *esetén* $\mu(A) \leq \mu(B)$;

c. μ *σ-szubadditív, azaz* $A, A_n \in \mathcal{P}$, $A = \cup_{n=1}^{\infty} A_n$ *esetén* $\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n)$;

d. μ*-végesen additív, azaz* $A_1, A_2 \in \mathcal{P}$ *olyan diszjunkt halmazok esetén, melyre* $A_1 \cup A_2 \in \mathcal{P}$ *is fennáll teljesül a* $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$ *egyenlőség*

feltételekkel.

Bizonyítás. A generált gyűrűre terjesszük ki μ-t (1.2.9). Láttuk, hogy egy gyűrűn értelmezett halmazfüggvény, amely az i. és a ii. feltételeket kielégíti, teljesíti a második feltételcsoport előírásait még a generált gyűrűn is (1.2.7), ezért a szűkebb $\mathcal{P}$ félgyűrűn is.

Legyen $A, A_n \in \mathcal{P}$, melyekre $A = \cup_{n=1}^{\infty} A_n$ diszjunkt előállítás. Jelölje $\hat{\mu}$ az $\mathrm{r}(\mathcal{P})$ generált gyűrűre való végesen additív kiterjesztést (1.2.10). Ekkor

$$\mu(A) = \hat{\mu}(A) \geq \hat{\mu}\left(\cup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \hat{\mu}(A_n) = \sum_{n=1}^{N} \mu(A_n)$$

Ez minden N mellett fennáll, ami a σ-szubadditivitással együtt a σ-additivitást jelenti. □

Persze a második feltételcsoport 4 állítása nem független egymástól. A legjobb, ha a monotonitásra úgy tekintünk, mint a végesen additivitás egy gyengítésére. Ez a gyengítés vezet a külső mérték fogalmához.

3.1.2. definíció (külső mérték)**.** Egy $\mu : \mathcal{H} \to \overline{\mathbb{R}}_+$ függvényt *külső mértéknek* nevezünk a $\mathcal{H}$ halmazrendszeren, ha

1. $\mu(\emptyset) = 0$,
2. monoton, azaz $H_1, H_2 \in \mathcal{H}$ és $H_1 \subseteq H_2 \Rightarrow \mu(H_1) \leq \mu(H_2)$,
3. σ-szubadditív, azaz $H = \cup_{n=1}^{\infty} H_n,\ H, H_n \in \mathcal{H}\ \forall n \in \mathbb{N} \Rightarrow \mu(H) \leq$ $\leq \sum_{n=1}^{\infty} \mu(H_n)$.

A fejezetben nagyon fontos a σ-additív halmazfüggvények értelmezési tartománya. A kényelmes szóhasználat okán továbbra is *mértéknek* nevezünk egy nem konstans σ-additív halmazfüggvényt, még akkor is ha nem egy σ-algebra annak értelmezési tartománya. Ebben az esetben az értelmezési tartományt mindig külön megjelöljük.

A külső mérték tehát a mértéknél gyengébb fogalom. A következő egy nagyon fontos példa külső mértékre. Még a definíció előtt emlékeztetünk arra, hogy az üreshalmaz infumumát $+\infty$-nek definiáltuk.

3.1.3. állítás (generált külső mérték)**.** *Legyen $\mathcal{H}$ egy tetszőleges X-beli halmazrendszer, amelyre $\varnothing \in \mathcal{H}$, $\mu : \mathcal{H} \to \overline{\mathbb{R}}_+$ halmazfüggvény, amelyre $\mu(\varnothing) = 0$. Definiálja tetszőleges $M \subseteq X$ mellett*

$$\mu^*(M) = \inf\left\{\sum_{n=1}^{\infty} \mu(H_n) : M \subseteq \cup_{n=1}^{\infty} H_n, H_n \in \mathcal{H}, n \in \mathcal{N}\right\}.$$

Ekkor $\mu^* : \mathcal{P}(X) \to \overline{\mathbb{R}}_+$ az egész hatványhalmazon értelmezett külső mérték, amelyre minden $H \in \mathcal{H}$ esetén $\mu^*(H) \leq \mu(H)$.

Ha $\nu : \mathcal{P}(X) \to \overline{\mathbb{R}}_+$ egy másik olyan külső mérték, amelyre $\nu(H) \leq \mu(H)$ fennáll minden $H \in \mathcal{H}$ mellett, akkor $\nu \leq \mu^*$, azaz μ^* a fenti típusú külső mértékek közül a maximális.

Bizonyítás. A $\mu^*(H) \leq \mu(H), H \in \mathcal{H}$ egyenlőtlenség következik abból, hogy $\{H\}$ is egy $\mathcal{H}$-beli legfeljebb megszámlálható lefedése H-nak, amiből persze $\mu^*(\varnothing) \leq \mu(\varnothing) = 0$ is látszik. A monotonitás is nyilvánvaló, hiszen bővebb halmaz infimuma kisebb.

A σ-szubadditivitáshoz tegyük fel, hogy $M = \cup_{n=1}^{\infty} M_n$.

Amennyiben $\sum_{n=1}^{\infty} \mu^*(M_n) = +\infty$, úgy készen is vagyunk. Ellenkező esetben rögzítsünk valamely $\varepsilon > 0$ számot. Persze $\mu^*(M_n) < +\infty$, ezért az $\varepsilon/2^n$ számhoz létezik $H_i^{(n)} \in \mathcal{H}$ $i \in \mathbb{N}$, $M_n \subseteq \cup_{i=1}^{\infty} H_i^{(n)}$ lefedés, amelyre $\mu^*(M_n) + \varepsilon/2^n > \sum_{i=1}^{\infty} \mu\left(H_i^{(n)}\right)$. Ekkor

$$\begin{aligned} &M = \cup_{n=1}^{\infty} M_n \subseteq \cup_{n=1}^{\infty} \cup_{i=1}^{\infty} H_i^{(n)}, \text{ ezért} \\ &\mu^*(M) \leq \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \mu\left(H_i^{(n)}\right) < \sum_{n=1}^{\infty} \left(\mu^*(M_n) + \varepsilon/2^n\right) = \sum_{n=1}^{\infty} \mu^*(M_n) + \varepsilon. \end{aligned}$$

Mivel ez minden $\varepsilon > 0$ mellett is fennáll, a fenti egyenlőtlenség a μ^* halmazfüggvény σ-szubadditivitását jelenti. Igazoltuk tehát, hogy μ^* valóban az egész hatványhalmazon értelmezett külső mérték.

Most legyen ν egy másik külső mérték, amelyre $H \in \mathcal{H}$ esetén $\nu(H) \leq \mu(H)$. Ha egy $M \subseteq X$ halmazra $\mu^*(M) = \infty$, akkor $\nu(M) \leq \mu^*(M)$ nyilvánvalóan teljesül. Egyébként létezik $M \subseteq \cup_n^{\infty} H_n$, $H_n \in \mathcal{H}$ befedés. A ν külső mérték, ezért

$$\nu(M) \leq \nu\left(\cup_{n=1}^{\infty} H_n\right) \leq \sum_{n=1}^{\infty} \nu(H_n) \leq \sum_{n=1}^{\infty} \mu(H_n).$$

Ez minden $\mathcal{H}$-beli lefedésre, ezért ezek infimumára is fennáll, azaz $\nu(M) \leq \mu^*(M)$. ☐

3.1.4. állítás (1. lépés)**.** *Legyen $\mathcal{H}$ metszet zárt és $\varnothing \in \mathcal{H}$. Ha $\mu : \mathcal{H} \to \overline{\mathbb{R}}_+$ külső mérték, akkor a fent definiált μ^* kiterjesztése is μ-nek, azaz minden $H \in \mathcal{H}$ mellett $\mu(H) = \mu^*(H)$.*

Bizonyítás. Az előző állítás szerint (3.1.3) csak a $\mu(H) \le \mu^*(H)$ egyenlőtlenség szorul indoklásra. Amennyiben $\mu^*(H) = +\infty$, úgy készen is vagyunk.

Ellenkező esetben van megszámlálható $\mathcal{H}$-beli lefedését H-nak, azaz $H \subseteq \cup_{n=1}^{\infty} H_n$, $H_n \in \mathcal{H}$. De $H = \cup_{n=1}^{\infty}(H \cap H_n)$ és $H \cap H_n \in \mathcal{H}$. Ezért a külső mérték σ-szubadditivitása és monotonitása szerint

$$\mu(H) \le \sum_{n=1}^{\infty} \mu(H \cap H_n) \le \sum_{n=1}^{\infty} \mu(H_n).$$

De ez minden befedésre, így ezek infumumára is fennáll, azaz $\mu(H) \le \mu^*(H)$. □

3.1.5. definíció (Caratheodory-értelemben mérhető halmaz)**.** Legyen $\mu : \mathcal{P}(X) \to \overline{\mathbb{R}}$ olyan halmazfüggvény, melynek értékkészlete a $+\infty$ és $-\infty$ közül csak az egyiket tartalmazza, azaz $\mu(A) + \mu(B)$ minden A, B halmazra értelmezett. Definiálja

$$\mathcal{A}_\mu = \{A \subseteq X : \forall M \subseteq X\text{-re } \mu(M) = \mu(M \cap A) + \mu(M \cap A^c)\}.$$

Az $\mathcal{A}_\mu$ halmazrendszer elemeit a μ-re nézve *Caratheodory-értelemben mérhető* halmazoknak nevezzük.

3.1.6.

Érdemes úgy tekintenünk a Caratheodory-értelemben mérhető halmazok rendszerére, mint arra a halmazrendszerre, amelyen az egész hatványhalmazon értelmezett halmazfüggvény legalább végesen additív.

Világos ugyanis, hogy ha A_1 és A_2 diszjunkt halmazok, amelyek közül legalább az egyik – mondjuk A_1 – Caratheodory-értelemben mérhető, akkor

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2),$$

hiszen $M = A_1 \cup A_2$ jelöléssel

$$\mu(M) = \mu(M \cap A_1) + \mu(M \cap A_1^c) = \mu(A_1) + \mu(A_2).$$

3.1.7. állítás (2. lépés)**.** *Legyen $\mu : \mathcal{P}(X) \to \overline{\mathbb{R}}_+$, melyre $\mu(\varnothing) = 0$. Ekkor az $\mathcal{A}_\mu$ halmazrendszer egy algebra; és a μ halmazfüggvénynek az $\mathcal{A}_\mu$-re való megszorítása egy végesen additív halmazfüggvény.*

Definiálja tetszőlegesen rögzített $M \subseteq X$ mellett $\mu_M(A) = \mu(M \cap A)$ az egész hatványhalmazon értelmezett μ_M halmazfüggvényt. Ekkor μ_M-nek $\mathcal{A}_\mu$-re való megszorítása is végesen additív.

Bizonyítás. Először azt mutatjuk meg, hogy az $\mathcal{A}_\mu$ halmazrendszerre: 1) $X \in \in \mathcal{A}_\mu$; 2) $A \in \mathcal{A}_\mu \Rightarrow A^c \in \mathcal{A}_\mu$; 3) $A_1, A_2 \in \mathcal{A}_\mu \Rightarrow A_1 \cap A_2 \in \mathcal{A}_\mu$. Ez első két állítás az $\mathcal{A}_\mu$ halmazrendszer definíciójának azonnali következménye.

A metszet-zártsághoz azt kell megmutatnunk, hogy tetszőleges $M \subseteq X$ halmazra

$$\mu(M) = \mu(M \cap A_1 \cap A_2) + \mu(M \cap (A_1 \cap A_2)^c). \tag{$\dagger$}$$

Induljunk ki a jobb oldali második halmazból. Az A_1 Caratheodory-mérhetőségét erre a halmazra felírva

$$\mu(M \cap (A_1 \cap A_2)^c) = \mu(M \cap (A_1 \cap A_2)^c \cap A_1) + \mu(M \cap (A_1 \cap A_2)^c \cap A_1^c).$$

A jobb oldalon megjelenő halmazokra

$$\begin{aligned} &M \cap (A_1 \cap A_2)^c \cap A_1 = (M \cap A_1) \cap (A_1^c \cup A_2^c) = M \cap A_1 \cap A_2^c, \\ &M \cap (A_1 \cap A_2)^c \cap A_1^c = M \cap ((A_1 \cap A_2) \cup A_1)^c = M \cap A_1^c. \end{aligned}$$

Így az igazolandó ($\dagger$) jobb oldalán végül is a

$$\mu(M \cap A_1 \cap A_2) + \mu(M \cap A_1 \cap A_2^c) + \mu(M \cap A_1^c)$$

összeg szerepel. De vegyük észre, hogy A_2 mérhetőségét az $M \cap A_1$ halmazra felírva az első két tag összege éppen $\mu(M \cap A_1)$, tehát ($\dagger$) jobb oldal már

$$\mu(M \cap A_1) + \mu(M \cap A_1^c) = \mu(M)$$

alakú az A_1 halmaz Caratheodory-mérhetősége szerint, ami éppen a kívánt ($\dagger$).

Ezt kellett tehát belátni ahhoz, hogy igazoljuk az $\mathcal{A}_\mu$ halmazrendszer algebra mivoltát.

Legyen most $M \subseteq X$ rögzítve, továbbá $A \in \mathcal{A}_\mu$ és $B \subseteq X$ diszjunkt halmazok. Ekkor alkalmazva az A halmaz Caratheodory-mérhetőségének definícióját azt kapjuk, hogy

$$\begin{aligned} &\mu_M(A \cup B) = \mu(M \cap (A \cup B)) = \\ &\qquad \mu((M \cap (A \cup B)) \cap A) + \mu((M \cap (A \cup B)) \cap A^c) = \\ &= \mu(M \cap A \cap (A \cup B)) + \mu(M \cap A^c \cap (A \cup B)) = \mu(M \cap A) + \mu(M \cap B) = \\ &\qquad = \mu_M(A) + \mu_M(B). \end{aligned}$$

Ebből következik a μ_M halmazfüggvény végesen additivitása az $\mathcal{A}_\mu$ halmazalgebra fölött. $\square$

3.1.8. állítás (3. lépés)**.** *Legyen* $\mu : \mathcal{P}(X) \to \overline{\mathbb{R}}_+$ *külső mérték. Ekkor az* $\mathcal{A}_\mu$ *halmazrendszer egy* σ*-algebra;* $\mu(A) = 0$ *esetén az* $A \in \mathcal{A}_\mu$ *is fennáll;* μ*-nek az* $\mathcal{A}_\mu$*-re való* $\tilde{\mu}$ *megszorítása egy* σ*-additív halmazfüggvény; és az* $(X, \mathcal{A}_\mu, \tilde{\mu})$ *mértéktér teljes.*

Bizonyítás. Láttuk, hogy $\mathcal{A}_\mu$ algebra elegendő tehát belátni, hogy $\mathcal{A}_\mu$ zárt a diszjunkt megszámlálható egyesítésre. Legyenek $A_n \in \mathcal{A}_\mu$ egymástól diszjunkt halmazok és legyen $A = \cup A_n$ ezek egyesítése. Legyen $B_N = \cup_{n=1}^N A_n$ $\forall N \in \mathbb{N}$ mellett. Mivel $\mathcal{A}_\mu$ algebra, ezért $B_N \in \mathcal{A}_\mu$. Triviális, hogy $B_N \subseteq A$. A B_N halmaz Caratheodory-mérhetősége miatt

$$\begin{aligned}\mu(M) = \mu(M \cap B_N) + \mu(M \cap B_N^c) = \mu_M(B_N) + \mu(M \cap B_N^c) = \\ = \sum_{n=1}^N \mu_M(A_n) + \mu(M \cap B_N^c) = \sum_{n=1}^N \mu(M \cap A_n) + \mu(M \cap B_N^c) \geq \\ \geq \sum_{n=1}^N \mu(M \cap A_n) + \mu(M \cap A^c).\end{aligned}$$

Egyrészt ez minden $N \in \mathbb{N}$ esetén fennáll, másrészt az

$$M = \left(\cup_{n=1}^\infty (M \cap A_n)\right) \cup (M \cap A^c)$$

egyenlőség szerint

$$\mu(M) \geq \sum_{n=1}^\infty \mu(M \cap A_n) + \mu(M \cap A^c) \geq \mu(M \cap A) + \mu(M \cap A^c) \geq \mu(M).$$

Persze ekkor fent mindenütt egyenlőség van, így $A \in \mathcal{A}_\mu$, ergo $\mathcal{A}_\mu$ valóban σ-algebra.

Az iménti egyenlőség tetszőleges $M \subseteq X$ halmazra teljesül, így speciálisan $M = A$-ra is. Ekkor

$$\begin{aligned}\mu(\cup_{k=1}^\infty A_k) = \mu(M) = \sum_{n=1}^\infty \mu(M \cap A_n) + \mu(M \cap A^c) = \\ = \sum_{n=1}^\infty \mu(A_n) + \mu(\varnothing) = \sum_{n=1}^\infty \mu(A_n),\end{aligned}$$

amivel igazoltuk μ σ-additivitását az $\mathcal{A}_\mu$ σ-algebrán.

Ha valamely $A \subseteq X$ mellett $\mu(A) = 0$, akkor minden $M \in \mathcal{P}(X)$ esetén $\mu(M) \leq \mu(M \cap A) + \mu(M \cap A^c) \leq \mu(M)$, tehát $\mu(M) = \mu(M \cap A) + \mu(M \cap A^c)$ is teljesül. Legyen most $B \subseteq A \in \mathcal{A}_\mu$ és $\tilde{\mu}(A) = 0$. Ekkor a külső mérték monotonitása szerint $\mu(B) = 0$ is fennáll, tehát a már bizonyítottak szerint $B \in \mathcal{A}_\mu$ is teljesül. Ez éppen azt jelenti, hogy $\tilde{\mu}$ teljes mérték az $\mathcal{A}_\mu$ σ-algebrán. □

3.1.9. tétel (4. lépés, kiterjesztési tétel)**.** *Legyen $\mathcal{P}$ egy félgyűrű és $\mu : \mathcal{P} \to \overline{\mathbb{R}}_+$ külső mérték. Jelölje μ^* a μ generálta külső mértéket, amely az egész hatványhalmazon értelmezett kiterjesztése μ-nek. Jelölje $\mathcal{A}_{\mu^*}$ a Caratheodory-értelemben μ^* mérhető halmazok σ-algebráját. Ekkor a $\mathcal{P} \subseteq \mathcal{A}_{\mu^*}$ tartalmazás teljesülése ekvivalens a μ halmazfüggvény végesen additivitásával.*

Bizonyítás. Tegyük fel, hogy $\mathcal{P} \subseteq \mathcal{A}_{\mu^*}$. Tudjuk, hogy $\mathcal{A}_{\mu^*}$ halmazrendszeren μ^* végesen additív (3.1.7), és mivel μ külső mérték $\mathcal{P}$ félgyűrűn, úgy $\mu = \mu^*|\mathcal{P}$ (3.1.4), tehát μ is végesen additív.

Most tegyük fel, hogy μ végesen additív külső mérték. Legyen $A \in \mathcal{P}$. Megmutatjuk, hogy A Caratheodory-mérhető (3.1.5). Világos, hogy tetszőleges M mellett $\mu^*(M) \leq \mu^*(M \cap A) + \mu^*(M \cap A^c)$ hiszen μ^* szubadditív. Ha $\mu^*(M) = +\infty$ lenne, akkor készen is volnánk a fordított irányú egyenlőtlenséggel. Ha $\mu^*(M) < \infty$, akkor azt kell megmutatnunk a generált külső mérték definíciója szerint (3.1.3), hogy tetszőlegesen rögzített $M \subseteq \cup_{n=1}^{\infty} A_n$, $A_n \in \mathcal{P}$, $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ tulajdonságú halmazokra

$$\mu^*(M \cap A) + \mu^*(M \cap A^c) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Legyenek a továbbiakban a fenti A_n halmazok rögzülve. Ekkor

$$M \cap A \subseteq \cup_{n=1}^{\infty} (A_n \cap A)$$

egy $\mathcal{P}$-beli halmazokból álló lefedés, ezért μ^* definíciója miatt

$$\mu^*(M \cap A) \leq \sum_{n=1}^{\infty} \mu(A_n \cap A). \tag{\dagger}$$

Hasonlóan $M \cap A^c \subseteq \cup_n (A_n \setminus A)$. Az $A_n \setminus A$ alakú halmazok előállnak mint diszjunkt $\mathcal{P}$-beli halmazok egyesítései, ezért $\cup_{n=1}^{\infty} (A_n \setminus A)$ előáll mint $\mathcal{P}$-beli megszámlálható lefedés. Formálisabban: minden $n \in \mathbb{N}$-hez létezik véges sok diszjunkt $H_i^{(n)} \in \mathcal{P}$, $i = 1, \ldots, k_n$, hogy $A_n \setminus A = \cup_{i=1}^{k_n} H_i^{(n)}$. Így

$$M \cap A^c \subseteq \cup_{n=1}^{\infty} (A_n \setminus A) = \cup_{n=1}^{\infty} \cup_{i=1}^{k_n} H_i^{(n)}.$$

Emlékezzünk vissza, hogy félgyűrűn értelmezett mérték, egyértelműen terjed ki a generált gyűrűre (1.2.9). Jelölje az $\mathrm{r}(\mathcal{P})$ generált gyűrűre kiterjesztett mértéket $\hat{\mu}$. A μ^* generált külső mérték definíciója miatt (3.1.3)

$$\begin{aligned}\mu^*\left(M\cap A^c\right) &\leq \sum_{n=1}^{\infty}\sum_{i=1}^{k_n}\mu\left(H_i^{(n)}\right)=\sum_{n=1}^{\infty}\hat{\mu}\left(A_n\setminus A\right)=\\ &=\sum_{n=1}^{\infty}\hat{\mu}\left(A_n\setminus\left(A_n\cap A\right)\right)=\sum_{n=1}^{\infty}\left(\hat{\mu}\left(A_n\right)-\hat{\mu}\left(A_n\cap A\right)\right)=\\ &=\sum_{n=1}^{\infty}\left(\mu\left(A_n\right)-\mu\left(A_n\cap A\right)\right). \qquad (\ddagger)\end{aligned}$$

Itt kihasználtuk a gyűrűn értelmezett mérték szubtraktivitását, ami $\hat{\mu}\left(A_n\right)=$ $=\mu\left(A_n\right)<\infty$ szerint teljesül. Összevetve a (†) és a (‡) becsléseket majd kihasználva a

$$\sum_{n=1}^{\infty}\mu\left(A_n\cap A\right)\leq\sum_{n=1}^{\infty}\mu\left(A_n\right)<\infty$$

konvergenciát, kapjuk, hogy

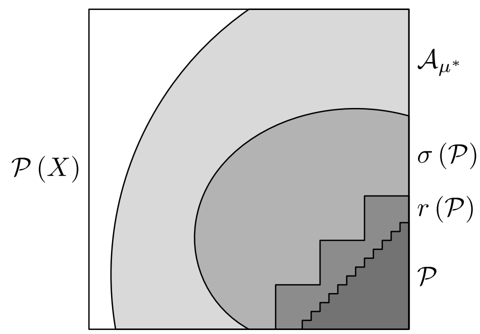
$$\begin{aligned}\mu^*\left(M\cap A\right)+\mu^*\left(M\cap A^c\right) &\leq \sum_{n=1}^{\infty}\mu\left(A_n\cap A\right)+\sum_{n=1}^{\infty}\left(\mu\left(A_n\right)-\mu\left(A_n\cap A\right)\right)=\\ &=\sum_{n=1}^{\infty}\mu\left(A_n\right).\end{aligned}$$

Ezt kellett belátni. □

Ne feledjük, hogy egy külső mértékre, ha az még végesen additív is, akkor a σ-additivitás is teljesül (3.1.1), ergo egy félgyűrűn értelmezett mértéket végül is kiterjesztettünk egy σ-algebrára. Az eddigi négy lépésünk – 3.1.4., 3.1.7., 3.1.8., 3.1.9. – összefoglalását mondjuk Caratheodory-féle kiterjesztési eljárásnak.

3.1.10. (Caratheodory-kiterjesztés (3.1. ábra))
A *Caratheodory-féle kiterjesztési eljáráson* az alábbi eljárást értjük. Tekintsünk egy $\mathcal{P}\subseteq\mathcal{P}(X)$ félgyűrűn értelmezett μ mértéket. Láttuk, hogy ez kiterjeszthető egy az egész hatványhalmazon értelmezett μ^* külső mértékké. Jelölje $\mathcal{A}_{\mu^*}$ az ezen μ^* külső mérték szerint Caratheodory-értelemben mérhető halmazok rendszerét. Tudjuk, hogy $\mathcal{A}_{\mu^*}$ egy a $\mathcal{P}$ félgyűrűt tartalmazó σ-algebra. Jelölje $\tilde{\mu}$ a μ^*-nak erre az $\mathcal{A}_{\mu^*}$ σ-algebrára való megszorítását. Láttuk, hogy $\tilde{\mu}$ egy σ-additív halmazfüggvény és az $(X,\mathcal{A}_{\mu^*},\tilde{\mu})$ teljes mértéktér, és $\tilde{\mu}$ a μ kiterjesztése.

3.1.11. (félgyűrűn értelmezett mérték kiterjesztése a generált σ-algebrára)
Tartsuk meg a fenti jelöléseket. Mivel $\mathcal{P}\subseteq\mathcal{A}_{\mu^*}$ és az $\mathcal{A}_{\mu^*}$ Caratheodory-mérhető halmazok σ-algebrát alkotnak, ezért $\sigma(\mathcal{P})\subseteq\mathcal{A}_{\mu^*}$. Ez azt jelenti,



3.1. ábra. A Caratheodory-kiterjesztés bővülő halmazrendszerei

hogy egy félgyűrűn értelmezett mérték kiterjeszthető a félgyűrű által generált σ-algebrára. Pusztán $\tilde{\mu}$ mértéket kell a generált σ-algebrára tovább szorítanunk. Ilyenkor persze a teljességet semmi nem garantálja!

Felmerül a kérdés, hogy a lehet-e esetleg más módszerrel másfajta mérték kiterjesztéseket gyártani. Látni fogjuk a Fubini-tétel bizonyításakor, hogy más módon is kiterjesztethetnénk egy a szorzat félgyűrűn értelmezett szorzatmértéket a generált szorzat σ-algebrára. A kérdés, hogy mi a kapcsolat a két kiterjesztés között? Azt fogjuk kapni, hogy σ-véges esetben minden, a generált σ-algebráig terjedő mértékkiterjesztés ugyanazt az eredményt adja, mint a fent megismert Caratheodory-féle kiterjesztési eljárás.

3.1.12. állítás. *A kiterjesztési eljárásban definiált $\tilde{\mu}$ a $\mathcal{P}$ félgyűrűn értelmezett mérték maximális kiterjesztése $\sigma(\mathcal{P})$-re azaz, ha ν egy másik mérték $\sigma(\mathcal{P})$-n, melyre $\nu = \mu$ a $\mathcal{P}$-beli halmazokon, akkor $\nu \leq \tilde{\mu}$.*

Bizonyítás. Legyen tehát $E \in \sigma(\mathcal{P})$ rögzítve. Amennyiben $\mu^*(E) = \infty$ úgy készen is vagyunk. Ha van E-nak $E \subseteq \cup_n A_n$ befedése melyre $A_n \in \mathcal{P}$ és $\sum_n \mu(A_n) < \infty$ teljesül, akkor a ν mérték σ-szubadditivitása miatt $\nu(E) \leq \nu(\cup_n A_n) \leq \sum_n \nu(A_n) = \sum_n \mu(A_n)$. Ez épp azt jelenti, hogy $\nu(E) \leq \mu^*(E) = \tilde{\mu}(E)$. □

3.1.13. definíció (σ-véges mérték)**.** Legyen $\mathcal{P}$ egy félgyűrű és μ egy mérték $\mathcal{P}$-n. Azt mondjuk, hogy a μ mérték *σ-véges mérték*, ha léteznek $X_n \in \mathcal{P}$ halmazok, amelyekre $X = \cup_{n=1}^{\infty} X_n$, $X \in \mathcal{P}$ és $\mu(X_n) < \infty$.

3.1.14. (diszjunktizáció)
Ha $X = \cup_{n=1}^{\infty} X_n$ egy $\mathcal{P}$-beli, véges mértékű halmazokkal való előállítás, akkor a diszjunktizációs lemmát használva (1.1.7) kapunk egy $X = \cup_{n=1}^{\infty} Y_n$ diszjunkt előállítást, melyre $Y_n \subseteq X_n$, és $Y_n \in \mathrm{r}(\mathcal{P})$ minden $n \in \mathbb{N}$ mellett.

De tudjuk, hogy a generált gyűrű elemei előállnak mint véges sok diszjunkt félgyűrűbeli elem összege (1.1.12), így minden egyes Y_n halmaz úgy tekinthető, mint véges sok $\mathcal{P}$-beli halmaz diszjunkt egyesítése. Azt kaptuk tehát, hogy egy félgyűrűn értelmezett σ-véges mérték esetén az alaphalmaz előáll megszámlálhatóan sok, diszjunkt, véges mértékű, félgyűrűbeli halmaz egyesítéseként.

3.1.15. tétel (kiterjesztés unicitása a generált σ-algebráig)**.** *Legyen a $\mathcal{P}$ félgyűrűn definiált μ mérték σ-véges. Ekkor a fent definiált $\tilde{\mu}$ a μ-nek egyetlen mérték kiterjesztése azaz, ha ν egy másik mérték $\sigma(\mathcal{P})$-n, melyre $\nu = \mu$ a $\mathcal{P}$-beli halmazokon, akkor $\nu = \tilde{\mu}$.*

Bizonyítás. Első lépésként tegyük fel, hogy $E \in \sigma(\mathcal{P})$ olyan halmaz, melyhez létezik egy $A \in \mathcal{P}$, melyre $E \subseteq A$ és $\mu(A) < \infty$. Ekkor $A = E \cup (A \setminus E)$ és $\nu \leq \mu^*$ miatt

$$\nu(A) = \nu(E) + \nu(A \setminus E) \leq \mu^*(E) + \mu^*(A \setminus E) = \mu^*(A) = \mu(A).$$

No de a $\mathcal{P}$-beli halmazokon a μ és a ν mértékek azonosak, ezért itt középen is egyenlőség van. Kihasználva, hogy a fent kiemelt sorban csupa véges szám szerepel azt kapjuk, hogy

$$0 = (\mu^*(E) - \nu(E)) + (\mu^*(A \setminus E) - \nu(A \setminus E)).$$

Ez viszont azt jelenti, hogy 0 számot előállítottuk két nem negatív szám összegeként, tehát $\tilde{\mu}(E) = \mu^*(E) = \nu(E)$.

Most tegyük fel, hogy $E \in \sigma(\mathcal{P})$ tetszőleges halmaz. Mivel μ σ-véges a $\mathcal{P}$-n, ezért léteznek $A_n \in \mathcal{P}, n \in \mathbb{N}$ diszjunkt halmazok, amelyekre $\mu(A_n) < \infty$ és $X = \cup_n A_n$. Világos, hogy $E = \cup_{n=1}^{\infty}(E \cap A_n)$ és az előzőek szerint $\tilde{\mu}(E \cap A_n) = \nu(E \cap A_n)$. Kihasználva $\tilde{\mu}$ és ν σ -additivitását kapjuk, hogy

$$\tilde{\mu}(E) = \sum_{n=1}^{\infty} \tilde{\mu}(E \cap A_n) = \sum_{n=1}^{\infty} \nu(E \cap A_n) = \nu(E). \qquad \square$$

3.2. Lebesgue-mérték

Alkalmazzuk a kiterjesztési eljárást, a korábbi 1.2.11. definíció esetében: $X = \mathbb{R}$, $\mathcal{P}$ a balról zárt jobbról nyílt korlátos intervallumok félgyűrűje, $\mu = \lambda_f$, ahol $\lambda_f([a,b)) = f(b) - f(a)$ az $f : \mathbb{R} \to \mathbb{R}$ monoton növekedő, balról folytonos függvény megváltozása az $[a,b)$ intervallum felett. Meggondoltuk (1.2.13), hogy ezzel a definícióval λ_f valóban egy σ-additív halmazfüggvény a $\mathcal{P}$ félgyűrűn, amely még σ-véges is. Az $f = \mathrm{id}$ függvény speciális esete különösen fontos. Ekkor $\lambda_{\mathrm{id}}[a,b)$ az intervallum hossza.

3.2.1. definíció (Lebesgue–Stieltjes-mérhető halmazok, Lebesgue–Stieltjes-mérték)**.** Legyen $\mathcal{P}$ balról zárt jobbról nyílt, korlátos intervallumok félgyűrűje, $f : \mathbb{R} \to \mathbb{R}$ egy monoton növekedő, balról folytonos függvény, $\lambda_f : \mathcal{P} \to \mathbb{R}_+$ az f megváltozása az 1.2.11. definíciónak megfelelően. Láttuk, hogy λ_f egy σ-additív halmazfüggvény a $\mathcal{P}$ félgyűrűn (1.2.13). Jelölje λ_f^* a λ_f generálta külső mértéket, amely az $\mathbb{R}$ egész hatványhalmazán van értelmezve, jelölje Λ_f a Caratheodory-értelemben λ_f^* mérhető halmazok σ-algebráját és $\tilde{\lambda}_f$ a λ_f^* külső mértéknek a Λ_f σ-algebrára való megszorítását.

A kiterjesztési tétel szerint $\mathcal{P} \subseteq \Lambda_f$, és $\tilde{\lambda}_f(E) = \lambda_f(E)$ minden $E \in \mathcal{P}$-re, tehát a balról zárt jobbról nyílt intervallumok félgyűrűjén értelmezett függvényt kiterjesztettük egy, az ezt a félgyűrűt tartalmazó, teljes σ-algebrára.

A λ_f^* külső mértéket *Lebesgue–Stieltjes-külső mértéknek*, a Λ_f σ-algebrát az f-re nézve *Lebesgue–Stieltjes-mérhető* halmazok σ-algebrájának nevezzük. Amikor az nem okoz félreértést az egyszerűség kedvéért továbbra is λ_f-fel jelöljük a kiterjesztett halmazfüggvényt és ezt nevezzük *Lebesgue–Stieltjes-mértéknek*.

3.2.2. definíció (Lebesgue-mérhető halmazok, Lebesgue-mérték)**.** Az $f = \mathrm{id}$ függvényre nézve Lebesgue–Stieltjes-mérhető halmazokat *Lebesgue-mérhető halmazoknak* nevezzük és Λ_{id} helyett Λ-val jelöljük. Hasonlóan, a λ_{id}^* Lebesgue–Stieltjes-külső mértéket *Lebesgue-külső mértéknek* nevezzük és λ^* módon jelöljük. Ugyanígy az identitás függvénnyel konstruált $\tilde{\lambda}_{\mathrm{id}}$ mértéket *Lebesgue-mértéknek* mondjuk, és amennyiben az nem okoz félreértést a nehézkes $\tilde{\lambda}_{\mathrm{id}}$ helyett λ jelöli a Lebesgue-mérhető halmazokon értelmezett Lebesgue-mértéket.

3.2.3.

A kiterjesztési eljárás szerint $\mathcal{P} \subseteq \mathcal{A}_{\mu^*}$. Tudjuk, hogy a számegyenes Borel-halmazait generálják a balról zárt jobbról nyílt intervallumok, azaz $\mathcal{B} = \sigma(\mathcal{P}) \subseteq \Lambda_f$. Így $\lambda_f : \Lambda_f \to \bar{\mathbb{R}}_+$ Lebesgue–Stieltjes-mérték értelmezett a Borel-halmazokon és $\lambda_f([a,b)) = f(b) - f(a)$. Hogyan alakul a többi korlátos intervallum Lebesgue–Stieltjes-mértéke? Az $(a,b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b\right)$ azonosságra alkalmazva a mérték monoton folytonossági tételét, azt kapjuk, hogy

$$\lambda_f((a,b)) = \lim_{n\to\infty} \left(f(b) - f\left(a + \frac{1}{n}\right)\right) = f(b) - f(a+),$$

ahol $f(a+)$ jelöli az f függvény a pontbeli jobb oldali határértékét. Hasonlóan, az $[a,b] = \bigcap_{n=1}^{\infty} \left[a, b + \frac{1}{n}\right)$ azonosság vezet az

$$\lambda_f([a,b]) = f(b+) - f(a)$$

értékhez, melynek speciális eseteként minden egyelemű $\{a\}$ halmazra,

$$\lambda_f(\{a\}) = f(a+) - f(a).$$

Ezekből már a végesen additivitás segítségével is adódik a

$$\lambda_f\left((a,b]\right) = f\left(b+\right) - f\left(a+\right)$$

formula. Az id függvény speciális esetében látjuk, hogy a valós egyenes valamennyi intervallumának Lebesgue-mértéke a szóban forgó intervallum hosszával egyezik meg.

3.2.4.
Mivel a balról zárt jobbról nyílt korlátos intervallumok generálják a valós egyenes $\mathcal{B}_{\mathbb{R}}$ Borel-halmazait, ezért

$$\mathcal{B}_{\mathbb{R}} \subseteq \Lambda_{\mathbb{R}}. \tag{3.1}$$

3.2.5. definíció (Lebesgue-mérhető függvény)**.** Az $f : \mathbb{R} \to \bar{\mathbb{R}}$ függvényt, *Lebesgue-mérhetőnek* nevezünk, ha $f^{-1}(A) \in \Lambda_{\mathbb{R}}$ teljesül minden $A \subseteq \mathbb{R}$ nyílt halmaz esetén. Az f függvény Lebesgue-mérhetősége azt jelenti, hogy f mérhető leképezés az $(\mathbb{R}, \Lambda_{\mathbb{R}})$ és az $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ mérhető terek közt a korábbi 1.3.1. definíció értelmében.

3.2.6.
Érdemes összevetnünk a mérhetőség (1.3.1), a Borel-mérhetőség (1.3.4), és a Lebesgue-mérhetőség (3.2.5) fogalmakat valós függvények esetében. A Borel-mérhetőség az

$$(\mathbb{R}, \mathcal{B}_{\mathbb{R}}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$$

mérhető terek közti mérhetőséget jelenti, míg Lebesgue-mérhetőségen a mérhető terek közti

$$(\mathbb{R}, \Lambda_{\mathbb{R}}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$$

mérhetőséget értjük. A (3.1) tartalmazás szerint minden Borel-mérhető valós függvény egyben Lebesgue-mérhető is.

Amíg nyilvánvaló módon teljesül, hogy f, g Borel-mérhető függvények esetén a $f \circ g$ kompozíció is Borel-mérhető, ugyanezt két Lebesgue-mérhető függvényre nem várhatjuk el, hiszen $g^{-1}\left(f^{-1}(F)\right) \in \Lambda_{\mathbb{R}}$-t nem indokolja semmi abban az esetben, mikor $f^{-1}(F) \in \Lambda_R \setminus \mathcal{B}_{\mathbb{R}}$. Látni fogjuk (3.2.12), hogy még g folytonossága mellett is előfordulhat, hogy $f \circ g$ egy nem Lebesgue-mérhető függvény. Ha viszont azt tesszük fel, hogy f Borel-mérhető és g Lebesgue-mérhető, akkor $f \circ g$ Lebesgue-mérhetősége magától értetődik.

A (3.1) tartalmazás valódi is, ezt egy konkrét Lebesgue- de nem Borel-mérhető halmaz megadásával látjuk a 3.2.12. pont alatt, de ennél sokkal több is igaz. Korábban szó volt arról (24. oldal), hogy mélyebb halmazelméleti ismeretek híján nem használjuk a generált σ-algebra belső reprezentációját. A Borel-halmazok számosságának kérdése olyan pont, ahol érdemes kivételt tenni.

3.2.7. (A Borel-halmazok számossága kontinuum)
Egy σ-algebra belső reprezentációjának leírásából az látszik, hogy egy megszámlálhatóan végtelen generátor rendszerrel rendelkező σ-algebra legfeljebb kontinuum számosságú. Fogadjuk el ezt egy másik diszciplína ismert állításaként. Mivel a racionális végpontú, balról zárt jobbról nyílt intervallumok halmaza egy legfeljebb megszámlálható számosságú generátor rendszere a $\mathcal{B}_{\mathbb{R}}$ σ-algebrának, ezért a Borel-halmazok halmazának számossága kontinuum számosságnál nem nagyobb. Viszont a valós egyenes előáll mint megszámlálhatóan végtelen sok balról zárt, jobbról nyílt intervallum, mint Borel-halmaz $\mathbb{R} = \cup_{n=1}^{\infty} P_n$ diszjunkt egyesítése, ezért $\mathbb{N}$ hatványhalmaza és $\mathcal{B}_{\mathbb{R}}$ Borel-halmazok között a

$$M \subseteq \mathbb{N} \qquad M \mapsto \cup_{n \in M} P_n$$

leképezés injekció. No de $\mathbb{N}$ hatványhalmaza kontinuum számosságú, ergo $\mathcal{B}_{\mathbb{R}}$ is éppen kontinuum számosságú halmazrendszer.

3.2.8. (a Lebesgue-mérhető halmazok számossága nagyobb, mint kontinuum)
A bevezető ismeretek közt definiáltuk a Cantor-halmazt (0.1.10). Mivel zárt halmazról van szó, ezért a Cantor-halmaz egy Borel-halmaz, ergo van Lebesgue-mértéke. A Cantor-halmaz definíciója szerint a [0,1] intervallumból megszámlálhatóan sok nyílt intervallum elhagyásával keletkezik, amelyeknek összhossza 1. Eszerint Cantor-halmaz Lebesgue-mértéke zérus. A külső mérték monotonitása szerint minden részhalmazának is zérus a Lebesgue-külső mértéke, tehát minden részhalmaza is Lebesgue-mérhető. Találtunk tehát egy kontinuum számosságú halmazt, amelynek minden részhalmaza is Lebesgue-mérhető. Mivel a hatványhalmaz számossága az eredeti halmaz számosságánál nagyobb, ezért a Lebesgue-mérhető halmazok számossága is nagyobb, mint kontinuum. A Lebesgue-mérhető halmazok halmaza $\mathbb{R}$ hatványhalmazának részhalmaza, ezért a Bernstein-tétel szerint a Lebesgue-mérhető halmazok számossága éppen azonos $\mathbb{R}$ hatványhalmazának számosságával.

Összefoglalva: $\mathbb{B}_{\mathbb{R}}$ halmazrendszer kontinuum (κ) számosságú, míg $\Lambda_{\mathbb{R}}$ halmazrendszer kontinuumnál nagyobb számosságú (2^{κ})!

3.2.9.
Az $\mathbb{R}$ hatványhalmaza tehát a $\Lambda_{\mathbb{R}}$ Lebesgue-mérhető halmazok halmazával ekvipotens, azonos számosságú. Látni fogjuk, hogy a Lebesgue-mérték eltolásinvariáns (3.2.16). Ezt elfogadva és visszatérve a fejezetet bevezető példára, lásd a a 98. oldalt, ha az ottani E halmaz Lebesgue-mérhető volna, akkor minden $r \in [0{,}1)$ mellett $E_r \in \Lambda$ is teljesülne és $\lambda(E_r) = \lambda(E)$ is fennállna, ami az $1 = 0$ vagy az $1 = +\infty$ egyenlőségekre vezetne.

Bebizonyítottuk tehát a következő állítást:

3.2.10. állítás (Példa nem Lebesgue-mérhető halmazra)**.** *Definiáljuk az alábbi ekvivalenciarelációt a* $[0{,}1)$ *intervallum felett. Az* (x, y) *pár pontosan akkor legyen a relációban, ha a különbségük racionális szám, azaz* $x - y \in \mathbb{Q}$*. Tekintsük az intervallum egy olyan részhalmazát, amely az ekvivalenciareláció minden ekvivalenciaosztályából pontosan egy elemet tartalmaz. Ez a halmaz nem Lebesgue-mérhető.*

Az itt bevezetett nem Lebesgue-mérhető halmaznak van még egy extrém tulajdonsága, nevezetesen, hogy annyira sok nem mérhető részhalmaza van, amennyire az csak lehetséges. Egy zérus Lebesgue-külső mértékű halmaz ugyanis egyben Lebesgue-mérhető is, amit úgy is interpretálhatunk, hogy az ilyen halmazok a kicsiségük okán mérhetők. Most meggondoljuk, hogy a fenti nem mérhető halmaznak nincs is más mérhető részhalmaza, csak olyan, amely a kicsisége okán mérhető.

3.2.11. állítás. *Jelölje* E *a 3.2.10. állítás nem Lebesgue-mérhető halmazát.*

1. *Egy* $A \subseteq E$ *halmaz pontosan akkor Lebesgue-mérhető, ha annak Lebesgue-külsőmértéke zérus.*

2. *A számegyenes minden pozitív Lebesgue-mértékű halmaza tartalmaz nem Lebesgue-mérhető részhalmazt.*

Bizonyítás. Jelölje A_r az A halmaz $[0{,}1-r)$ és $[1-r{,}1)$ intervallumokba eső részeinek felcserélésével kapott halmazt ugyanúgy, mint az a nem mérhető halmaz definíciójában már szerepelt. Tehát

$$A_r = (A \cap [0{,}1-r) + r) \cup (A \cap [1-r{,}1) + r - 1)\,.$$

Világos, hogy $A_r \subseteq E_r$ minden $r \in [0{,}1)$ mellett, így az $A_r \cap A_s = \varnothing$ minden $r \neq s$ esetén. Ha tehát A egy Lebesgue-mérhető halmaz, akkor minden A_r is az, a Lebesgue-mérték eltolás-invarianciája szerint. Ugyan e miatt minden szóba jövő r mellett $\lambda(A_r) = \lambda(A)$ is fennáll. Tehát a σ-additivitás következtében:

$$(+\infty) \cdot \lambda(A) = \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(A_r) = \lambda\left(\bigcup_{r \in \mathbb{Q} \cap [0,1)} A_r\right) \leq \lambda([0{,}1)) = 1.$$

Ez csak úgy lehetséges, ha $\lambda(A) = 0$. Megmutattuk tehát, hogy E nem mérhető halmaznak pusztán a nulla mértékű részhalmazai lehetnek mérhetők.

Most legyen $B \in \Lambda$, $\lambda(B) > 0$. A $B \cap [k, k+1)$ halmazok legalább egyike valamely $k \in \mathbb{Z}$ egész mellett pozitív Lebesgue-mértékű. Az eltolás-invariancia miatt feltehető, hogy $k = 0$, vagy ami ugyanaz, feltehető, hogy

$B \subseteq [0,1)$. Mivel az E_r halmazok diszjunkt módon lefedik a teljes $[0,1)$ intervallumot, ezért

$$B = \bigcup_{r \in \mathbb{Q} \cap [0,1)} E_r \cap B.$$

Ha a jobb oldali $E_r \cap B$ halmazok mindegyike mérhető lenne, akkor a bal oldal pozitív mértékűsége szerint, legalább az egyik r mellett $\lambda(E_r \cap B) > 0$ is teljesülne. Visszacserélés után kapnánk az E nem mérhető halmaznak mérhető és pozitív Lebesgue-mértékű részhalmazát, ami a már igazolt első állítás miatt képtelenség. Megmutattuk tehát, hogy legalább egy $r \in \mathbb{Q} \cap [0,1)$ esetén az $E_r \cap B$ halmaz nem Lebesgue-mérhető. □

3.2.12. (Lebesgue-mérhető, de nem Borel-halmaz)
Legyen $f : [0,1] \to [0,1]$ a Cantor-függvény. Tekintsük a $g(x) = f(x) + x$ függvényt. Egy monoton növekedő és egy szigorúan monoton növő függvény összege is szigorúan monoton nő, ezért $g : [0,1] \to [0,2]$ injekció. Mivel $g(0) = 0$, $g(1) = 2$ és g folytonos, így Bolzano-tétele szerint g minden értéket felvesz a $[0,2]$ intervallumból, ergo g folytonos bijekció. Így $g^{-1} : [0,2] \to [0,1]$ inverz is folytonos. Legyen C a Cantor-halmaz és tekintsük a $g(C)$ direktképet. A g^{-1} inverz folytonossága szerint e $g(C)$ direktkép Borel-halmaz. Mivel a Cantor-halmaz definíciójában a kimaradó nyílt intervallumok összhossza 1, és egy ilyen kimaradó intervallumnak a g képe, egy a kimaradó intervallummal azonos hosszú intervallum, $\lambda(g(C^c)) = 1$. Ebből az következik, hogy $\lambda(g(C)) = 1$ is teljesül. Mivel minden pozitív mértékű halmaznak van nem Lebesgue-mérhető részhalmaza, létezik $E \subseteq g(C)$, $E \notin \Lambda_{\mathbb{R}}$. Tekintsük az $F = g^{-1}(E)$ ősképet. Egyrészt világos, hogy $F \in \Lambda_{\mathbb{R}}$ egy Lebesgue-mérhető halmaz, hiszen $F = g^{-1}(E) \subseteq g^{-1}(g(C)) = C$, de a Cantor-halmaz minden részhalmaza is Lebesgue-mérhető. Másrészt, ha F egy Borel-halmaz lenne, akkor a g^{-1} függvény folytonossága szerint $\left(g^{-1}\right)^{-1}(F) = g(F) = E$ halmaz is Borel-mérhető lenne, ami ellentmondana $E \notin \Lambda_{\mathbb{R}}$ választásnak.

Konkrét példát mutattunk tehát

1. Lebesgue-mérhető halmazra, amely nem Borel-mérhető, hiszen $F \in \Lambda_{\mathbb{R}}$, de $F \notin \mathcal{B}_{\mathbb{R}}$;

2. $h = g^{-1}$ folytonos függvényre, amelyre a $h^{-1}(F)$ inverzkép annak ellenére nem Lebesgue-mérhető, hogy az F halmaz Lebesgue-mérhető;

3. Lebesgue-mérhető és folytonos valós függvényekre, amelyek kompozíciója nem Lebesgue-mérhető, hiszen $F \in \Lambda_{\mathbb{R}}$ szerint χ_F Lebesgue-mérhető és a $h = g^{-1}$ folytonos függvényre, a $\chi_F \circ h$ kompozíciója nem Lebesgue-mérhető, a $(\chi_F \circ h)^{-1}(\{1\}) = h^{-1}\left(\chi_F^{-1}\{1\}\right) = h^{-1}(F) = E \notin \Lambda_{\mathbb{R}}$ tartalmazás szerint.

Az n-dimenziós Lebesgue-mérték

3.2.13. definíció (Lebesgue-mérhető halmaz és Lebesgue-mérték)**.** Jelölje $\mathcal{P}$ az $\mathbb{R}^n$ balról zárt jobbról nyílt intervallumainak halmazát. Világos $\mathcal{P}$ az n-szeres szorzata az $\mathbb{R}$ balról zárt jobbról nyílt intervallumait tartalmazó félgyűrűnek, ezért maga is félgyűrű. A monoton konvergenciatétel egyik következményeképpen láttuk, hogy σ-additív halmazfüggvények szorzata is σ-additív, ezért a

$$\lambda^{(n)}\left(\prod_{i=1}^{n}[a_i,b_i)\right)=\times_{i=1}^{n}\lambda\left([a_i,b_i)\right)=\times_{i=1}^{n}\left(b_i-a_i\right) \tag{3.2}$$

módon definiált halmazfüggvény mérték $\mathcal{P}$ félgyűrűn (2.2.16). Alkalmazhatjuk tehát a kiterjesztési eljárást az $\left(\mathbb{R}^n,\mathcal{P},\lambda^{(n)}\right)$ félgyűrűn értelmezett mértékre. Így azt kapjuk, hogy létezik olyan $\Lambda^{(n)}$ σ-algebra melyre $\mathcal{P}\subseteq\Lambda^{(n)}$ és létezik olyan továbbra is $\lambda^{(n)}$-nel jelölt mérték kiterjesztése az eredeti mértéknek a $\Lambda^{(n)}$ σ-algebrára, amely teljes is. A $\Lambda^{(n)}$ σ-algebrát nevezzük az *n-dimenziós Lebesgue-mérhető halmazok* σ-algebrájának és a kiterjesztett $\lambda^{(n)}$ mértéket pedig *n-dimenziós Lebesgue-mértéknek.*

Mikor az kényelmesebb az n-dimenziós Lebesgue-mérhető halmazok halmazát $\Lambda_{\mathbb{R}^n}$ módon is jelöljük, és ha az a szövegkörnyezetből nyilvánvaló az n-dimenziós Lebesgue-mértéket is λ jelöli.

3.2.14.
Mivel az n-dimenziós félig zárt, félig nyílt kockák generálta σ-algebra az n-dimenziós Borel-halmazok σ-algebrája, ezért a

$$\mathcal{B}_{\mathbb{R}^n}\subseteq\Lambda_{\mathbb{R}^n} \tag{3.3}$$

tartalmazás (3.1)-hez hasonlóan most is nyilvánvaló. A kiterjesztési eljárás unicitása (3.1.15) miatt, $\lambda^{(n)}|\mathcal{B}_{\mathbb{R}^n}$ mérték az egyetlen σ-additív kiterjesztése (3.2)-nek az n-dimenziós Borel-halmazok σ-algebrájára.

3.2.15.
Érdemes látnunk, hogy mi a Borel-tégláknak mint a lehető legegyszerűbb Borel-halmazoknak a Lebesgue-mértéke. Ha $B_1\in\mathcal{B}_{\mathbb{R}^s}$ és $B_2\in\mathcal{B}_{\mathbb{R}^t}$ egy-egy s- és t-dimenziós Borel-halmaz, akkor az ezekből alkotott $B_1\times B_2$ tégla egy $n=s+t$-dimenziós Borel-halmaz. Ezek n-dimenziós Lebesgue-mértékére

$$\lambda^{(n)}\left(B_1\times B_2\right)=\lambda^{(s)}\left(B_1\right)\lambda^{(t)}\left(B_2\right). \tag{3.4}$$

Ugyanis 2.2.16. szerint a fenti definíció jobb oldala egy σ-additív halmazfüggvényt definiál a Borel-téglák $\mathcal{B}_{\mathbb{R}^s}\times\mathcal{B}_{\mathbb{R}^t}$ félgyűrűjén. Márpedig a definícióbeli kiterjesztési eljárás az n-dimenziós balról zárt jobbról nyílt kockák generálta

σ-algebráig, tehát $\mathcal{B}_{\mathbb{R}^n}$-ig, egyetlen egy mértéket ír elő. Így pusztán azt használva, hogy $\mathcal{B}_{\mathbb{R}^s} \times \mathcal{B}_{\mathbb{R}^t} \subseteq \mathcal{B}_{\mathbb{R}^n}$ a σ-additív $\lambda^{(n)}$ halmazfüggvénynek egybe kell esnie a Borel-téglákon a (3.4) jobb oldalával.

Speciálisan, ha $B_1, \ldots, B_n \in \mathcal{B}$ egydimenziós Borel-halmazok, akkor ezek szorzatának mértékére

$$\lambda^{(n)}\left(\prod_{i=1}^{n} B_i\right) = \times_{i=1}^{n} \lambda\left(B_i\right).$$

Speciális esetként az is adódik, hogy (3.2) nem csak balról zárt jobbról nyílt, hanem akármilyen intervallumok szorzatára is fennáll, azaz intervallumok szorzatának Lebesgue-mértéke éppen az alkotó intervallumok hosszainak szorzata.

3.2.16. állítás (Lebesgue-mérték eltolás-invariáns)**.** *Az $\mathbb{R}^n$ hatványhalmazán értelmezett Lebesgue-külső mérték eltolás-invariáns, azaz minden $M \subseteq \mathbb{R}^n$ halmazra és minden $x \in \mathbb{R}^n$ vektorra $\lambda^*(E) = \lambda^*(x+E)$.*

A Lebesgue-mérték is eltolás-invariáns, azaz $E \in \Lambda^{(n)}$ és $x \in \mathbb{R}^n$ esetén $(x+E) \in \Lambda^{(n)}$, és $\lambda(x+E) = \lambda(E)$.

Bizonyítás. Először gondoljuk meg, hogy a λ^* Lebesgue-külső mérték eltolás-invariáns. Ha $E \subseteq \mathbb{R}^n$ tetszőleges halmaz, $M_n \in \mathcal{P}, n \in \mathbb{N}$ balról zárt jobbról nyílt intervallumok és $x \in \mathbb{R}^n$ vektor, akkor $x + E \subseteq \cup_{n=1}^{\infty} M_n$ pontosan akkor áll fenn, ha $E \subseteq \cup_{n=1}^{\infty}(M_n - x)$. Az is nyilvánvaló, hogy $\lambda(M_n - x) =$ $= \lambda(M_n)$ ami igazolja, hogy a Lebesgue-külső mértékre $\lambda^*(x+E) = \lambda^*(E)$ tetszőleges $E \subseteq \mathbb{R}^n$ mellett.

Tetszőleges $M, E \subseteq \mathbb{R}^n$ mellett

$$\begin{aligned} M \cap (E+x) &= x + ((M-x) \cap E), \\ M \cap (E+x)^c &= M \cap (E^c + x) = x + ((M-x) \cap E^c). \end{aligned}$$

Ez azt jelenti, hogy az $\lambda^*(M) = \lambda^*(M \cap (E+x)) + \lambda^*(M \cap (E+x)^c)$ és a $\lambda^*(M-x) = \lambda^*((M-x) \cap E) + \lambda^*((M-x) \cap E^c)$ azonosságok egyszerre teljesülnek. Emiatt az E halmaz Lebesgue-mérhetősége mellett $E + x$ is Lebesgue-mérhető. □

3.2.17. állítás (Lebesgue-féle külső mérték kívülről nyílt-reguláris)**.** *Legyen μ^* az $\mathbb{R}^n$ tér Lebesgue-külső mértéke tetszőleges $n \in \mathbb{N}$ mellett, vagy külön az $n = 1$ esetben valamely monoton növő balról folytonos függvénnyel képzett Lebesgue–Stieltjes-külső mérték. Hasonlóan μ az n-dimenziós Lebesgue-mérték, vagy $n = 1$ esetben a megfelelő Lebesgue–Stieltjes-mérték. Ekkor minden $M \subseteq \mathbb{R}^n$ esetén*

$$\mu^*(M) = \inf\{\mu(G) : M \subseteq G, G \text{ nyílt}\}. \tag{3.5}$$

Bizonyítás. Rögzített $M \subseteq \mathbb{R}^n$ részhalmaz mellett jelölje $a \in \overline{\mathbb{R}}_+$ a fent szereplő infimumot. Ha $M \subseteq G$, akkor a külső mérték monotonitása miatt $\mu^*(M) \leq \mu^*(G) = \mu(G)$ fennáll minden G nyílt halmazra, így azok infimumára is, tehát $\mu^*(M) \leq a$.

A fordított irányú egyenlőtlenség nyilván fennáll $\mu^*(M) = \infty$ esetben. Egyébként nézzük először az $M \in \mathcal{P}$ speciális esetet, ahol $\mathcal{P}$ a balról zárt jobbról nyílt (esetleg n-dimenziós) intervallumok félgyűrűje.

Ha a $\mu^* = \lambda^*{}_f$ Lebesgue–Stieltjes-külső mértékről van szó az f függvénnyel, akkor $\lambda_f : \mathcal{P} \to \mathbb{R}_+$ függvény regularitási tulajdonsága szerint (1.2.12), $\varepsilon > 0$-hoz létezik $M \subseteq (a, b)$ nyílt intervallum, hogy

$$\lambda_f^*(M) + \varepsilon = \lambda_f(M) + \varepsilon > f(b) - f(a) \geq f(b) - f(a+) = \lambda_f((a, b)),$$

ami igazolja a Lebesgue–Stieltjes-mérték esetét.

Ha $\mu = \lambda$ az n-dimenziós Lebesgue-mérték, akkor az $M = \prod_{i=1}^n [a_i, b_i)$ balról zárt jobbról nyílt kockához és $\varepsilon > 0$-hoz, a szorzás művelet folytonossága szerint, létezik $\delta > 0$, hogy $\Pi_{i=1}^n (b_i - a_i + \delta) < \Pi_{i=1}^n (b_i - a_i) + \varepsilon$. Így találtunk $M \subseteq \prod_{i=1}^n (a_i - \delta, b_i)$ nyílt kockát, amelyre

$$\begin{aligned}\lambda^*(M) + \varepsilon = \lambda(M) + \varepsilon = \Pi_{i=1}^n (b_i - a_i) + \varepsilon > \\ > \Pi_{i=1}^n (b_i - a_i + \delta) = \lambda\left(\prod_{i=1}^n (a_i - \delta, b_i)\right),\end{aligned}$$

ami az n-dimenziós Lebesgue-mérték esetében is igazolja a kívánt (3.5) egyenlőséget az $M \in \mathcal{P}$ speciális esetben.

Az általános esetben egy rögzített $\varepsilon > 0$ szám mellett, vegyünk $H_k \in \mathcal{P}$ balról zárt jobbról nyílt, esetleg n-dimenziós intervallumokat, amelyekre

$$M \subseteq \cup_{k=1}^\infty H_k \quad \text{és} \quad \sum_{k=1}^\infty \mu(H_k) < \mu^*(M) + \frac{\varepsilon}{2}.$$

Most minden egyes $H_k \in \mathcal{P}$ halmazra alkalmazzuk a már igazolt állítást, azaz válasszunk olyan $G_k \subseteq \mathbb{R}^n$ nyílt halmazokat, hogy

$$H_k \subseteq G_k \quad \text{és} \quad \mu(G_k) < \mu(H_k) + \frac{\varepsilon}{2^{k+1}}.$$

Jelölje $G = \cup_{k=1}^\infty G_k$ ezen nyílt halmazok egyesítését. Világos, hogy $G \subseteq \mathbb{R}^n$ nyílt halmaz és $M \subseteq G$. No de

$$\mu(G) \leq \sum_{k=1}^\infty \mu(G_k) < \sum_{k=1}^\infty \mu(H_k) + \frac{\varepsilon}{2} < \mu^*(M) + \varepsilon.$$

Így $a \leq \mu^*(M) + \varepsilon$ minden pozitív ε mellett fennáll, ergo $a \leq \mu^*(M)$ is teljesül. $\square$

3.2.18. definíció ($\mathcal{G}_\delta$ és $\mathcal{F}_\sigma$ halmazok)**.** $\mathcal{G}_\delta$ halmaznak nevezünk egy topologikus térbeli halmazt, amennyiben az előáll megszámlálhatóan sok nyílt halmaz metszeteként, és $\mathcal{F}_\sigma$-nak, ha az előáll megszámlálhatóan sok zárt halmaz egyesítéseként.

3.2.19.
Világos, hogy ha $\mathcal{B}$ jelöli valamely topologikus tér Borel-halmazait, akkor $\mathcal{G}_\delta \subseteq \sigma(\mathcal{G}) = \mathcal{B}$ és hasonlóan $\mathcal{F}_\sigma \subseteq \sigma(\mathcal{F}) = \mathcal{B}$.

Az $\mathcal{F}_\sigma$ és a $\mathcal{G}_\delta$ halmazrendszerek elemeire úgy tekinthetünk mint az egyik legegyszerűbb típusú Borel-halmazokra.

3.2.20. állítás (Lebesgue-mérhető halmaz közelítése)**.** *Legyen μ az n-dimenziós Lebesgue-mérték tetszőleges $n \in \mathbb{N}$ mellett, vagy $n = 1$ esetben egy Lebesgue–Stieltjes-mérték. Ekkor minden $M \subseteq \mathbb{R}^n$ Lebesgue-mérhető halmazhoz, vagy a Lebesgue–Stieltjes-mérték esetén minden $M \subseteq \mathbb{R}$ Lebesgue–Stieltjes-mérhető halmazhoz*

1. *és minden $\varepsilon > 0$ számhoz létezik G nyílt halmaz, amelyre*
$$M \subseteq G, \text{ és } \mu(G \setminus M) < \varepsilon.$$
2. *és minden $\varepsilon > 0$-hoz létezik olyan G nyílt és F zárt halmaz, hogy*
$$F \subseteq M \subseteq G, \text{ valamint } \mu(G \setminus F) < \varepsilon.$$
3. *létezik $G \in \mathcal{G}_\delta$ és létezik $F \in \mathcal{F}_\sigma$ halmaz, amelyekre*
$$F \subseteq M \subseteq G, \text{ valamint } \mu(G \setminus F) = 0.$$

Bizonyítás. Sorban igazoljuk az állításokat.

1. Ha $\mu(M) < \infty$, akkor az előző állítás szerint a tételben rögzített $\varepsilon > 0$-hoz létezik olyan nyílt $M \subseteq G$ halmaz, amelyre $\mu(G) < \mu(M) + \varepsilon$. Az M halmaz Lebesgue-mérhetősége (Lebesgue–Stieltjes-mérhetősége) szerint a $G \setminus M$ is Lebesgue-mérhető (Lebesgue–Stieltjes-mérhető), $\mu(M)$ végessége és a μ mérték végesen additivitása miatt
$$\mu(G \setminus M) = \mu(G) - \mu(M) < \varepsilon.$$
Ha $\mu(M) = \infty$, akkor a Lebesgue-mérték (Lebesgue–Stieltjes-mérték) σ-végessége szerint $M = \cup_{k=1}^{\infty} M_k$ alakban áll elő, ahol M_k Lebesgue-mérhető (Lebesgue–Stieltjes-mérhető) minden $k \in \mathbb{N}$ mellett és $\mu(M_k) < \infty$. Így a már igazolt gondolatot felhasználva kapunk $M_k \subseteq G_k$ nyílt halmazokat, amelyekre $\mu(G_k \setminus M_k) < \frac{\varepsilon}{2^k}$. Bevezetve a $G = \cup_{k=1}^{\infty} G_k$ halmazt a G olyan nyílt halmaz, amelyre $M \subseteq G$ és $G \setminus M \subseteq \cup_{k=1}^{\infty} (G_k \setminus M_k)$, ezért
$$\mu(G \setminus M) \leq \sum_{k=1}^{\infty} \mu(G_k \setminus M_k) < \varepsilon.$$

2. Rögzített $\varepsilon > 0$ mellett alkalmazzuk az eddig igazoltakat, de az M és az M^c halmazokra! Kapunk tehát $M \subseteq G$ halmazt, amelyre G nyílt és $\mu(G \setminus M) < \frac{\varepsilon}{2}$, valamint $M^c \subseteq F^c$ halmazt, amelyre F zárt és $\mu(F^c \setminus M^c) < \frac{\varepsilon}{2}$. Persze $F^c \setminus M^c = M \setminus F$, így az $F \subseteq M$ olyan zárt halmaz, amelyre $\mu(M \setminus F) < \frac{\varepsilon}{2}$ is fennáll. Ekkor a $G \setminus F \subseteq (G \setminus M) \cup (M \setminus F)$ tartalmazás szerint
$$\mu(G \setminus F) \leq \mu(G \setminus M) + \mu(M \setminus F) < \varepsilon.$$
3. Alkalmazzuk $\varepsilon = \frac{1}{k}$ mellett az előző állítást. Léteznek tehát $F_k \subseteq M \subseteq G_k$ halmazok, amelyekre F_k zárt, G_k nyílt és $\mu(G_k \setminus F_k) < \frac{1}{k}$. Jelölje $G = \cap_{k=1}^{\infty} G_k$ és $F = \cup_{k=1}^{\infty} F_k$. Világos, hogy $G \in \mathcal{G}_\delta$, $F \in \mathcal{F}_\sigma$, $F \subseteq M \subseteq G$. No de $G \setminus F \subseteq G_k \setminus F_k$ minden $k \in \mathbb{N}$ mellett, ezért $\mu(G \setminus F) \leq \mu(G_k \setminus F_k) < \frac{1}{k}$ is fennáll, azaz csak $\mu(G \setminus F) = 0$ lehetséges. □

3.2.21. állítás. *Szorítsuk meg az n-dimenziós Lebesgue-mértéket a Borel-mérhető halmazokra. Az így kapott $(\mathbb{R}^n, \mathcal{B}, \lambda|_{\mathcal{B}})$ mértéktér teljessé tétele a Lebesgue-féle $(\mathbb{R}^n, \Lambda, \lambda)$ mértéktér.*

Hasonlóan, az $n = 1$-dimenziós esetben, ha megszorítunk egy Lebesgue–Stieltjes-mértéket a valós egyenes Borel-mérhető halmazaira, és a kapott $(\mathbb{R}, \mathcal{B}, \lambda_f|_{\mathcal{B}})$ mértékteret teljessé tesszük, akkor a Lebesgue–Stieltjes-féle $(\mathbb{R}, \Lambda_f, \lambda_f)$ mértékteret kapjuk.

Bizonyítás. Jelölje $\left(\mathbb{R}^n, \bar{\mathcal{B}}, \bar{\lambda}\right)$ az n-dimenziós Lebesgue-mérték Borel-halmazokra megszorítottjának a teljessé tételét. Visszaemlékezve a mértéktér teljessé tételének konstrukciójára (2.2.31) világos, hogy az előző állítás (3.2.20) éppen $\Lambda \subseteq \bar{\mathcal{B}}$ tartalmazást jelenti. Mivel a teljessé tétel a legszűkebb teljes kiterjesztés, így $\bar{\mathcal{B}} = \Lambda$.

Ha $M \in \Lambda$ és $E, F \in \mathcal{B}$ olyan halmazok, amelyekre
$$E \subseteq M \subseteq F, \quad \text{és} \quad \lambda(F \setminus E) = 0,$$
akkor a teljessé tett mérték definíciója szerint $\bar{\lambda}(M) = \lambda(E)$. Persze $M = E \cup (M \setminus E)$ és a Lebesgue-mérték teljessége miatt $\lambda(M \setminus E) = 0$, így
$$\bar{\lambda}(M) = \lambda(E) = \lambda(E) + \lambda(M \setminus E) = \lambda(M).$$
Ezzel igazoltuk a $\bar{\lambda} = \lambda$ azonosságot is, azaz az $\left(\mathbb{R}^n, \bar{\mathcal{B}}, \bar{\lambda}\right)$ és a $(\mathbb{R}^n, \Lambda, \lambda)$ mértékterek egybeesnek. A fenti okoskodáshoz egyedül a Lebesgue-mérhető halmazok Borel-halmazokkal való közelíthetőségét használtuk (3.2.20), ami Lebesgue–Stieltjes-mértékre is igaz. Így a fentiek szó szerinti ismétlésével kapjuk a Lebesgue–Stieltjes-mértékre vonatkozó állítás indoklását. □

Természetesen merül fel a kérdés, hogy a Lebesgue-mérhető halmazokból alkotott téglák, Lebesgue-mérhetőek-e?

3.2.22. állítás. *Legyen $A_1 \in \Lambda_{\mathbb{R}^s}$ egy s-dimenziós, és $A_2 \in \Lambda_{\mathbb{R}^t}$ egy t-dimenziós Lebesgue-mérhető halmaz. Ekkor az ezekből alkotott tégla egy $n = s + t$-dimenziós Lebesgue-mérhető halmaz, azaz*

$$A_1 \times A_2 \in \Lambda_{\mathbb{R}^n},$$

továbbá ezek Lebesgue-mértékeire

$$\lambda^{(n)}(A_1 \times A_2) = \lambda^{(s)}(A_1)\,\lambda^{(t)}(A_2).$$

Bizonyítás. Az $A_1 \in \Lambda_{\mathbb{R}^s}$ miatt léteznek $E_1, F_1 \in \mathcal{B}_{\mathbb{R}^s}$ Borel-halmazok, melyekre

$$E_1 \subseteq A_1 \subseteq F_1 \quad \text{és} \quad \lambda^{(s)}(F_1 \setminus E_1) = 0.$$

Hasonlóan vannak olyan t-dimenziós $E_2, F_2 \in \mathcal{B}_{\mathbb{R}^t}$ Borel-halmazok, amelyekre

$$E_2 \subseteq A_2 \subseteq F_2 \quad \text{és} \quad \lambda^{(t)}(F_2 \setminus E_2) = 0.$$

Világos, hogy az $\mathbb{R}^n = \mathbb{R}^s \times \mathbb{R}^t$ szorzaton fennáll az

$$E_1 \times E_2 \subseteq A_1 \times A_2 \subseteq F_1 \times F_2,$$

tartalmazás, és a két Borel-tégla különbségére

$$(F_1 \times F_2) \setminus (E_1 \times E_2) = ((F_1 \setminus E_1) \times F_2) \cup (F_1 \times (F_2 \setminus E_2))$$

teljesül. Tekintetbe véve, hogy Borel-téglák mértéke az alkotó Borel-halmazok mértékeinek szorzata (3.2.15), ezért a bal oldali halmaz n-dimenziós Lebesgue-mértéke zérus. Igazoltuk, hogy az $A_1 \times A_2$ halmaz az n-dimenziós Borel-halmazok $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \lambda^{(n)}|\mathcal{B}_{\mathbb{R}^n}\right)$ mértéktere teljessé tételének, tehát az n-dimenziós Lebesgue-mérhető halmazok halmazának egy eleme, és e halmaz Lebesgue-mértékére

$$\lambda^{(n)}(A_1 \times A_2) = \lambda^{(n)}(E_1 \times E_2) = \lambda^{(s)}(E_1)\,\lambda^{(t)}(E_2) = \lambda^{(s)}(A_1)\,\lambda^{(t)}(A_2).$$

Ezt kellett belátni. □

3.2.23.
Persze az iménti állítás úgy formalizálható, hogy $\Lambda_{\mathbb{R}^s} \times \Lambda_{\mathbb{R}^t} \subseteq \Lambda_{\mathbb{R}^n}$, ahol $n = s + t$. Ebből azonnal következik, hogy a generált σ-algebrára is, azaz a szorzat σ-algebrára is, fennáll az

$$\Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t} = \sigma(\Lambda_{\mathbb{R}^s} \times \Lambda_{\mathbb{R}^t}) \subseteq \Lambda_{\mathbb{R}^n} \tag{3.6}$$

tartalmazás. Felmerül, hogy esetleg egyenlőséget is írhatnánk itt hasonlóan ahhoz, ahogyan Borel-halmazok szorzata is kiadja a teljes Borel-halmazok halmazát (1.1.29). Ha $A \subseteq \mathbb{R}^s$ egy tetszőleges nem üres halmaz, és $B \notin \Lambda_{\mathbb{R}^t}$ egy nem Lebesgue-mérhető halmaz, akkor látni fogjuk a 4.1.4. állítás következményeként, hogy az $A \times B$ tégla nem lehet a $\Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}$ szorzat-σ-algebrára nézve sem mérhető. Ez azt jelenti, hogy ha $A \in \Lambda_{\mathbb{R}^s}$ nem üres nullmértékű halmazt választunk, akkor az $A \times \mathbb{R}^t \in \Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}$ egy olyan n-dimenziós Lebesgue-nullmértékű halmaz, amelynek az $A \times B$ egy nem mérhető részhalmaza. Ez azt jelenti, hogy a (3.6) bal oldali szorzat-σ-algebrája az n-dimenziós Lebesgue-mérték megszorításával egy nem teljes mértéktér, ergo (3.6) bal és jobb oldala nem azonos. Konkrétan $A \times B$ a jobb oldalnak eleme, de a bal oldalnak nem.

Érdemes összefoglalnunk a Borel- és Lebesgue-mérhető halmazok, valamint ezek szorzatának kapcsolatát:

$$\mathcal{B}_{\mathbb{R}^n} = \mathcal{B}_{\mathbb{R}^s} \otimes \mathcal{B}_{\mathbb{R}^t} \subseteq \Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t} \subseteq \Lambda_{\mathbb{R}^n},$$

ahol mindkét tartalmazás szigorú.

3.2.24. állítás. *Jelölje $n = s + t$ pozitív egészek mellett $\Sigma = \Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}$ a Lebesgue-mérhető halmazok σ-algebrájának szorzatát. Legyen μ az e szorzat σ-algebrára szűkített n-dimenziós Lebesgue-mérték, azaz $\mu = \lambda^{(n)}|_{\Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}}$. Ekkor az $(\mathbb{R}^n, \Sigma, \mu)$ mértéktér teljessé tétele az n-dimenziós $\left(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda^{(n)}\right)$ Lebesgue-mértéktér.*

Bizonyítás. Láttuk, hogy a $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \lambda^{(n)}|_{\mathbb{B}_{\mathbb{R}^n}}\right)$ Borel-mértéktér teljessé tétele a tételbeli Lebesgue-mértéktér. Mivel $\mathcal{B}_{\mathbb{R}^n} \subseteq \Sigma$ és mivel $\mu|_{\mathcal{B}_{\mathbb{R}^n}} = \lambda^{(n)}|_{\mathcal{B}_{\mathbb{R}^n}}$ ezért $(\mathbb{R}^n, \Sigma, \mu)$ teljessé tétele is $\left(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda^{(n)}\right)$. $\square$

Kapcsolat a Riemann-integrállal

A Lebesgue-mértékkel kapott absztrakt integrál elmélet egybeesik a klasszikus Riemann-féle integrál-elmélettel olyan függvényekre, melyekre ez utóbbinak értelme van:

3.2.25. állítás (Riemann-integrálhatóság Lebesgue-kritériuma)**.** *Legyen $f : [a, b] \to \mathbb{R}$ egy korlátos és zárt intervallumon értelmezett korlátos függvény.*

1. *Ha f Riemann-integrálható, akkor f Lebesgue-mérhető is és a Lebesgue-integrál azonos a Riemann-integrállal, azaz*

$$\int_a^b f(x)\, dx = \int_{[a,b]} f\, d\lambda,$$

2. *Az f függvény pontosan akkor Riemann-integrálható, ha szakadási helyeinek halmaza Lebesgue null-mértékű.*

Bizonyítás. A bizonyítás alapja a Darboux-tétel, mely szerint tetszőleges nullához finomodó felosztás-sorozatra az alsó közelítő összegek sorozata az alsó Riemann-integrálhoz tart, míg a felső közelítő összegek sorozata a felső Riemann-integrálhoz.

Legyen $\mathcal{F}_n$ az $[a,b]$ intervallum olyan egyforma hosszúságú 2^n darab intervallumából álló partíciója úgy, hogy $\mathcal{F}_{n+1}$ intervallumai az $\mathcal{F}_n$ intervallumainak felezésével keletkezzenek. Definiálja

$$G_n = \sum_{\kappa \in \mathcal{F}_n} \left(\sup_{\kappa} f \right) \chi_\kappa, \qquad g_n = \sum_{\kappa \in \mathcal{F}_n} \left(\inf_{\kappa} f \right) \chi_\kappa.$$

Mivel az intervallumok Borel-mérhetők, ezért g_n, G_n Borel-mérhető függvények, melyekre $\mathcal{F}_n$ partíció volta miatt $g_n \leq f \leq G_n$. Persze $\int_{[a,b]} G_n \, d\lambda$ az $\mathcal{F}_n$ felosztáshoz tartozó felső közelítő összeg és $\int_{[a,b]} g_n \, d\lambda$ az alsó. A Darboux-tétel szerint az alsó közelítő összegek monoton növő sorozatának és a felső közelítőö sszegek monoton fogyó sorozatának a határértéke az alsó, illetve a felső Riemann-integrál. Most vegyük észre, hogy $G_{n+1} \leq G_n$ és $g_{n+1} \geq g_n$ monotonitások is fennállnak, hiszen úgy particionáltuk az $[a,b]$ intervallumot, hogy $\mathcal{F}_{n+1}$ minden intervalluma részhalmaza az $\mathcal{F}_n$ valamely intervallumának. Jelölje

$$G = \lim_{n \to \infty} G_n, \qquad g = \lim_{n \to \infty} g_n$$

Borel-mérhető függvényeket. No persze $g \leq f \leq G$ az egész intervallumon fennáll. Mivel f korlátos és az értelmezési tartomány Lebesgue-mértéke véges, ezért alkalmazhatjuk a dominált konvergenciatételt (2.3.21). Azt kapjuk tehát, hogy

$$\int_{[a,b]} G \, d\lambda = \lim_{n \to \infty} \int_{[a,b]} G_n \, d\lambda = \int_a^{*b} f(x) \, dx,$$
$$\int_{[a,b]} g \, d\lambda = \lim_{n \to \infty} \int_{[a,b]} g_n \, d\lambda = \int_{*a}^{b} f(x) \, dx.$$

Hogy a Riemann-integrálhatóság mellett a folytonosságot is megfogjuk, vezessük be a

$$H(x) = \lim_{n \to \infty} \sup_{(x-1/n, x+1/n)} f, \qquad h(x) = \lim_{n \to \infty} \inf_{\left(x - \frac{1}{n}, x + \frac{1}{n}\right)} f$$

függvényeket. Az f korlátossága miatt a fenti határértékek monoton fogyó, illetve növő sorozatok pontonkénti határértékei, tehát $H, h : [a,b] \to \mathbb{R}$ jól definiált függvények. Világos, hogy $h(x) \leq f(x) \leq H(x)$ minden $x \in [a,b]$ mellett, és

$$H(x) - h(x) = \lim_{n \to \infty} \sup_{u,v \in \left(x - \frac{1}{n}, x + \frac{1}{n}\right)} |f(u) - f(v)|.$$

Ebből leolvasható, hogy az f függvény pontosan akkor folytonos az x pontban, ha $h(x) = H(x)$ egyenlőség teljesül.

Olyan pont, amely egyetlen n-re sem osztópontja egyetlen $\mathcal{F}_n$ partíciónak sem, belső pontja az összes partíció valamely intervallumocskájának. Ilyen pontra $H(x) = G(x)$ és $h(x) = g(x)$ teljesül. Mivel megszámlálható halmaz Lebesgue-mértéke zérus, ezért $H = G\ [\lambda]$ és $h = g\ [\lambda]$, így figyelembe véve a Lebesgue-mérték teljességét, H és h is Lebesgue-mérhető függvények (2.2.29). Összefoglalva az eddigieket azt találtuk, hogy

$$\int_a^{*b} f(x)\ dx - \int_{*a}^{b} f(x)\ dx = \int_{[a,b]} G - g\, d\lambda = \int_{[a,b]} H - h\, d\lambda.$$

Eszerint a függvény Riemann-integrálhatósága azzal ekvivalens, hogy $h = H\ [\lambda]$ teljesül, azaz f szakadási helyeinek halmaza Lebesgue null-mértékű, amivel az második állítás ekvivalenciáját meggondoltuk.

Riemann-integrálható f esetén tehát $h = f = H\ [\lambda]$ áll fenn. A Lebesgue-mérték teljessége szerint f Lebesgue-mérhető (2.2.29), és

$$\int_a^b f(x)\ dx = \int_{[a,b]} g\, d\lambda \le \int_{[a,b]} f\, d\lambda \le \int_{[a,b]} G\, d\lambda = \int_a^b f(x)\ dx. \qquad \square$$

3.2.26. (Példa Riemann-integrálható nem Borel-mérhető függvényre) Nem igaz, hogy egy Riemann-integrálható függvény Borel-mérhető is lenne. Tekintsünk ugyanis egy korlátos és zárt, kontinuum számosságú, de zérus Lebesgue-mértékű halmazt, mondjuk a [0,1] intervallumon a Cantor-halmazt. Mivel ennek hatványhalmaza az összes Borel-halmazok halmazánál is nagyobb számosságú, ezért e C halmaznak van nem Borel-mérhető $A \subseteq C$ részhalmaza. Világos, hogy χ_A a C^c halmazon, ergo egy nyílt halmazon konstans, ezért annak minden pontjában folytonos. A Lebesgue-kritérium szerint χ_A Riemann-integrálható függvény, ami persze nem Borel-mérhető.

3.3. Lebesgue-mérték és determináns

3.3.1. Szinguláris felbontási tétel

3.3.1. állítás. *Legyen $(X, \|\cdot\|)$ és $(Y, \|\cdot\|)$ véges dimenziós skaláris szorzatos tér és $T : X \to Y$ lineáris operátor. Ekkor léteznek $\{v_1, v_2, \ldots, v_m\} \subseteq X$ és $\{u_1, u_2, \ldots, u_n\} \subseteq Y$ ortonormált bázisok, és léteznek $\sigma_1, \sigma_2, \ldots, \sigma_l \ge 0$ nem negatív valós számok, amelyekre*

$$Tv_i = \sigma_i u_i$$

fennáll minden $i = 1, \ldots, l$ mellett, ahol $l = \min\{m, n\}$.

Bizonyítás. Tekintsük a $T^*T : X \to X$ lineáris operátort. Világos, hogy T^*T szimmetrikus így a spektrál felbontási tétel szerint van $\{v_1, v_2, \ldots, v_m\} \subseteq$ $\subseteq X$ az X térben ortonormált bázis, és vannak $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ valós számok, amelyekre

$$T^*Tv_i = \lambda_i v_i$$

fennáll minden $i = 1, \ldots, m$ esetén. Vegyük észre, hogy

$$\langle Tv_i; Tv_j\rangle = \langle T^*Tv_i; v_j\rangle = \langle \lambda_i v_i; v_j\rangle = \lambda_i\langle v_i; v_j\rangle.$$

Ebből azonnal látszik, hogy $\{Tv_i : i = 1, \ldots, m\}$ egy ortogonális rendszer Y-ban, másrészt $0 \leq \|Tv_i\|^2 = \lambda_i\|v_i\|^2 = \lambda_i$. Esetleges átindexelés után feltehető, hogy $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m \geq 0$. Most definiáljuk az Y-beli ortonormált rendszert. Mivel

$$\{Tv_i : Tv_i \neq 0\} = \{Tv_i : \lambda_i \neq 0\}$$

bázisa a T operátor képterének ezért, ha r a T operátor rangja, akkor $m \geq$ $\geq i > r$ mellett $\lambda_i = 0$, és minden $i \leq r$ mellett $\lambda_i > 0$. Legyen tehát $i \leq r$ esetén

$$\sigma_i = \sqrt{\lambda_i}, \quad u_i = \frac{1}{\sigma_i}Tv_i.$$

Ilyen módon $\{u_1, \ldots, u_r\}$ ortonormált, amely persze T képterének bázisa is. Ha T nem lenne szürjektív, akkor egészítsük ki Gramm–Schmidt-ortogonalizációval a képtér iménti bázisát az Y normált tér $\{u_1, ..., u_r, u_{r+1}, u_{r+2}, ..., u_n\}$ ortonormált bázisává, és $r < i \leq \min\{m, n\}$ mellett legyen $\sigma_i = 0$.

Világos, hogy a definiált ortonormált bázisok és σ_i számok megfelelnek a tétel követelményeinek. Hiszen $i \leq r$ mellett a fent kiemelt sor éppen $Tv_i =$ $= \sigma_i u_i$ egyenlőséget jelenti, és $r < i \leq \min\{m, n\}$ esetben $\|Tv_i\| = \sqrt{\lambda_i} = 0$, ergo $Tv_i = 0 = \sigma_i u_i$ most is fennáll. □

3.3.2. következmény. *Minden $T : X \to Y$ véges dimenziós skaláris szorzatos terek közt ható lineáris operátorhoz van az X térnek és az Y térnek olyan ortonormált bázisa, amely bázisokban felírva T mátrixát e mátrixnak csak a diagonálisában vannak nemzérus elemek és a diagonálisban minden elem nem negatív.*

Speciálisan, ha $T : X \to X$ egy lineáris bijekció, akkor van a térnek egy-egy ortonormált bázisa, amelyekben felírva a transzformáció mátrixát e mátrix diagonális és a diagonálisban minden elem pozitív.

3.3.3. következmény (SVD)**.** *Legyen $T \in \mathbb{R}^{n\times m}$ tetszőleges mátrix. Ekkor létezik $V \in \mathbb{R}^{m\times m}$ ortogonális mátrix, és létezik $U \in \mathbb{R}^{n\times n}$ ortogonális mátrix, és létezik $S \in \mathbb{R}^{n\times m}$ diagonális mátrix, amelynek diagonálisában csupa nem negatív elem szerepel és*

$$T = USV^*.$$

Bizonyítás. A szokásos belső szorzás mint skalárszorzás mellett alkalmazzuk az előző állítást. Létezik tehát $\{v_1, \ldots, v_m\} \subseteq \mathbb{R}^m$ ortonormált bázis és létezik $\{u_1, ..., u_n\} \subseteq \mathbb{R}^n$ ortonormált bázis és léteznek $\{\sigma_i \geq 0 : i = 1, ..., \min\{m, n\}\}$ nem negatív valós számok, amelyre

$$Tv_i = \sigma_i u_i$$

minden szóba jövő i mellett.

Álljon $V \in \mathbb{R}^{m\times m}$ mátrix a $v_1, v_2, \ldots, v_m$ oszlopvektorokból és hasonlóan az $U \in \mathbb{R}^{n\times n}$ mátrix az $u_1, u_2, \ldots, u_n$ oszlopvektorokból. Világos, hogy U és V ortogonális mátrixok. Legyen $S \in \mathbb{R}^{n\times m}$ diagonális mátrix, melyben az i-edik diagonális elem σ_i minden $i = 1, \ldots, \min\{m, n\}$. Ekkor mátrix szorzás definíciója szerint

$$TV = US.$$

Persze $V^* = V^{-1}$ miatt készen is vagyunk. □

3.3.4. állítás. *Legyenek μ és ν mértékek az $\mathbb{R}^n$ Borel-halmazain értelmezve, amelyek minden racionális végpontú I intervallumon egybeesnek és végesek, azaz*

$$\mu(I) = \nu(I) < \infty.$$

Ekkor a két mérték azonos, tehát minden A Borel-halmaz esetén is

$$\mu(A) = \nu(A).$$

Bizonyítás. Tudjuk, hogy σ-véges mértékek mellett a mérték kiterjesztése egyértelmű a félgyűrűről a generált gyűrűig. Mivel a racionális végpontú balról zárt jobbról nyílt intervallumok félgyűrűje a Borel-halmazokat generálja, ezért ha a kiterjesztési eljárást alkalmazzuk a két mértékre, azt kapjuk, hogy minden A Borel-halmaz mellett $\mu(A) = \mu^\star(A) = \nu(A^\star) = \nu(A)$. □

3.3.5. állítás (Eltolás invariáns Borel-mérték)**.** *Legyen μ egy eltolás-invariáns mérték az $\mathbb{R}^n$ Borel-halmazain és λ a Lebesgue-mérték. Ekkor létezik $c \geq 0$ konstans, hogy minden A Borel-halmaz mellett*

$$\mu(A) = c\lambda(A).$$

Bizonyítás. Legyen I a [0,1) intervallum n-szeres szorzata és $c = \mu(I)$. Világos, hogy

$$\mu(I) = c \cdot 1 = c\lambda(I) \tag{3.7}$$

teljesül. Ha megmutatjuk, hogy a fenti (3.7) egyenlőség minden más racionális végpontú balról zárt, jobbról nyílt intervallumra is fennáll, akkor az előző tétel szerint a μ és a cm Borel mértékek azonosak.

Legyenek ez okból a $l_1, l_2, ..., l_n$ pozitív egészek rögzítve. Jelölje $I_{l_1,l_2,...,l_n} = [0,1/l_1) \times [0,1/l_2) \times \ldots \times [0,1/l_n)$. Az I intervallum előáll mint a $l_1 l_2 \cdots l_n$ darab $I_{l_1,...,l_n}$ eltolt diszjunkt egyesítése. Így $c = \mu(I) = l_1 l_2 \cdots l_n \mu(I_{l_1,l_2,...,l_n})$, amiből kapjuk, hogy

$$\mu(I_{l_1,l_2,...,l_n}) = c \frac{1}{l_1} \frac{1}{l_2} \cdots \frac{1}{l_n} = c\lambda(I_{l_1,l_2,...,l_n})$$

Ez azt jelenti, hogy (3.7) minden $I_{l_1,l_2,...,l_n}$ típusú intervallumra fennáll.

Legyenek most $b_1, \ldots, b_n \in \mathbb{Q}$ racionális számok, és $b_n = \frac{k_i}{l_i}$. Látható, hogy ha $J = [0, b_1) \times [0, b_2] \times \ldots \times [0, b_n)$, akkor J előáll mint az $I_{l_1,l_2,...,l_n}$ intervallumnak $k_1 k_2 \cdots k_n$ darab diszjunkt eltoltja. Ilyen módon

$$\mu(J) = k_1 k_2 \cdots k_n \mu(I_{l_1,...,l_n}) = k_1 k_2 \cdots k_n c\lambda(I_{l_1,...,l_n}) = c\lambda(J).$$

Így (3.7) mindent fenti típusú J intervallum mellett is fennáll.

Ha most $L = [a_1, b_2) \times [a_2, b_2) \times \ldots \times [a_n, b_n)$ alakú, ahol a végpontok rendre racionális számok, akkor

$$L = (a_1, a_2, \ldots, a_n) + [0, b_1 - a_1) \times [0, b_2 - a_2] \times \ldots \times [0, b_n - a_n)$$

hiszen az összeadást koordinátánként kell elvégezni. Tehát a mértékek eltolás-invarianciája szerint

$$\begin{aligned}\mu(L) &= \mu([0, b_1 - a_1) \times [0, b_2 - a_2] \times \ldots \times [0, b_n - a_n)) = \\ &= c\lambda([0, b_1 - a_1) \times [0, b_2 - a_2] \times \ldots \times [0, b_n - a_n)) = c\lambda(L).\end{aligned}$$

Ezt kellett belátni. □

3.3.6. állítás. *Legyen $T : \mathbb{R}^n \to \mathbb{R}^n$ reguláris lineáris transzformáció. Ekkor minden A Borel-mérhető halmaz mellett a $T(A)$ képhalmaz is Borel-mérhető, és*

$$\lambda(T(A)) = |\det(T)|\lambda(A).$$

Bizonyítás. Mivel véges dimenziós téren értelmezett lineáris operáció mindig folytonos, ezért a T^{-1} inverz leképezés folytonos lévén Borel-mérhető is. Ez éppen azt jelenti, hogy minden $A \in \mathcal{B}$ Borel-halmaz esetén a $T(A) = \left(T^{-1}\right)^{-1}(A)$ halmaz is Borel-halmaz marad.

Most definiálja reguláris T lineáris transzformációra

$$\mu_T(A) = \lambda(T(A)).$$

A fentiek szerint μ_T értelmezett a Borel-halmazokon. Látható, hogy μ_T eltolás-invariáns Borel-mérték. Ugyanis, injektiv függvény diszjunkt halmazokat diszjunkt halmazokra képez, ezért

$$\mu_T\left(\cup_{n=1}^{\infty} A_n\right) = \lambda\left(T\left(\cup_{n=1}^{\infty} A_n\right)\right) = \lambda\left(\cup_{n=1}^{\infty} T\left(A_n\right)\right) =$$
$$= \sum_{n=1}^{\infty} \lambda\left(T\left(A_n\right)\right) = \sum_{n=1}^{\infty} \mu_T\left(A_n\right)$$

egy Borel-mérték. A linearitás, valamint a Lebesgue-mérték eltolás-invarianciája miatt

$$\mu_T\left(A+x\right) = \lambda\left(T\left(A+x\right)\right) = \lambda\left(T\left(A\right)+T\left(x\right)\right) = \lambda\left(T\left(A\right)\right) = \mu_T\left(A\right)$$

is teljesül minden A Borel-halmazra és minden $x \in \mathbb{R}^n$ vektorra.

Az eltolás-invariáns Borel-mértékek unicitási tétele miatt minden T reguláris lineáris transzformációhoz létezik egyetlen $c_T \geq 0$ valós szám, amelyre

$$\mu_T = c_T \cdot \lambda$$

Ha T és S két ilyen lineáris transzformáció, akkor minden A Borel-halmazra

$$c_{TS} \cdot \lambda\left(A\right) = \mu_{TS}\left(A\right) = \lambda\left(T\left(S\left(A\right)\right)\right) = c_T \cdot \lambda\left(S\left(A\right)\right) = c_T \cdot c_S \cdot \lambda\left(A\right)$$

Ez csak úgy lehetséges, ha

$$c_{TS} = c_T \cdot c_S$$

teljesül.

Számoljuk ki a c_T számot egy ortogonális T transzformáció esetében. Egy ortogonális transzformáció a $T^* = T^{-1}$ azonosság miatt normatartó is:

$$\|T^*x\|^2 = \langle T^*x; T^*x\rangle = \langle x; TT^*x\rangle = \langle x; x\rangle = \|x\|^2.$$

Ha tehát B jelöli a tér egységgömbjét, akkor $T\left(B\right) \subseteq B$ azonnal adódik. Persze minden $x \in B$ mellett $x = T\left(T^*x\right)$, amiből $T\left(B\right) = B$ következik. Így $\mu_T\left(B\right) = \lambda\left(T\left(B\right)\right) = \lambda\left(B\right)$, ergo

$$c_T = 1$$

teljesül minden T ortogonális lineáris transzformáció mellett.

Számoljuk ki a c_T számot egy olyan T lineáris transzformáció esetében, amelynek valamely ortonormált bázisban felírt mátrixa diagonális, és a diagonálisban $d_1, d_2, \ldots, d_n$ pozitív számok állnak. Jelölje $Q = [0{,}1)^n$ kockát, ahol n a tér dimenziója. Ezt a kockát T a

$$[0, d_1) \times [0, d_2) \times \ldots [0, d_n)$$

intervallumba képezi, melynek Lebesgue-mértéke $d_1 d_2 \ldots d_n$. Így

$$d_1 d_2 \ldots d_n = \lambda\left(T\left(Q\right)\right) = \mu_T\left(Q\right) = c_T \lambda\left(Q\right) = c_T.$$

Ha tehát a T lineáris transzformáció szinguláris felbontása $T = USV$, ahol U, V ortogonális transzformáció, és V mátrixa diagonális a $\sigma_1, \ldots, \sigma_n$ pozitív számokkal, akkor figyelembe véve, hogy egy ortogonális transzformáció determinánsa csak 1 vagy -1 lehet,

$$|\det T| = |\det(USV)| = |\det U||\det S||\det V| = |\det S| =$$
$$= |\sigma_1 \cdots \sigma_n| = \sigma_1 \cdots \sigma_n = c_S = c_U c_S c_V = c_{USV} = c_T.$$

Ezt kellett belátni. □

3.3.7. állítás. *Legyen S az $\mathbb{R}^n$ vektortér egy n-nél alacsonyabb dimenziós altere. Ekkor az S halmaz n-dimenziós Lebesgue-mértéke zérus.*

Bizonyítás. Először azt gondoljuk meg, hogy az

$$M = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_n = 0\}$$

$n-1$-dimenziós altér n-dimenziós Lebesgue-mértéke zérus. Jelölje $Q_k = [0,1)^{n-1} \times [0,1/k)$ és $Q = [0,1)^{n-1} \times \{0\}$ n-szeres szorzatotokat. Világos, hogy minden $k \in \mathbb{N}$ mellett a Q_k halmaz n-dimenziós mértéke $1/k$ és

$$\cap_{k=1}^{\infty} Q_k = Q.$$

Így a mérték monoton folytonossági tétele miatt Q mértéke zérus. Mivel M előáll mint a Q nullmértékű halmaz megszámlálhatóan sok eltoltjának egyesítéseiként, ezért az M altér n-dimenziós Lebesgue-mértéke is zérus.

Az S altér mértékékének kiszámítására térve, azt gondoljuk meg, hogy mivel lineáris transzformáció egy bázison tetszőlegesen előírható, ezért létezik egy $T : \mathbb{R}^n \to \mathbb{R}^n$ reguláris lineáris transzformáció, amelyre $T(S) \subseteq M$. Így

$$|\det T|\lambda(S) = \lambda(T(S)) \leq \lambda(M) = 0,$$

amiből $\lambda(S) = 0$ a T regularitása szerint már adódik. □

3.3.8. állítás. *Legyen $T : \mathbb{R}^n \to \mathbb{R}^n$ tetszőleges lineáris operátor, és $L \in \Lambda_{\mathbb{R}^n}$ Lebesgue-mérhető halmaz. Ekkor ennek $T(L)$ képe is Lebesgue-mérhető, valamint az n-dimenziós Lebesgue-mértékére*

$$\lambda(T(L)) = |\det T|\lambda(L).$$

Bizonyítás. Tegyük fel először, hogy T szinguláris, tehát $|\det T| = 0$. Ekkor $T(L)$ egy n-nél alacsonyabb dimenziós altér részhalmazaként, azaz egy az n-dimenziós Lebesgue-mértékre nézve nullmértékű halmaz részhalmazaként, maga is nullmértékű, ezért Lebesgue-mérhető.

Ha T reguláris, akkor vegyünk $A, B \in \mathcal{B}_{\mathbb{R}^n}$ Borel-halmazokat, amelyekre

$$A \subseteq L \subseteq B \quad \text{de} \quad \lambda(B \setminus A) = 0.$$

Világos, hogy ekkor $T(A), T(B) \in \mathcal{B}_{\mathbb{R}^n}$ Borel-halmazok, és

$$T(A) \subseteq T(L) \subseteq T(B).$$

A T injektivitása miatt $T(B) \setminus T(A) = T(B \setminus A)$, így

$$\lambda(T(B) \setminus T(A)) = \lambda(T(B \setminus A)) = |\det T| \cdot \lambda(B \setminus A) = 0.$$

Ez éppen azt jelenti, hogy $T(L)$ is Lebesgue-mérhető, és Lebesgue-mértékére

$$\lambda(T(L)) = \lambda(T(A)) = |\det T| \cdot \lambda(A) = |\det T| \cdot \lambda(L).$$

Ezt kellett belátni. □

A fenti tételben nagyon fontos, hogy azonos dimenziós euklideszi tereknek azonos dimenziós Lebesgue-mértékéről van szó. Nem marad érvényben a tétel, már $\mathbb{R}^2 \to \mathbb{R}$ projekció esetében sem. Például vegyünk egy $E \subseteq \mathbb{R}$ nem Lebesgue-mérhető halmazt. Képezzük az $\{(x,0) : x \in E\}$ szorzatot. Mivel a szorzat valamely egydimenziós altér részhalmaza, ezért annak kétdimenziós Lebesgue-mértéke zérus, ergo kétdimenziós értelemben Lebesgue-mérhető. Természetesen az első változó szerinti merőleges vetület éppen E, amely választásunk szerint egydimenziós értelemben nem Lebesgue-mérhető.

3.3.9. állítás. *Tekintsük az n-dimenziós Lebesgue-mértéket $(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda)$. Legyen $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ Lebesgue-mérhető függvény és $T : \mathbb{R}^n \to \mathbb{R}^n$ egy reguláris lineáris transzformáció. Ekkor az $f \circ T$ kompozíció függvény is Lebesgue-mérhető, továbbá fennáll az*

$$\int_{\mathbb{R}^n} f \, d\lambda = |\det T| \int_{\mathbb{R}^n} f \circ T \, d\lambda \tag{3.8}$$

azonosság feltéve, hogy $f \geq 0$, vagy $f \in L_1(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda)$. Ez utóbbi esetben $f \circ T \in L_1(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda)$ is teljesül.

Bizonyítás. Először a kompozíció függvény Lebesgue-mérhetőségét gondoljuk meg. Ha $B \subseteq \mathbb{R}$ egy Borel-halmaz, akkor f Lebesgue-mérhetősége miatt $f^{-1}(B) \in \Lambda_{\mathbb{R}^n}$. No de a T^{-1} lineáris transzformáció Lebesgue-mérhető halmazt Lebesgue-mérhető halmazra képez, így

$$(f \circ T)^{-1}(B) = T^{-1}\left(f^{-1}(B)\right) \in \Lambda_{\mathbb{R}^n}$$

valóban teljesül.

Legyen most $f = \chi_L$ karakterisztikus függvény, valamely $L \in \Lambda_{\mathbb{R}^n}$ Lebesgue-mérhető halmaz mellett. Ekkor $f \circ T = \chi_{T^{-1}(L)}$, ezért

$$\int_{\mathbb{R}^n} f \circ T \, d\lambda = \lambda\left(T^{-1}(L)\right) = \frac{1}{|\det T|}\lambda(L) = \frac{1}{|\det T|}\int_{\mathbb{R}^n} f \, d\lambda,$$

ami éppen a bizonyítandó állítás egy speciális $f = \chi_L$ Lebesgue-mérhető függvény esetében.

Ha $f \geq 0$ Lebesgue-mérhető függvény, akkor a mérhető függvény alaptétele szerint $f = \sum_{n=1}^{\infty} \alpha_n f_n$ alakú, ahol $\alpha_n \geq 0$ és f_n valamely Lebesgue-mérhető halmazzal képzett karakterisztikus függvény minden $n \in \mathbb{N}$ mellett. Így a monoton konvergenciatétel sorokra igazolt alakját kétszer használva kapjuk, hogy

$$\begin{aligned}
&\int_{\mathbb{R}^n} f \, d\lambda = \sum_{n=1}^{\infty} \alpha_n \int_{\mathbb{R}^n} f_n \, d\lambda = \\
&\quad = \sum_{n=1}^{\infty} \alpha_n |\det T| \int_{\mathbb{R}^n} f_n \circ T \, d\lambda = |\det T| \int_{\mathbb{R}^n} \sum_{n=1}^{\infty} \alpha_n \left(f_n \circ T\right) d\lambda = \\
&\qquad = |\det T| \int_{\mathbb{R}^n} \left(\sum_{n=1}^{\infty} \alpha_n f_n\right) \circ T \, d\lambda = |\det T| \int_{\mathbb{R}^n} f \circ T \, d\lambda.
\end{aligned}$$

Legyen most $f \in L_1\left(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda\right)$. Ekkor az $|f|$ nem negatív függvényre teljesül a már igazolt azonosság. Így

$$\int_{\mathbb{R}^n} |f| \, d\lambda = |\det T| \int_{\mathbb{R}^n} |f| \circ T \, d\lambda = |\det T| \int_{\mathbb{R}^n} |f \circ T| \, d\lambda.$$

A bal oldal végességéből a jobb oldal végessége következik, ami azt jelenti, hogy $f \circ T \in L_1\left(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda\right)$ valóban fennáll. Most az f függvény f^+ pozitív és f^- negatív részére alkalmazzuk a már igazolt azonosságot.

$$\begin{aligned}
\int_{\mathbb{R}^n} f \, d\lambda &= \int_{\mathbb{R}^n} f^+ \, d\lambda - \int_{\mathbb{R}^n} f^- \, d\lambda = \\
&= |\det T| \left(\int_{\mathbb{R}^n} f^+ \circ T \, d\lambda - \int_{\mathbb{R}^n} f^- \circ T \, d\lambda\right) = \\
&= |\det T| \int_{\mathbb{R}^n} \left(f^+ \circ T\right) - \left(f^- \circ T\right) d\lambda = \\
&= |\det T| \int_{\mathbb{R}^n} \left(f^+ - f^-\right) \circ T \, d\lambda = |\det T| \int_{\mathbb{R}^n} f \circ T \, d\lambda.
\end{aligned}$$

Ezt kellett belátni. $\square$

4. fejezet

Szorzatmérték

A probléma a következő. Tegyük fel, hogy egy kettős integrál kiszámítása a feladatunk. Először is tisztázzuk mit jelent ez. Adott egy $(X, \mathcal{M}, \mu)$ és egy $(Y, \mathcal{N}, \lambda)$ mértéktér, és egy $f : X \times Y \to \mathbb{R}$ függvény. Az

$$\iint_{X,Y} f \, d\mu \, d\lambda = \int_X \int_Y f(x, y) \, \mu(dx) \, \lambda(dy) \tag{4.1}$$

integrálás ebben a sorrendben azt jelenti, hogy rögzített $x \in X$ mellett próbáljuk integrálni az $f(x, \cdot) : Y \to \mathbb{R}$ függvényt. Ha ez sikerült, a kapott értékre mint x függvényére tekintsünk:

$$x \mapsto \int_Y f(x, y) \, \lambda(dy)$$

Ez persze egy $X \to \mathbb{R}$ függvény, így ennek X-en való integrálja (4.1) kettős integrál.

Persze a változókat a fordított sorrendben is rögzíthetjük. Nézzük tehát az

$$\iint_{Y,X} f \, d\lambda \, d\mu = \int_Y \int_X f(x, y) \, \lambda(dy) \, \mu(dx) \tag{4.2}$$

esetet. Ezt úgy értelmezzük, hogy először a rögzített $y \in Y$ változó mellett tekintsük az $f(\cdot, y) : X \to \mathbb{R}$ függvényt. Most ennek X-en számolt integrálját tekintsük mint az y változó függvényét, tehát az

$$y \mapsto \int_X f(x, y) \, \mu(dx)$$

függvényt. Ha ezt a függvényt integráljuk az y változó szerint, akkor kapjuk a (4.2) iterált integrált.

Az első sorrendű iteráció nagyban különbözik a második sorrendű iterációtól, így természetesen merül fel a kérdés, hogy mi a két különböző sorrendű iterált integrál kapcsolata. Persze általánosabban is megfogalmazhatjuk a problémát: Ha egy n változós függvényünk van, akkor a változók rögzítése már $n!$-féleképpen lehetséges. A kérdés, hogy milyen feltételeket kell tennünk ahhoz, hogy az így számolt $n!$ különböző iterált integrál értéke a változók rögzítésének sorrendjétől függetlenül mind-mind azonos legyen.

4.1. A Fubini-tétel

Először is tisztázzuk, hogy a fenti ismertetett iterált integráloknak milyen feltevések mellet van értelme, majd két tételben adunk elegendő feltételt (4.1.14 és 4.1.15) arra, hogy a különböző sorrendű iterált integrálok értéke a sorrendtől független legyen. Az egyszerűség kedvéért a kétváltozós függvények esetére szorítkozunk. Az általános eset a kétváltozós függvényekre vonatkozó eredmény többszöri alkalmazásával könnyen következik.

4.1.1. definíció (szorzat-mérhetőség)**.** Legyen $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető tér és $(X \times Y, \Sigma)$ ezek szorzata, ahol $\Sigma = \sigma(\mathcal{M} \times \mathcal{N})$ a mérhető téglák generálta σ-algebra. Ha adott egy az $X \times Y$ szorzaton értelmezett függvény, akkor ennek mérhetősége, vagy kicsit pontosabban *szorzat-mérhetősége* a fenti Σ σ-algebrára vonatkozó mérhetőséget jelenti.

Tudjuk, hogy az $\mathcal{M}$ és $\mathcal{N}$ σ-algebráknak mint félgyűrűknek az $\mathcal{M} \times \mathcal{N}$ szorzata a mérhető téglák félgyűrűje (1.1.3), így Σ egy félgyűrű által generált σ-algebraként van definiálva.

4.1.2. definíció (szorzattérbeli halmaz szelete)**.** Az $E \subseteq X \times Y$ halmaz és $x \in X$ esetén legyen $E_x = \{y \in Y : (x, y) \in E\}$, valamint hasonlóan $y \in Y$ mellett $E^y = \{x \in X : (x, y) \in E\}$. Az $E_x \subseteq Y$ és $E^y \subseteq X$ halmazokat az E halmaz *x-metszetének* vagy *x-szeletének* nevezzük.

Az itt bevezetett halmaz-operáció nagyon kényelmesen viselkedik a szokásos halmazelméleti műveletekkel kapcsolatosan.

4.1.3.
A definíció egyszerű alkalmazásával kapjuk, hogy

$$\left(\cup_{n=1}^{\infty} A_n\right)_x = \cup_{n=1}^{\infty} (A_n)_x\,; \quad \left(\cap_{n=1}^{\infty} A_n\right)_x = \cap_{n=1}^{\infty} (A_n)_x\,; \quad (E^c)_x = (E_x)^c\,.$$

Az is világos, hogy

$$(A \times B)_x = \begin{cases} B, & \text{ha } x \in A \\ \varnothing, & \text{ha } x \notin A. \end{cases}$$

4.1.4. állítás (mérhető halmaz metszete mérhető)**.** *Szorzat-mérhető halmaz minden metszete mérhető, azaz minden* $F \in \Sigma$ *esetén* $F_x \in \mathcal{N}$ *és* $F^y \in \mathcal{M}$ *minden* $x \in X$ *és* $y \in Y$ *mellett.*

Bizonyítás. Nézzük például az x-szeletekre vonatkozó állítást. Az y-szeletek mérhetősége ezzel analóg módon adódik. Az 1.1.22.-ben ismertetett σ-indukció elvét követjük. Jelölje $\mathcal{H} \subseteq \mathcal{P}(X)$ azt halmazrendszert, amelynek elemeire az állítás igaz, azaz

$$\mathcal{H} = \{E \subseteq X \times Y : E_x \in \mathcal{N}\ \forall x \in X\}.$$

Azt kell belátnunk, hogy

1. $\mathcal{M} \times \mathcal{N} \subseteq \mathcal{H}$;
2. minden $E \in \mathcal{H}$ esetén $E^c \in \mathcal{H}$;
3. ha $E_n \in \mathcal{H}$ egy megszámlálható halmazsorozat, akkor $\cup_{n=1}^{\infty} E_n \in \mathcal{H}$ is teljesül.

Az igazolásokhoz 4.1.3. tulajdonságokat használjuk.
1. Ha $A \times B \in \mathcal{M} \times \mathcal{N}$, akkor $(A \times B)_x$ vagy $\varnothing \in \mathcal{N}$, vagy $B \in \mathcal{N}$, ezért $A \times B \in \mathcal{H}$.
2. Ha $E \in \mathcal{H}$, akkor $E_x \in \mathcal{N}$, ezért $(E^c)_x = (E_x)^c \in \mathcal{N}$, az $\mathcal{N}$ σ-algebra tulajdonsága szerint.
3. Ha az E_n, $n \in \mathbb{N}$ halmazokra $(E_n)_x \in \mathcal{N}$, akkor $(\cup_{n=1}^{\infty} E_n)_x = \cup_{n=1}^{\infty} (E_n)_x$. Az $\mathcal{N}$ halmazrendszer egy σ-algebraként zárt a megszámlálható egyesítésre, így $\cup_{n=1}^{\infty} E_n \in \mathcal{H}$ valóban fennáll. □

Azért, hogy szóhasználatbeli analógiát kapjunk, most függvényekre is bevezetjük a szelet fogalmát. Az f függvény x-szelete voltaképpen a rögzített x pont melletti a második változó szerinti parciális függvény, azaz $f_x = f(x, \cdot)$.

4.1.5. definíció (függvény szelete)**.** Legyen f az $X \times Y$-on értelmezett függvény. Tetszőlegesen rögzített $x \in X$ mellett f_x legyen az az Y-on értelmezett függvény, melyre minden $y \in Y$ mellett $f_x(y) = f(x, y)$. Az f_x függvényt az f függvény x-*szeletének* nevezzük. Hasonlóan, rögzített $y \in Y$ mellett minden $x \in X$-re $f^y(x) = f(x, y)$ az f függvény y-*szelete.*

Világos, hogy a fenti definíció nem új, pusztán egy jelölés bevezetéséről van szó. Bizonyos helyzetben kellemetlen ugyanis az $f_x = f(x, \cdot)$, illetve $f^y = f(\cdot, y)$ pont-jelölés. Szép analógiát kapunk halmazokra és a függvényekre bevezetett szelet-fogalomra, mivel halmaz-szelet és függvény-szelet kommutál az őskép-művelettel.

4.1.6.
Vegyük észre, hogy tetszőleges $V \subseteq \mathbb{R}$ halmazra

$$\left(f^{-1}(V)\right)_x = f_x^{-1}(V),$$

ha f az $X \times Y$ szorzaton értelmezett. Ugyanis $y \in f^{-1}(V)_x$ pontosan akkor, ha $f(x, y) \in V$, ami pontosan akkor teljesül, ha $y \in f^{-1}(x, \cdot)(V) = f_x^{-1}(V)$.

Az is nyilvánvaló, hogy a függvényszelet-operátor is szépen viselkedik az algebrai műveletekkel szemben. Ha f, g az $X \times Y$ szorzaton értelmezett függvények, akkor

$$(\alpha f + \beta g)_x = \alpha f_x + \beta g_x$$

minden $x \in X$ mellett.

Speciálisan, egy szorzattérbeli halmaz karakterisztikus függvényének szelete éppen a szorzathalmaz szeletének karakterisztikus függvénye. Tehát, ha $Q \subseteq X \times Y$ és $x \in X$, akkor

$$(\chi_Q)_x = \chi_{(Q_x)}.$$

4.1.7. állítás (mérhető függvény szelete mérhető)**.** *Legyenek* $(X, \mathcal{M})$ *és* $(Y, \mathcal{N})$ *mérhető terek, valamint* $(X \times Y, \Sigma)$ *ezek szorzata. Tegyük fel, hogy* $f : X \times \times Y \to \overline{\mathbb{R}}$ *egy mérhető függvény. Ekkor minden* $x \in X$ *rögzített érték mellett az* $f_x : Y \to \overline{\mathbb{R}}$ *függvény is mérhető.*

Bizonyítás. Legyen $V \subseteq \overline{\mathbb{R}}$ egy nyílt halmaz. Láttuk az imént, hogy $f_x^{-1}(V) = f^{-1}(V)_x$. No de az f függvény mérhetősége szerint $f^{-1}(V) \in \Sigma$, és a mérhető halmaz szeletének mérhetőségét biztosító 4.1.4. szerint $f^{-1}(V)_x \in \mathcal{N}$. □

Most bevezetjük a mértékterek szorzatának fogalmát. Emlékezzünk arra, hogy a monoton konvergenciatétel egy fontos következményeként (2.2.16) láttuk, hogy mértékek szorzata egy a σ-algebrák szorzatán mint félgyűrűn értelmezett σ-additív halmazfüggvény. Tudjuk, hogy a kiterjesztési eljárás szerint félgyűrűn értelmezett megszámlálhatóan additív halmazfüggvény egyértelműen terjeszthető ki mértékként a generált σ-algebrára (3.1.15), feltéve, hogy a teljes alaptér előáll mint megszámlálhatóan sok félgyűrűbeli, véges mértékű halmaz egyesítése.

4.1.8. definíció (mértékterek szorzata)**.** Legyenek $(X, \mathcal{M}, \mu)$ és $(Y, \mathcal{N}, \lambda)$ σ-véges mértékterek. Definiálja $A \times B \in \mathcal{M} \times \mathcal{N}$ mérhető téglára

$$(\mu \times \lambda)(A \times B) = \mu(A)\lambda(B)$$

az $\mathcal{M} \times \mathcal{N}$ félgyűrűn értelmezett megszámlálhatóan additív halmazfüggvényt (2.2.16). Tekintsük a generált $\Sigma = \sigma(M \times \mathcal{N})$ σ-algebrára való egyértelmű mérték-kiterjesztését. Jelölje ezt $\mu \otimes \lambda$. Az így kapott $(X \times Y, \Sigma, \mu \otimes \lambda)$ mértékteret a *két mértéktér szorzatának* mondjuk.

4.1.9. definíció (Fubini-tulajdonság)**.** Legyenek $(X, \mathcal{M}, \mu)$ és $(Y, \mathcal{N}, \lambda)$ σ-véges mértékterek, valamint $(X \times Y, \Sigma, \mu \otimes \lambda)$ e két mértéktér szorzata. Legyen $f : X \times Y \to \mathbb{R}_+$ egy nem negatív Σ-mérhető függvény. Azt mondjuk, hogy ez a függvény *első változójában Fubini-tulajdonságú*, ha az alább definiált $\varphi : X \to \mathbb{R}_+$

$$\varphi_f(x) = \int_Y f_x \, d\lambda \tag{†}$$

függvény $\mathcal{M}$-mérhető és

$$\int_{X \times Y} f \, d(\mu \otimes \lambda) = \int_X \varphi \, d\mu. \tag{‡}$$

A célunk az, hogy megmutassuk minden nem negatív függvény rendelkezik az itt definiált Fubini-tulajdonsággal.

4.1.10. állítás. *Legyenek $(X, \mathcal{M}, \mu)$ és $(Y, \mathcal{N}, \lambda)$ σ-véges mértékterek, valamint $(X \times Y, \Sigma, \mu \otimes \lambda)$ e két mértéktér szorzata. Jelölje Φ a nem negatív mérhető Fubini-tulajdonságú függvények halmazát. E halmazrendszer rendelkezik az alábbi tulajdonságokkal:*

1. *Minden $f, g \in \Phi$ és $\alpha, \beta \geq 0$ mellett $\alpha f + \beta g \in \Phi$.*
2. *Minden $f_n \in \Phi$, $f_n \leq f_{n+1}$, $n \in \mathbb{N}$ monoton növő függvénysorozat pontonkénti határértékére $\lim_{n\to\infty} f_n \in \Phi$.*
3. *Minden olyan $f_n \in \Phi$, $f_n \geq f_{n+1}$, $n \in \mathbb{N}$ monoton fogyó függvénysorozatra, amelyre $\int_{X \times Y} f_1 \, d(\mu \otimes \lambda) < \infty$ a pontonkénti határértékre $\lim\limits_{n\to\infty} f_n \in \Phi$.*
4. *Minden $Q \in \mathcal{M} \times \mathcal{N}$ esetén $\chi_Q \in \Phi$.*

Bizonyítás. Az alábbiakat kell meggondolnunk:

1. Világos, hogy $\varphi(x) = \int_Y (\alpha f + \beta g)_x d\,\lambda = \alpha \int_Y f_x \, d\lambda + \beta \int_Y g_x \, d\lambda = = \alpha\varphi_f(x) + \beta\varphi_g(x)$, ahol φ_f és φ_g az $f, g \in \Phi$-hez tartozó függvények (†)-nek megfelelően. Persze nem negatív mérhető függvények kúp kombinációja is ilyen, sőt

$$\int_X \varphi \, d\mu = \int_X \alpha\varphi_f + \beta\varphi_g \, d\mu = \alpha \int_X \varphi_f \, d\mu + \beta \int_X \varphi_g \, d\mu =$$
$$= \alpha \int_{X \times Y} f d\,(\mu \otimes \lambda) + \beta \int_{X \times Y} g \, d(\mu \otimes \lambda) = \int_{X \times Y} \alpha f + \beta g \, d(\mu \otimes \lambda).$$

2. Hasonlóan, legyen $\varphi_n(x) = \int_Y (f_n)_x \, d\lambda$ és az $f = \lim_{n\to\infty} f_n$ függvényre $\varphi(x) = \int_Y f_x \, d\lambda$. Persze $\varphi_n \to \varphi$ monoton növekedőleg. Ezért először is φ mérhető, másrészt a monoton konvergenciatételt kétszer alkalmazva

$$\int_X \varphi_n \, d\mu \to \int_X \varphi \, d\mu, \text{ valamint } \int_{X \times Y} f_n \, d(\mu \otimes \lambda) \to \int_{X \times Y} f \, d(\mu \otimes \lambda).$$

Ezzel a 2. állítást be is láttuk, hiszen a fenti konvergens sorozatok az $f_n \in \Phi$ feltétel szerint azonosak, így a határértékük is azonos.

3. A fenti jelöléseket megtartva $\int_{X\times Y} f_1 d(\mu\times\lambda) = \int_X \varphi_1\,d\mu < \infty$. Ez azt jelenti, hogy az (f_n) függvénysorozatnak van $f_1 \in L_1(X\times Y, \Sigma, \mu\otimes\lambda)$ majoránsa, és a (φ_n) függvénysorozatnak van $\varphi_1 \in L_1(X, \mathcal{M}, \mu)$ majoránsa. Ezért kétszer alkalmazva a majorált konvergenciatételt kapjuk, hogy $\lim_{n\to\infty} f_n \in \Phi$ is fennáll.

4. Világos, hogy $\varphi(x) = \int_Y (\chi_Q)_x\,d\lambda = \int_Y \chi_{Q_x}\,d\lambda = \lambda(Q_x) = \lambda(B)\chi_A(x)$, ahol $Q = A\times B \in \mathcal{M}\times\mathcal{N}$ alakú. Így φ egy mérhető függvény konstans szorosaként maga is mérhető, továbbá

$$\int_X \varphi\,d\mu = \int_X \lambda(B)\chi_A\,d\mu = \lambda(B)\mu(A) = (\mu\times\lambda)(A\times B) =$$
$$= (\mu\otimes\lambda)(A\times B) = \int_{X\times Y} \chi_Q\,d(\mu\otimes\lambda). \quad \square$$

4.1.11.
A fenti állítás első két pontjának azonnali következménye, hogy nem csak a véges összeg, hanem a végtelen sorösszeg művelet sem vezet ki a Fubini-tulajdonságú függvények köréből. Ha $f_n \in \Phi$ minden $n \in \mathbb{N}$ mellett, akkor az első tulajdonság szerint

$$s_n = \sum_{k=1}^{n} f_k \in \Phi.$$

Az itt szereplő függvények nem negatívitása miatt $s_n \leq s_{n+1}$ minden $n \in \in \mathbb{N}$ mellett, tehát a második pontban meggondoltak szerint a részletösszeg-sorozat pontonkénti határértékeként adódó végtelen összegre is

$$\sum_{n=1}^{\infty} f_n = \lim_{n\to\infty} s_n \in \Phi.$$

A fenti négy tulajdonságból és a Dynkin-tételből a Fubini-tétel halmazokra vonatkozó alakja már könnyen látszik.

4.1.12. állítás (nem negatív szorzat-mérhető függvény Fubini-tulajdonságú)**.** *Legyenek $(X,\mathcal{M},\mu)$ és $(Y,\mathcal{N},\lambda)$ σ-véges mértékterek, valamint $(X\times Y,\Sigma,\mu\times\lambda)$ e két mértéktér szorzata. Ekkor minden $Q \in \Sigma$ szorzat-mérhető halmaz χ_Q karakterisztikus függvénye Fubini-tulajdonságú.*

Bizonyítás. A tegyük fel, hogy $X = \cup_{n\in\mathbb{N}} X_n$ valamint $Y = \cup_{m\in\mathbb{N}} Y_m$, ahol az egyes $X_n \in \mathcal{M}$ és $Y_m \in \mathcal{N}$ halmazok diszjunktak, mérhetők, valamint $\mu(X_n) < \infty$ és $\lambda(Y_m) < \infty$ minden szóba jövő n és m esetén. Legyen

$$\Omega = \left\{Q \subseteq X\times Y : \chi_{Q\cap(X_n\times Y_m)} \in \Phi \quad \forall n,m\in\mathbb{N}\right\},$$

ahol Φ a Fubini-tulajdonságú függvények halmaza. Megmutatjuk, hogy $\Sigma \subseteq$ $\subseteq \Omega$. Ehhez az 1.1.23. m-indukció módszerét használjuk. Azt kell tehát meggondolnunk, hogy

1. $\mathrm{r}\,(\mathcal{M} \times \mathcal{N}) \subseteq \Omega$
2. Ha $Q_k \in \Omega$ monoton bővülő sorozat, akkor $\cup_{k=1}^{\infty} Q_k \in \Omega$;
3. Ha $Q_k \in \Omega$ monoton szűkülő sorozat, akkor $\cap_{n=1}^{\infty} Q_k \in \Omega$;

Sorjában az indoklások:

1. Ha $Q \in \mathrm{r}\,(\mathcal{M} \times \mathcal{N})$, akkor a gyűrű belső reprezentációja (1.1.12) szerint

$$Q \cap (X_n \times Y_m) = \cup_{k=1}^{N} \left(Q_k \cap (X_n \times Y_m)\right)$$

alakú, ahol a $Q_k \in \mathcal{M} \times \mathcal{N}$ halmazok diszjunktak. Láttuk (4.1.10), hogy $\mathcal{M} \times \mathcal{N}$-beli halmaz karakterisztikus függvénye Fubini-tulajdonságú. No de $\mathcal{M} \times \mathcal{N}$ metszet zárt, hiszen még félgyűrű is, emiatt $\chi_{Q_k \cap (X_n \times Y_m)} \in \Phi$. Azt viszont tudjuk (4.1.10), hogy Φ zárt a véges összegre, emiatt

$$\chi_{Q \cap (X_n \times Y_m)} = \sum_{n=1}^{N} \chi_{Q_n \cap (X_n \times Y_m)} \in \Phi$$

is fennáll, ergo $Q \in \Omega$.

2. Jelölje $Q = \cup_{k=1}^{\infty} Q_k$ a bővülő $Q_k \in \Omega$ halmazok egyesítését.
A $(Q_k \cap (X_n \times Y_m))_{k \in \mathbb{N}}$ sorozat is bővülő tetszőlegesen rögzített $n, m \in \mathbb{N}$ mellett, és egyesítésük a $Q \cap (X_n \times Y_m)$ halmaz. Így

$$\chi_{Q_k \cap (X_n \times Y_m)} \to \chi_{Q \cap (X_n \times Y_m)}$$

monoton növekedőleg pontonként. De Φ zárt monoton növő pontonkénti konvergenciára nézve (4.1.10), ezért $\chi_{Q \cap (X_n \times Y_m)} \in \Phi$. Mivel ez tetszőleges n, m-re elmondható, így $Q \in \Omega$ valóban fennáll.

3. A $Q_k \in \Omega$ szűkülő halmazok esete szinte szó szerint azonos a monoton bővülő esettel. Az egyetlen különbség, hogy ellenőriznünk kell a

$$\int_{X \times Y} \chi_{Q_1 \cap (X_n \times Y_m)}\, d\,(\mu \otimes \lambda) < \infty$$

feltétel teljesülését, mikor a Φ függvényosztály zártságát használjuk a pontonkénti monoton fogyó konvergenciára nézve (4.1.10). Ez viszont a $\mu\,(X_n) < \infty$ és a $\nu\,(Y_m) < \infty$ feltételek következménye.

Azt kaptuk tehát, hogy minden $Q \in \Sigma$ szorzat-mérhető halmaz esetén $\chi_{Q \cap (X_n \times Y_m)} \in \Phi$ tetszőleges $n, m \in \mathbb{N}$ mellett fennáll. No persze

$$Q = \cup_{n,m}^{\infty} \left(Q \cap (X_n \times Y_m)\right)$$

diszjunkt egyesítés, ezért

$$\chi_Q = \sum_{n,m} \chi_{Q \cap (X_n \times Y_m)}.$$

De a Fubini-tulajdonságú függvények halmaza zárt a végtelen összeg képzésére (4.1.11). Igazoltuk tehát, hogy $\chi_Q \in \Phi$ teljesül minden $Q \in \Sigma$ szorzatmérhető halmaz esetén. □

4.1.13.
A fentiekhez hasonló módon $(X, \mathcal{M}, \mu)$ és $(Y, \mathcal{N}, \lambda)$ σ-véges mértékterek mellett most legyen rögzített $y \in Y$-ra

$$\psi(y) = \int_X f^y \, d\mu.$$

A 4.1.9 definícióval analóg módon azt mondhatnánk, hogy szorzaton értelmezett nem negatív mérhető függvény a *második változójában Fubini-tulajdonságú*, ha ψ mérhető az $\mathcal{N}$ σ-algebrára nézve és

$$\int_Y \psi \, d\lambda = \int_{X \times Y} f \, d\,(\mu \otimes \lambda)\,.$$

A Fubini-tulajdonságot taglaló 4.1.10. és annak 4.1.11. következménye szó szerint ismételhető φ helyett ψ függvényre. Hasonlóan a 4.1.12. állítás is szó szerint ismétlődik az új definíció mellett. Így tehát azt kapjuk, hogy minden $Q \in \Sigma$ mérhető halmazra φ egy $\mathcal{M}$-mérhető, ψ egy $\mathcal{N}$-mérhető függvény és

$$\int_X \varphi \, d\mu = \int_{X \times Y} \chi_Q \, d\,(\mu \otimes \lambda) = \int_Y \psi \, d\lambda.$$

Most megmutatjuk, hogy minden nem negatív szorzat-mérhető függvény az első és a második változójában is Fubini-tulajdonságú, ezért ezt az ideiglenesen bevezetett terminológiát a továbbiakban nem használjuk.

4.1.14. tétel (Fubini tétele nem negatív mérhető függvényre)**.** *Legyenek az $(X, \mathcal{M}, \mu)$ és az $(Y, \mathcal{N}, \lambda)$ mértékterek σ-végesek és $f : X \times Y \to \overline{\mathbb{R}}_+$ nem negatív, szorzat-mérhető függvény. Jelölje*

$$\varphi\,(x) = \int_Y f_x \, d\lambda, \quad \psi\,(y) = \int_X f^y \, d\mu.$$

Ekkor a $\varphi : X \to \overline{\mathbb{R}}_+$ függvény mérhető az $\mathcal{M}$- és a $\psi : Y \to \overline{\mathbb{R}}_+$ függvény mérhető az $\mathcal{N}$ σ-algebrára nézve, továbbá az

$$\int_X \varphi \, d\mu = \int_{X \times Y} f \, d\,(\mu \otimes \lambda) = \int_Y \psi \, d\lambda$$

integrálok egyenlősége is fennáll.

Bizonyítás. A mérhető függvények alaptétele (1.3.18) szerint minden nem negatív szorzat-mérhető függvény előáll

$$f = \sum_{n=1}^{\infty} \alpha_n \chi_{Q_n}$$

alakban, ahol $Q_n \in \Sigma$ szorzat-mérhető halmaz, és $\alpha_n \geq 0$ valós. De épp az imént láttuk (4.1.12, 4.1.10 és 4.1.13), hogy minden egyes $\alpha_n \chi_{Q_n}$ mind az első, mind a második változója szerint Fubini-tulajdonságú, és azt is tudjuk, (4.1.11), hogy Fubini-tulajdonságú függvényekből alkotott függvénysor összegfüggvénye is Fubini-tulajdonságú marad. □

Ha tehát egy szorzat-mérhető függvény nem vált előjelet, akkor semmi további feltételt nem kell ellenőriznünk ahhoz, hogy az integrálás sorrendjét szabadon felcserélhessük. Azt kapjuk, hogy bármelyik sorrendű integrálás a szorzat-mérték szerinti integrált adja.

Most nézzük mit mondhatunk előjelváltó függvényekre.

4.1.15. tétel (Fubini tétele L_1-beli függvényre). *Legyenek az* $(X, \mathcal{M}, \mu)$ *és* $(Y, \mathcal{N}, \lambda)$ *mértékterek* σ*-végesek és* $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \lambda)$ *ezek szorzata. Legyen*

$$f \in L_1 \left(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \lambda\right).$$

Ekkor μ*-m.m.* $x \in X$ *esetén* $f_x \in L_1(Y, \mathcal{N}, \lambda)$ *és* λ*-m.m.* $y \in Y$ *mellett* $f^y \in L_1(X, \mathcal{M}, \mu)$*. Bevezetve az*

$$\varphi(x) = \int_Y f_x \, d\lambda, \qquad \psi(y) = \int_X f^y \, d\mu$$

jelöléseket $\varphi \in L_1(X, \mathcal{M}, \mu)$ *valamint* $\psi \in L_1(Y, \mathcal{N}, \lambda)$*; továbbá*

$$\int_X \varphi \, d\mu = \int_{X \times Y} f \, d\left(\mu \otimes \lambda\right) = \int_Y \psi \, d\lambda.$$

Bizonyítás. Jelölje $\varphi^*(x) = \int_Y |f_x| \, d\lambda$. A nem negatív esetre már bizonyított Fubini-tétel szerint φ^* mérhető és

$$\int_X \varphi^* \, d\lambda = \int_{X \times Y} |f| \, d(\mu \otimes \lambda) < \infty.$$

Ekkor tehát a tétel feltétele szerint $\varphi^* \in L_1(X, \mathcal{M}, \mu)$, ezért μ-m.m. $x \in X$ mellett

$$\varphi^*(x) < \infty.$$

Ez éppen azt jelenti, hogy μ-m.m. $x \in X$ mellett $f_x \in L_1(Y, \mathcal{N}, \lambda)$, ezért a tétel kimondásában definiált φ függvény majdnem minden $x \in X$ mellett értelmes. A többi pontban az értéke legyen nulla. Definiálja most

$$\varphi_1(x) = \int_Y \left(f^+\right)_x \, d\lambda \quad \text{és} \quad \varphi_2(x) = \int_Y \left(f^-\right)_x \, d\lambda.$$

A nem negatív esetre vonatkozó tétel miatt φ_1 és φ_2 mérhető függvények és a szorzatmérték szerinti L_1-feltétel szerint

$$\int_X \varphi_1 \, d\mu = \int_{X \times Y} f^+ \, d(\mu \otimes \lambda) < \infty,$$

valamint

$$\int_X \varphi_2 \, d\mu = \int_{X \times Y} f^- \, d(\mu \otimes \lambda) < \infty.$$

Ezek szerint $\varphi_1, \varphi_2 \in L_1(X, \mathcal{M}, \mu)$. No de μ-majdnem minden $x \in X$ mellett

$$\varphi_1(x) - \varphi_2(x) = \int_X \left(f^+\right)_x - \left(f^-\right)_x \, d\mu = \int_X f_x \, d\mu = \varphi(x).$$

Ez azt jelenti, hogy φ két mérhető függvény különbségeként valóban mérhető, sőt két $L_1(X, \mathcal{M}, \mu)$-beli függvény különbségeként valóban az $L_1(X, \mathcal{M}, \mu)$ vektortér egy elemét reprezentálja. Igaz továbbá az is, hogy

$$\begin{aligned}\int_X \varphi \, d\mu &= \int_X \varphi_1 \, d\mu - \int_X \varphi_2 \, d\mu = \\ &= \int_{X \times Y} f^+ \, d(\mu \otimes \lambda) - \int_{X \times Y} f^- \, d(\mu \otimes \lambda) = \int_{X \times Y} f \, d(\mu \otimes \lambda).\end{aligned}$$

A ψ-re vonatkozó állítás igazolása az eddigiekkel analóg módon folyik. $\square$

Ahhoz, hogy a fenti tételt alkalmazzuk, azt kell tehát ellenőriznünk, a függvény abszolút értékének a szorzatmérték szerinti integrálja véges legyen. Persze a szorzatmérték szerint nincs integrálási technikánk, viszont a függvény abszolút értéke nyilvánvalóan nem negatív, így az integrált a nem negatív függvényekre vonatkozó Fubini-tétel (4.1.14) szerint az integrálás tetszőleges sorrendjében becsülhetjük.

Azt mutattuk meg tehát, hogy amennyiben egy kettős integrál kiszámítására van szükségünk, akkor a legegyszerűbb esetben két tételre kell emlékeznünk.

1. Ha a szorzat-mérhető függvényünk nem vált előjelet, akkor tetszőleges sorrendben iterálhatjuk az integrálást, az eredmény ugyan az lesz, nevezetesen a szorzat-mérték szerinti integrál.

2. Ha a kérdéses függvény nem-negativitását vagy nem-pozitivitását nem lehet garantálni, akkor a másik eszközünk az L_1-beli függvényekre vonatkozó Fubini-tétel (4.1.15). Ennek használatához elegendő az egyik sorrendben becsülni az iterált integrálokat. Ha az egyik sorrenddel sikerül a függvény abszolút értéke iterált integráljának végességét biztosítani, akkor nem csak a függvény abszolút értékének, hanem magának az eredeti függvénynek a szorzatmérték szerinti integrálja is azonos akár melyik sorrendű iterált integrál értékével, ergo az iterált integrálok értéke független az iterálás sorrendjétől.

Most nézzünk egy olyan példát, mikor a kettős integrál értéke függ az iterálás sorrendjétől.

4.1.16. (ellenpélda a Fubini-tételre)
Legyen $(\delta_n)_{n\in\mathbb{N}}$ olyan szigorúan monton növő valós sorozat, amelyre $\delta_1 = 0$ és $\delta_n \to 1$. A $g_n : [0,1] \to \mathbb{R}$ folytonos függvények legyenek úgy megadva, hogy $\{x \in \mathbb{R} : g(x) \neq 0\} \subseteq (\delta_n, \delta_{n+1})$ és $\int_0^1 g_n = 1$. Legyen $f : [0,1] \times [0,1] \to \mathbb{R}$ a következő függvény:

$$f(x,y) = \sum_{n=1}^{\infty} (g_n(x) - g_{n+1}(x))\, g_n(y).$$

Azt mutatjuk meg, hogy

$$\int_0^1 \left(\int_0^1 f(x,y)\, dy \right) dx = 1 \neq 0 = \int_0^1 \left(\int_0^1 f(x,y)\, dx \right) dy$$

Világos, hogy adott y legfeljebb csak az egyik (δ_n, δ_{n+1}) intervallumocskába eshet, ezért az összegnek csak egy tagja nem nulla, így az f folytonos az $(1,1)$ ponttól eltekintve az egész egységnégyzeten.

Bizonyítás. Kezdjük a bal oldallal: Rögzített $x \in [0,1]$ mellett:

$$\int_0^1 f(x,y)\, dy = \int_{\cup_{i=1}^{\infty}(\delta_i,\delta_{i+1})} f(x,y)\, dy = \sum_{i=1}^{\infty} \int_{\delta_i}^{\delta_{i+1}} f(x,y)\, dy.$$

Viszont a (δ_i, δ_{i+1}) intervallumon kívül a g_i függvény 0, így az f-nek egyedül az i-edik tagja jön szóba:

$$\begin{aligned}
\int_0^1 f(x,y)\, dy &= \sum_{i=1}^{\infty} \int_{\delta_i}^{\delta_{i+1}} (g_i(x) - g_{i+1}(x))\, g_i(y)\, dy = \\
&= \sum_{i=1}^{\infty} \left((g_i(x) - g_{i+1}(x)) \int_{\delta_i}^{\delta_{i+1}} g_i(y)\, dy \right) = \\
&= \sum_{i=1}^{\infty} (g_i(x) - g_{i+1}(x)) = \lim_{N\to\infty} (g_1(x) - g_N(x)) = g_1(x).
\end{aligned}$$

Azt kaptuk tehát, hogy

$$\int_0^1 \left(\int_0^1 f(x,y)\, dy \right) dx = \int_0^1 g_1(x)\, dx = 1.$$

Most nézzük a jobb oldalt: Rögzített $y \in [0,1]$ mellett az első intervallumot különírva:

$$\int_0^1 f(x,y)\, dx = \int_{\delta_1}^{\delta_2} f(x,y)\, dx + \sum_{i=2}^{\infty} \int_{\delta_i}^{\delta_{i+1}} f(x,y)\, dx.$$

Most azt vegyük észre, hogy a (δ_1, δ_2) intervallumon az f-nek csak az első tagjának első fele nem zérus, míg az i-edik intervallumon csak az $i-1$-edik tag második felét és az i-edik tag első felét kell figyelembe venni. Tehát:

$$\int_{\delta_1}^{\delta_2} f(x,y)\, dx = \int_{\delta_1}^{\delta_2} g_1(x)\, g_1(y)\, dx = g_1(y) \int_{\delta_1}^{\delta_2} g_1(x)\, dx = g_1(y),$$

valamint

$$\begin{aligned}\sum_{i=2}^{\infty} \int_{\delta_i}^{\delta_{i+1}} f(x,y)\, dx &= \sum_{i=2}^{\infty} \int_{\delta_i}^{\delta_{i+1}} -g_i(x)\, g_{i-1}(y) + g_i(x)\, g_i(y)\, dx = \\ &= \sum_{i=2}^{\infty} \left(g_i(y) - g_{i-1}(y)\right) \int_{\delta_i}^{\delta_{i+1}} g_i(x)\, dx = \\ &= \sum_{i=2}^{\infty} (g_i(y) - g_{i-1}(y)) = \lim_{n\to\infty} g_n(y) - g_1(y) = -g_1(y).\end{aligned}$$

Összefoglalva:

$$\int_0^1 \left(\int_0^1 f(x,y)\, dx \right) dy = \int_0^1 g_1(y) - g_1(y)\, dy = \int_0^1 0 dy = 0.$$

Ezt kellett megmutatni. □

Az előbbiek szerint kiterjesztési eljárás nélkül is definiálható lenne az n-dimenziós ($n > 1$) Lebesgue-mérték a Fubini-tétel segítségével a következő két lépésben. Az első lépés a halmazokra vonatkozó Fubini-tétel, amivel definiáltuk az egydimenziós Lebesgue-mérték szorzatát az $\mathbb{R}^n$ Borel-halmazaira. Második lépésként teljessé tesszük az így kapott mértékteret. Így a Borel-halmazoknak e mérték szerinti teljessé tétele az n-dimenziós Lebesgue-mérhető halmazok halmaza, és a teljessé tett mérték, az n-dimenziós Lebesgue-mérték.

4.1.17.
Fontos látnunk, hogy mit jelentenek a bizonyított tételek $\mathbb{R}^n \to \overline{\mathbb{R}}$ Borel-mérhető függvények speciális esetében. Látható, hogy az $\left(\mathbb{R}^s, \mathcal{B}_{\mathbb{R}^s}, \lambda^{(s)}\right)$ és az $\left(\mathbb{R}^t, \mathcal{B}_{\mathbb{R}^t}, \lambda^{(t)}\right)$ Borel-mértékterek szorzata éppen az $n = s + t$-dimenziós $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \lambda^{(n)}\right)$ Borel-mértéktér. Tudjuk ugyanis (1.1.29), hogy $\mathcal{B}_{\mathbb{R}^s} \otimes \mathcal{B}_{\mathbb{R}^t} = \mathcal{B}_{\mathbb{R}^n}$, valamint ha $\mu = \lambda^{(s)} \otimes \lambda^{(t)}$ a szorzatmérték, akkor μ a balról zárt, jobbról nyílt intervallumokon mint egy s- és egy t-dimenziós tégla szorzatán azonos az n-dimenziós $\lambda^{(n)}$ Lebesgue-mértékkel. A kiterjesztési eljárás egyértelműsége miatt a két mérték a generált σ-algebrán, azaz $\mathcal{B}_{\mathbb{R}^n}$-en is egybeesik.

Meggondoltuk tehát, hogy a szakasz két fő tétele (4.1.14, 4.1.15) minden további nélkül alkalmazható az

$$
\begin{aligned}
(X, \mathcal{M}, \mu) &= \left(\mathbb{R}^s, \mathcal{B}_{\mathbb{R}^s}, \lambda^{(s)} | \mathcal{B}_{\mathbb{R}^s}\right), \\
(X, \mathcal{N}, \lambda) &= \left(\mathbb{R}^t, \mathcal{B}_{\mathbb{R}^t}, \lambda^{(t)} | \mathcal{B}_{\mathbb{R}^t}\right), \\
(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \lambda) &= \left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \lambda^{(n)} | \mathcal{B}_{\mathbb{R}^n}\right)
\end{aligned}
$$

speciális esetben. A tételek feltételei mellett tehát Borel-mérhető függvényekből képzett Lebesgue-mérték szerinti iterált integrál függvények is Borel-mérhetők, az iterált integrálok felcserélhetők, az értékük megegyezik az eredeti Borel-mérhető függvénynek a Lebesgue-mérték szerinti integráljával.

4.2. A Fubini-tétel kiterjesztése teljes mértéktérre

Természetesen merül fel a gondolat, hogy alkalmazható-e a Fubini-tétel $\mathbb{R}^n \to \overline{\mathbb{R}}$ Lebesgue-mérhető függvények esetén is. Az iménti (4.1.17) gondolat bizonyosan nem ismételhető Lebesgue-mérhető függvények esetére, mivel tudjuk, hogy a

$$\mathcal{B}_{\mathbb{R}^s} \otimes \mathcal{B}_{\mathbb{R}^t} = \mathcal{B}_{\mathbb{R}^n}$$

egyenlőséggel (1.1.29) ellentétben a Lebesgue σ-algebrák szorzatára csak a

$$\Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t} \subseteq \Lambda_{\mathbb{R}^n}$$

szigorú tartalmazás teljesül (3.6), ahol $n = s + t$. Semmi okunk feltételezni, hogy egy n-dimenziós értelemben Lebesgue-mérhető függvény első s változóját rögzítve a maradék t változó szerinti parciális függvény t-dimenziós értelemben Lebesgue-mérhető lenne. Emiatt az iterált integrál felírása sem lehetséges.

A megoldás kulcsa abban rejlik, hogy a fent kiemelt bal oldali σ-algebrának a Lebesgue-mértékre vonatkozó teljessé tétele a jobb oldali σ-algebra (3.2.24).

Mivel a teljessé tett σ-algebrára vonatkozó mérhetőséget vissza tudjuk vezetni az eredeti σ-algebrára vonatkozó mérhetőségre (2.2.32), ezért lehetőségünk van a Fubini-tételt is megfogalmazni teljes σ-algebrák szorzatának teljessé tételére.

Először tisztázni kell a mérhető függvények szeleteinek mérhetőségét biztosító állítást a teljes mérték terek szorzatának teljessé tételére.

4.2.1. lemma. *Legyenek $(X, \mathcal{M}, \mu)$ és $(Y, \mathcal{N}, \lambda)$ mértékterek σ-végesek és teljesek. Jelölje, mint eddig is, $(X{\times}Y, \Sigma, \mu{\otimes}\nu)$ ezek szorzatát és $(X{\times}Y, \bar{\Sigma}, \mu\bar{\otimes}\nu)$ a szorzat teljessé tételét. Legyen $h : X \times Y \to \overline{\mathbb{R}}$ egy $\bar{\Sigma}$-mérhető függvény, amely a $\mu \bar{\otimes} \lambda$ mértékre nézve majdnem mindenütt zérus.*

Ekkor a μ mértékre nézve majdnem minden $x \in X$ pontban a h_x metszet függvény mérhető az $\mathcal{N}$ σ-algebrára nézve, és az ilyen x értékekre $h_x = 0$ a λ mértékre nézve majdnem mindenütt.

Hasonlóan, a λ mértékre nézve majdnem minden $y \in Y$ pontban a h^y metszet függvény mérhető az $\mathcal{M}$ σ-algebrára nézve, és az ilyen y értékek mellett $h^y = 0$ a μ mértékre nézve majdnem mindenütt.

Bizonyítás. Jelölje

$$P = \{(x, y) \in X \times Y : h(x, y) \neq 0\}.$$

Világos, hogy $P \in \bar{\Sigma}$ a h mérhetősége szerint és $\left(\mu \bar{\otimes} \lambda\right)(P) = 0$. A teljessé tétel konstrukciója szerint létezik $Q \in \Sigma$ halmaz, amelyre $P \subseteq Q$ és $(\mu \otimes \lambda)(Q) = 0$. Alkalmazzuk most a χ_Q karakterisztikus függvényre a nemnegatív függvényekre vonatkozó 4.1.14. Fubini-tételt. Azt kapjuk, hogy az $x \mapsto \lambda(Q_x)$ függvény $\mathcal{M}$-mérhető, és fennáll az

$$\int_X \lambda(Q_x)\, d\mu = \int_{X \times Y} \chi_Q \, d(\mu \otimes \lambda) = 0.$$

egyenlőség. Mivel egy nem negatív mérhető függvény integrálja csak úgy lehet zérus, ha az integrandus majdnem mindenütt zérus (2.2.21), ezért az

$$A = \{x \in X : \lambda(Q_x) > 0\}$$

halmaz $\mathcal{M}$-mérhető, és $\mu(A) = 0$. Ha $x \notin A$, akkor $\lambda(Q_x) = 0$, így $P_x \subseteq Q_x$ és az $(Y, \mathcal{N}, \lambda)$ mértéktér teljessége miatt $P_x \in \mathcal{N}$. Az eddigieket összefoglalva: Találtunk egy $A \in \mathcal{M}, \mu(A) = 0$ halmazt, amelyre minden $x \notin A$ esetén a $h_x : Y \to \overline{\mathbb{R}}$ függvény a $P_x \in \mathcal{N}, \lambda(P_x) = 0$ halmazon kívül konstans zérus. Mivel $(Y, \mathcal{N}, \lambda)$ egy teljes mértéktér, ezért az ilyen x pontokra h_x $\mathcal{N}$-mérhető (2.2.29), és persze λ-ra nézve majdnem mindenütt zérus. Beláttuk tehát az első állítást. A másik állítás igazolása ezzel analóg. □

4.2.2. tétel (Fubini tétele nem negatív mérhető függvényre és teljes mértéktérre)**.** *Legyenek az $(X, \mathcal{M}, \mu)$ és az $(Y, \mathcal{N}, \lambda)$ teljes mértékterek σ-végesek.*

Jelölje e mértékterek szorzatának teljessé tételét $\left(X \times Y, \bar{\Sigma}, \mu \mathbin{\bar{\otimes}} \lambda\right)$. *Legyen* $f : X \times Y \to \overline{\mathbb{R}}_+$ *egy nem negatív,* $\bar{\Sigma}$*-mérhető függvény. Ekkor* f_x μ*-majdnem minden* $x \in X$ *mellett* $\mathcal{N}$*-mérhető,* f^y λ*-majdnem minden* $y \in Y$ *mellett* $\mathcal{M}$*-mérhető. Az ilyen* x *és* y *értékekre legyenek*

$$\varphi(x) = \int_Y f_x \, d\lambda, \quad \psi(y) = \int_X f^y \, d\mu.$$

Ekkor a $\varphi : X \to \overline{\mathbb{R}}_+$ *függvény mérhető az* $\mathcal{M}$*, és a* $\psi : Y \to \overline{\mathbb{R}}_+$ *függvény mérhető az* $\mathcal{N}$ σ*-algebrára nézve, továbbá az*

$$\int_X \varphi \, d\mu = \int_{X \times Y} f \, d\left(\mu \mathbin{\bar{\otimes}} \lambda\right) = \int_Y \psi \, d\lambda$$

integrálok egyenlősége is fennáll.

Bizonyítás. Láttuk (2.2.32), hogy létezik $g \geq 0$ az Σ szorzat σ-algebrára nézve mérhető függvény, és létezik h az $\bar{\Sigma}$-mérhető függvény, amelyre

$$f = g + h \quad \text{és} \quad h = 0 \, [\mu \mathbin{\bar{\otimes}} \lambda].$$

Alkalmazzuk g-re a nem negatív mérhető függvényekre vonatkozó 4.1.14. Fubini-tételt. Így φ_g egy $\mathcal{M}$-mérhető függvény és

$$\int_X \varphi_g \, d\mu = \int_{X \times Y} g \, d(\mu \otimes \lambda).$$

Mivel az iménti 4.2.1. lemma szerint μ-majdnem minden $x \in X$ pontban a h_x függvény $\mathcal{N}$-mérhető, ezért az ilyen x-ek mellett az $f_x = g_x + h_x$ két mérhető függvény összegeként maga is mérhető. Persze $h_x = 0 \, [\lambda]$ miatt továbbra is az ilyen x értékekre

$$\varphi_f(x) = \int_Y f_x \, d\lambda = \int_Y g_x + h_x \, d\lambda = \int_Y g_x \, d\lambda = \varphi_g(x).$$

Így a tétel kimondásában definiált φ_f függvény egy nullmértékű halmaz komplementerén megegyezik a φ_g, egy a Fubini-tétel szerint az $\mathcal{M}$-re nézve mérhető függvénnyel. Az $(X, \mathcal{M}, \mu)$ mértéktér teljessége (2.2.29) miatt φ_f valóban $\mathcal{M}$-mérhető. Az integrálokra pedig

$$\int_X \varphi_f \, d\lambda = \int_X \varphi_g \, d\lambda = \int_{X \times Y} g \, d(\mu \otimes \lambda) =$$
$$\int_{X \times Y} g \, d\left(\mu \mathbin{\bar{\otimes}} \lambda\right) = \int_{X \times Y} g + h \, d\left(\mu \mathbin{\bar{\otimes}} \lambda\right) = \int_{X \times Y} f \, d\left(\mu \mathbin{\bar{\otimes}} \lambda\right).$$

A bizonyítás a másik változóra ezzel analóg. □

4.2.3. tétel (Fubini tétele L_1-beli függvényre és teljes mértéktérre)**.** *Legyenek az $(X, \mathcal{M}, \mu)$ és $(Y, \mathcal{N}, \lambda)$ mértékterek σ-végesek és teljesek. Jelölje ezek szorzatának teljessé tételét $\left(X \times Y, \bar{\Sigma}, \mu \mathbin{\bar{\otimes}} \lambda\right)$. Legyen*

$$f \in L_1\left(X \times Y, \bar{\Sigma}, \mu \mathbin{\bar{\otimes}} \lambda\right).$$

Ekkor μ-m.m. $x \in X$ esetén $f_x \in L_1(Y, \mathcal{N}, \lambda)$ és λ-m.m. $y \in Y$ mellett $f^y \in L_1(X, \mathcal{M}, \mu)$. Bevezetve az

$$\varphi(x) = \int_Y f_x \, d\lambda, \qquad \psi(y) = \int_X f^y \, d\mu$$

jelöléseket $\varphi \in L_1(X, \mathcal{M}, \mu)$ valamint $\psi \in L_1(Y, \mathcal{N}, \lambda)$; továbbá

$$\int_X \varphi \, d\mu = \int_{X \times Y} f \, d\left(\mu \mathbin{\bar{\otimes}} \lambda\right) = \int_Y \psi \, d\lambda.$$

Bizonyítás. Analóg a szorzatra vonatkozó 4.1.15. tétel indoklásával. □

Térjünk vissza a az $\mathbb{R}^n \to \overline{\mathbb{R}}$ Lebesgue-mérhető függvények esetére.

4.2.4.
Ha $\lambda^{(s)}$ és $\lambda^{(t)}$ jelöli az s- és a t-dimenziós Lebesgue-mértéket, akkor e két mérték szorzata a $\sigma\left(\Lambda_{\mathbb{R}^s} \times \Lambda_{\mathbb{R}^t}\right) = \Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t} \subseteq \Lambda_{\mathbb{R}^n}$ σ-algebrán értelmezett, $n = s + t$, és

$$\lambda^{(s)} \otimes \lambda^{(t)} = \lambda^{(n)}|_{\Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}}.$$

Ugyanis a szorzat definíciója és a Lebesgue-téglák n-dimenziós Lebesgue-mértékére kapott képlet (3.2.22) szerint

$$\left(\lambda^{(s)} \otimes \lambda^{(t)}\right)(A \times B) = \lambda^{(s)}(A)\, \lambda^{(t)}(B) = \lambda^{(n)}(A \times B),$$

tetszőleges $A \in \Lambda_{\mathbb{R}^s}$ és $B \in \Lambda_{\mathbb{R}^t}$ mellett. Az n-dimenziós Lebesgue-mérték ezért a Lebesgue-téglák $\Lambda_{\mathbb{R}^s} \times \Lambda_{\mathbb{R}^t}$ félgyűrűjén egybeesik a szorzatmértékkel. A kiterjesztési tétel egyértelműsége miatt a generált σ-algebrán is azonos e két mérték.

Mivel az $\left(\mathbb{R}^n, \Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}, \lambda^{(n)}|_{\Lambda_{\mathbb{R}^s} \otimes \Lambda_{\mathbb{R}^t}}\right)$ teljessé tétele az n-dimenziós $\left(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda^{(n)}\right)$ Lebesgue-mértéktér, ezért az r- és az s-dimenziós Lebesgue-mértékterek szorzatának teljessé tétele az n-dimenziós Lebesgue-mértéktér.

Meggondoltuk tehát, hogy a szakasz két fő tétele (4.2.2, 4.2.3) minden további nélkül alkalmazható az

$$\begin{aligned}
(X, \mathcal{M}, \mu) &= \left(\mathbb{R}^s, \Lambda_{\mathbb{R}^s}, \lambda^{(s)}\right), \\
(X, \mathcal{N}, \lambda) &= \left(\mathbb{R}^t, \Lambda_{\mathbb{R}^t}, \lambda^{(t)}\right), \\
\left(X \times Y, \mathcal{M} \mathbin{\bar{\otimes}} \mathcal{N}, \mu \mathbin{\bar{\otimes}} \lambda\right) &= \left(\mathbb{R}^n, \Lambda_{\mathbb{R}^n}, \lambda^{(n)}\right)
\end{aligned}$$

teljes mértékterek speciális esetében. A tételek feltételei mellett tehát Lebesgue-mérhető függvényekből képzett Lebesgue-mérték szerinti iterált integrál függvények is majdnem mindenütt Lebesgue-mérhetők, az iterált integrálok felcserélhetők, az értékük megegyezik az eredeti Lebesgue-mérhető függvénynek az n-dimenziós Lebesgue-mérték szerinti integráljával.

5. fejezet

Mértékfelbontási tételek

5.1. Az L_p Lebesgue-terek

5.1.1. állítás (Jensen-egyenlőtlenség)**.** *Legyen* $(\Omega, \sum, \mu)$ *egy valószínűségi mértéktér, azaz* $\mu(\Omega) = 1$. *Ha* $f \in L_1(\Omega, \sum, \mu)$, *és* $\varphi : \mathbb{R} \to \mathbb{R}$, *konvex függvény, akkor*

$$\varphi\left(\int_\Omega f\,d\mu\right) \leq \int_\Omega \varphi \circ f\,d\mu.$$

Bizonyítás. Mivel egy valós konvex függvény az értelmezési tartományának minden belső pontjában folytonos, ezért φ folytonos függvény. Világos, lásd 3.2.6., hogy egy folytonos függvénynek egy Lebesgue-mérhető függvénnyel való kompozíciója is Lebesgue-mérhető, ezért a $\varphi \circ f$ kompozíció függvény Lebesgue-mérhető.

Ismert az is, hogy a φ konvexitása szerint tetszőleges ponthoz tartozó különbségi hányados függvény monoton nő. Ebből következik, hogy minden rögzített $t \in \mathbb{R}$-re a

$$\beta = \inf\left\{\frac{\varphi(s) - \varphi(t)}{s - t} : t < s \in \mathbb{R}\right\}$$

jelölés mellett

$$\varphi(s) \geq \varphi(t) + \beta(s - t)$$

fennáll minden $s \in \mathbb{R}$ esetén. Alkalmazva ezt $s = f(x)$ és $t = \int_\Omega f\,d\mu$ mellett, azt kapjuk, hogy $\varphi(f(x)) \geq \varphi(t) + \beta(f(x) - t)$ amelynek integrálásával:

$$\int_\Omega \varphi \circ f\,d\mu \geq \varphi(t) + \beta(t - t) = \varphi\left(\int_\Omega f\,d\mu\right).$$

Ezt kellett belátni. □

5.1.2. állítás (Hölder-egyenlőtlenség)**.** *Legyen* $1 < p, q < +\infty$, *melyre* $\frac{1}{p} + \frac{1}{q} = 1$, *valamint* $f, g : X \to \mathbb{R}_+$ *nem negatív mérhető függvények. Ekkor*

$$\int_X fg\, d\mu \le \left(\int_X f^p\, d\mu\right)^{\frac{1}{p}} \left(\int_X g^q\, d\mu\right)^{\frac{1}{q}}.$$

5.1.3. állítás (Minkowski-egyenlőtlenség)**.** *Legyen* $1 \le p < \infty$ *valamint* $f, g :$ $: X \to \mathbb{R}_+$ *nem negatív mérhető függvények. Ekkor*

$$\left(\int_X (f+g)^p\, d\mu\right)^{\frac{1}{p}} \le \left(\int_X f^p\, d\mu\right)^{\frac{1}{p}} + \left(\int_X g^p\, d\mu\right)^{\frac{1}{p}}.$$

Bizonyítás. A $p = 1$ esetben az állítás nyilvánvaló, egyébként legyen $q > 1$, amelyre $\frac{1}{p} + \frac{1}{q} = 1$. Tekintsük az $(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}$ felírást. Alkalmazva $f(f+g)^{p-1}$ szorzatra a Hölder-egyenlőtlenséget azt kapjuk, hogy

$$\int_X f(f+g)^{p-1}\, d\mu \le \left(\int_X f^p\, d\mu\right)^{\frac{1}{p}} \cdot \left(\int_X (f+g)^{q(p-1)}\, d\mu\right)^{\frac{1}{q}} =$$
$$= \left(\int_X f^p\, d\mu\right)^{\frac{1}{p}} \cdot \left(\int_X (f+g)^p\, d\mu\right)^{\frac{1}{q}}.$$

Hasonlóan az f helyett g-t és g helyett f-et gondolva azt kapjuk, hogy:

$$\int_X g(g+f)^{p-1}\, d\mu \le \left(\int_X g^p\, d\mu\right)^{\frac{1}{p}} \cdot \left(\int_X (g+f)^p\, d\mu\right)^{\frac{1}{q}}.$$

Összességében tehát

$$\int_X (f+g)^p\, d\mu \le \left(\int_X (f+g)^p\, d\mu\right)^{\frac{1}{q}} \left(\left(\int_X f^p\, d\mu\right)^{\frac{1}{p}} + \left(\int_X g^p\, d\mu\right)^{\frac{1}{p}}\right).$$

Amennyiben $0 < \int_X (f+g)^p\, d\mu < +\infty$, úgy $\left(\int_X (f+g)^p\, d\mu\right)^{\frac{1}{q}}$ számmal osztva készen is vagyunk. Amennyiben $\int_X (f+g)^p\, d\mu = 0$, akkor $f = g = 0$ μ-m.m, tehát nincs mit bizonyítanunk. Amennyiben pedig $\int_X (f+g)^p\, d\mu =$ $= +\infty$ lenne, úgy az $x \mapsto x^p$ függvény konvexitása szerint $\left(\frac{f+g}{2}\right)^p \le \frac{f^p+g^p}{2}$, tehát $(f+g)^p \le 2^{p-1}(f^p + g^p)$. Ez szerint $\int_X f^p\, d\mu$ és $\int g^p\, d\mu$ közül legalább az egyik $+\infty$ értékű. Ezt kellett belátni. □

5.1.4. definíció (p-edik hatványon integrálható függvények: L_p)**.** Legyen $p > 1$ rögzített valós szám, és $(X, \mathcal{M}, \mu)$ mértéktér. Tekintsük a következő vektorteret

$$\mathcal{L}_p(X, \mathcal{M}, \mu) = \left\{ f : X \to \overline{\mathbb{R}} : \ f \text{ mérhető}, \ \int_X |f|^p\, d\mu < +\infty \right\}$$

Világos, hogy $\int_X |f|^p \, d\mu = 0$ pontosan akkor, ha $f(x) = 0$ μ .m.m. $x \in \in X$ esetén. Ha $\mathcal{L}^0$ jelöli azon mérhető függvényeket amelyek a konstans 0 függvénytől csak egy a μ-re nézve null-mértékű halmazban különböznek, akkor világos, hogy $\mathcal{L}^0$ a fenti $\mathcal{L}^p$ térnek altere. Jelölje L_p az $\mathcal{L}_p$ térnek ezen $\mathcal{L}^0$ altere szerinti faktorterét, azaz

$$L_p(X, \mathcal{M}, \mu) = \mathcal{L}_p(X, \mathcal{M}, \mu) / \mathcal{L}^0(X, \mathcal{M}, \mu).$$

Ez azt jelenti, hogy L_p elemei olyan ekvivalenciaosztályok, hogy az egyes ekvivalenciaosztályok olyan mérhető függvényekből állnak, melyek egymástól csak null-mértékű halmazban különböznek. Továbbra is f-fel jelölve azt az ekvivalenciaosztályt, amely az f függvényt tartalmazza

$$\|f\|_{L_p} = \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}.$$

5.1.5. állítás (majdnem mindenütt korlátos függvények: L_∞)**.** *Jelölje $\mathcal{L}_\infty(X, \mathcal{M}, \mu)$ azon f mérhető függvények halmazát, amelyekhez létezik A mérhető, null-mértékű halmaz, hogy $\sup_{X \setminus A} |f| < +\infty$. Világos, hogy $\mathcal{L}_\infty(X, \mathcal{M}, \mu)$ is vektortér, amelynek $\mathcal{L}^0$ egy altere. Az $\mathcal{L}_\infty(X, \mathcal{M}, \mu)$-nek az $\mathcal{L}^0$ szerinti faktorterét nevezzük $L_\infty(X, \mathcal{M}, \mu)$ térnek. Minden $f \in L_\infty$ ekvivalenciaosztály esetén*

$$\inf_{\mu(A)=0} \sup_{X \setminus A} |f| = \inf \{K \geq 0 : \mu(\{x \in X : |f(x)| > K\}) = 0\}.$$

A fenti egyenlőségben mindkét oldalon az infimum helyett minimum is írható és ezt a számot nevezzük az f „függvény” L_∞ normájának. Ez azt jelenti, hogy létezik olyan A mérhető, $\mu(A) = 0$ halmaz, amelyre

$$\|f\|_{L_\infty} = \sup_{X \setminus A} |f| = \min \{K \geq 0 : \mu(\{x \in X : |f(x)| > K\}) = 0\}.$$

Bizonyítás. Jelölje $\kappa = \inf \left\{ \sup_{X \setminus A} |f| : A \in \mathcal{M}, \mu(A) = 0 \right\}$, és legyen K egy szám a jobb oldali halmazból. Megmutatjuk, hogy $\kappa \leq K$. Jelölje ugyanis $A = X(|f| > K)$. Ekkor K választása szerint $\mu(A) = 0$, ezért valóban $\kappa \leq \sup_{X \setminus A} |f| \leq K$. Most azt mutatjuk meg, hogy κ eleme is a jobb oldali halmaznak. Válasszunk ehhez megszámlálhatóan sok $A_n \in \mathcal{M}$ halmazt, amelyekre $\mu(A_n) = 0$ minden $n \in \mathbb{N}$ mellett és

$$\inf_n \left\{ \sup_{X \setminus A_n} |f| \right\} = \kappa.$$

Legyen $A = \cup_n A_n$. Világos, hogy $\mu(A) = 0$, és tetszőleges $x \in A^c = \cap_n A_n^c$ esetén $|f(x)| \leq \sup_{A_n^c} |f|$ teljesül minden $n \in \mathbb{N}$-re, ezért $|f(x)| \leq \kappa$. Ez azt jelenti, hogy

$$X(|f| > \kappa) \subseteq A,$$

amiből persze $\mu(X(|f| > \kappa)) = 0$, tehát κ eleme a tételben kiemelt jobb oldali halmaznak is. Ebből már látszik, hogy a jobb oldali halmaz minimuma éppen κ, sőt a fent konstruált A halmazra $\sup_{A^c} |f| = \kappa$. Ezt kellett belátni. $\square$

5.1.6. állítás. *Legyen tetszőleges $1 \le p, q \le +\infty$, $1/p + 1/q = 1$ mellett $f \in L_p(X, \mathcal{M}, \mu)$, és $g \in L_q(X, \mathcal{M}, \mu)$. Ekkor $fg \in L_1$, és*

$$\|fg\|_{L_1} \le \|f\|_{L_p} \cdot \|g\|_{L_q}.$$

Ha $f, q \in L_p(X, \mathcal{M}, \mu)$, akkor

$$\|f + g\|_{L_p} \le \|f\|_{L_p} + \|g\|_{L_p}.$$

Bizonyítás. Ha $1 < p < \infty$, akkor az első egyenlőtlenség éppen a Hölder-egyenlőtlenség. Ha $p = 1$, akkor $q = \infty$, tehát $K = \|g\|_\infty$ jelöléssel $|g| \le \le K$ majdnem mindenütt teljesül. Ekkor persze $\int_X |fg|\, d\mu \le \int_X |f| K\, d\mu = = K \int_X |f|\, d\mu = \|f\|_p \|g\|_\infty$.

Teljesen hasonlóan, ha $1 < p < \infty$, akkor a második egyenlőtlenség a Minkowski-egyenlőtlenség. Viszont $p = \infty$ esetben f és g majdnem mindenütt korlátosak. Látható, hogy $K = \|f\|_\infty + \|g\|_\infty$ egy olyan szám, amelyre

$$X(|f + g| > K) \subseteq X(|f| > \|f\|_\infty) \cup X(|g| > \|g\|_\infty),$$

ezért $\mu(X(|f + g| > K)) = 0$, amiből már következik is, hogy $\|f + g\|_\infty \le \le K = \|f\|_\infty + \|g\|_\infty$. $\square$

5.1.7. következmény. *Az $\left(L_p(X, \mathcal{M}, \mu), \|\cdot\|_{L_p}\right)$ normált tér, tetszőleges $1 \le p < +\infty$ mellett.*

5.1.8.

A fent bevezetett $L_\infty(X, \mathcal{M}, \mu)$ vektorteret ellátva az

$$\begin{aligned}\|f\|_{L_\infty} &= \min\left\{\sup_{X \setminus A} |f| : A \in \mathcal{M}, \mu(A) = 0\right\} = \\ &= \min\{K \ge 0 : \mu(\{x \in X : |f(x)| > K\}) = 0\}\end{aligned}$$

normával egy normált teret kapunk. Ebben a normált térben $\|f\|_{L_\infty} = 0$, pontosan akkor, ha f majdnem mindenütt zérus, tehát ha f az L_∞ tér nullelemét mint ekvivalenciaosztályt reprezentálja. Látható, hogy ebben a normált térben $f_n \to f$, azaz $\|f_n - f\|_\infty \to 0$, azt jelenti, hogy van olyan $A \in \mathcal{M}$ null-mértékű halmaz, hogy $f_n \to f$ az $X \setminus A$ halmazon egyenletesen.

Most a Lebesgue-tétel sorozatokra és sorokra vonatkozó alakját általánosítjuk.

5.1.9. tétel (Lebesgue)**.** *Legyen* $p \in \mathbb{R}, p \geq 1$ *rögzített. Legyen* $(f_n)_{n\in\mathbb{N}} \in \in L_p(X, \mathcal{M}, \mu)$ *olyan függvények sorozata, amely pontonként konvergál egy* f *függvényhez. Tegyük fel, hogy létezik* $h \in L_p(X, \mathcal{M}, \mu)$ *függvény, melyre* $|f_n| \leq h$ *minden* $n \in \mathbb{N}$ *esetén.*

Ekkor $f \in L_p$*, továbbá* $f_n \to f$ *az* L_p *tér topológiájában, azaz* $\|f - f_n\|_p \to 0$.

Bizonyítás. Az $|f_n| \leq h$ majoráns tulajdonság szerint

$$\int_X |f_n|^p \, d\mu \leq \int_X |h|^p \, d\mu < \infty.$$

Így a $h \in L_p$ feltétel szerint $f_n \in L_p$, minden n mellett. A pontonkénti konvergencia és a háromszög-egyenlőtlenség szerint $|f_n - f| \leq |f_n| + |f| \leq \leq 2h \in L_p$, így $f - f_n \in L_p$, amiből L_p vektortér volta miatt már következik is, hogy $f \in L_p$. Alkalmazzuk most az $(|f - f_n|^p)_{n\in\mathbb{N}}$ függvénysorozatra a majoráns konvergenciatételt. Ez megtehető, hiszen $|f - f_n|^p \leq 2^p h^p \in L_1$. Ekkor $(f - f_n)^p$ sorozat az L_1 térben konvergál zérushoz, ami éppen ugyanaz, mint az $(f - f_n)$ sorozat L_p-beni zéruskonvergenciája, hiszen

$$\|f - f_n\|_p^p = \int_X |f - f_n|^p \, d\mu = \| (f - f_n)^p \|_1.$$

Ezt kellett belátni. □

5.1.10. tétel (Riesz–Fischer-tétel)**.** *Legyen* $p \in \mathbb{R}, p \geq 1$ *rögzített. Tekintsük az* $f_n \in L_p(X, \mathcal{M}, \mu)$ *függvényeket. Ha*

$$\sum_{n=1}^{\infty} \left(\int_X |f_n|^p \, d\mu \right)^{\frac{1}{p}} < \infty,$$

akkor a $\sum_{n=1}^{\infty} f_n$ *függvénysor* μ*-majdnem mindenütt pontonként abszolút konvergens. Ha* f *jelöli a pontonkénti összegfüggvényt, akkor* $f \in L_p$ *és a sor* L_p*-ben is konvergál* f*-hez, azaz*

$$\|f - \sum_{k=1}^{n} f_k\|_{L_p} \to 0$$

midőn $n \to \infty$.

Bizonyítás. Legyen $\varphi = \sum_{n=1}^{\infty} |f_n|$, $M = \sum_{n=1}^{\infty} \|f_n\|_p$ és $\varphi_n = \sum_{k=1}^{n} |f_k|$. Persze $\varphi_n \in L_p$ és $\varphi_n^p \to \varphi^p$ pontonként monoton növekedőleg, így a monoton konvergenciatétel szerint

$$\|\varphi_n\|_p^p = \int_X \varphi_n^p \, d\mu \to \int_X \varphi^p \, d\mu.$$

No de a Minkowski-egyenlőtlenség szerint $\|\varphi_n\|_p \le M$ minden $n \in \mathbb{N}$ mellett, így

$$\int_X \varphi^p \, d\mu \le M^p.$$

Ez egyrészt azt jelenti, hogy $\varphi \in L_p$; másrészt, hogy φ majdnem minden pontban véges értékű, ergo a $\sum_{n=1}^{\infty} f_n(x)$ függvénysorozat majdnem mindenütt abszolút konvergens. A valós egyenes teljessége szerint egy abszolút konvergens sor konvergens is, ezért majdnem minden $x \in X$ mellett a $\sum_{n=1}^{\infty} f_n(x)$ sor konvergens. Jelölje f az összegfüggvényt. Ha s_n a $\sum_k f_k$ sor n-edik részletösszege, akkor $s_n \to f$ majdnem mindenütt pontonként, és $|s_n| \le \varphi \in L_p$. Alkalmazhatjuk tehát az L_p-beni majoráns konvergenciatételt, így

$$f \in L_p, \quad \text{és} \quad \|f - s_n\|_p \to 0.$$

Ezt kellett belátni. □

5.1.11.
Gondoljuk pontosan végig, hogy ha $f_n \in L_\infty$ függvények, amelyekre

$$\sum_{n=1}^{\infty} \|f_n\|_\infty < \infty,$$

akkor a $\sum_{n=1}^{\infty} f_n$ függvénysorozat majdnem mindenütt pontonként konvergens, mi több a részletösszegek sorozata majdnem mindenütt egyenletesen is konvergál az összegfüggvényhez.

5.1.12. állítás (Riesz–Fischer). *Az* $\left(L_p(X, \mathcal{M}, \mu), \|\cdot\|_{L_p}\right)$ *normált tér, tetszőleges* $1 \le p \le +\infty$ *mellett teljes, azaz Banach-tér.*

Bizonyítás. Éppen azt gondoltuk meg az előzőekben, hogy az L_p-térbeli abszolút konvergens sorok a normált tér topológiájában is konvergálnak. Tudjuk – lásd a 7. oldalon lévő 0.1.7. tételt –, hogy ez a tulajdonság ekvivalens a tér teljességével. □

Konstruálhatunk a [0,1] intervallumon értelmezett nem negatív f_n függvénysorozatot, amelyre $\int f_n \, d\mu \to 0$, de $f_n(x) \to 0$ egyetlen $x \in [0,1]$ mellett sem áll fenn. Ez azt jelenti, hogy az L_1-beli konvergenciából a pontonkénti konvergencia nem következhet. Ennél kicsit kevesebb viszont igaz. Ha az $f_n \to f$ az L_p tér metrikája szerint, akkor az f_n függvénysorozatnak van olyan f_{n_k} részsorozata, amelyre $f_{n_k} \to f$ majdnem mindenütt pontonként. A Riesz-Fischer-tétel e következményét, szokás Riesz-lemmának nevezni.

5.1.13. lemma (Riesz). *Legyen* $f \in L_p(X, \mathcal{M}, \mu)$ *egy az* L_p *térben Cauchy-sorozat. Ekkor létezik* $f \in L_p$ *függvény, és létezik* f_{n_k} *részsorozat, amelyre* $\|f_n - f\|_p \to 0$ *és* $f_{n_k}(x) \to f(x)$ *majdnem minden* $x \in X$*-re pontonként.*

Bizonyítás. Az L_p teljessége szerint már igazoltuk, hogy létezik $f \in L_p$ függvény, amelyre $f_n \to f$ az L_p normában. Mivel f_n egy L_p-beli Cauchy-sorozat, ezért választhatunk f_{n_k} részsorozatát, amelyre a

$$\sum_{k=1}^{\infty} \left(f_{n_{k+1}} - f_{n_k}\right)$$

függvénysor az L_p normával abszolút konvergens. A Riesz-Fischer-tétel szerint e sor L_p-ben és majdnem mindenütt pontonként is tart ugyanazon L_p-beli függvényhez. Ez annyit tesz, hogy a részletösszegek s_K sorozatára a konvergencia mind az L_p tér metrikája értelmében, mind a majdnem mindenütt pontonkénti értelemben is fennáll. No de $s_K = f_{n_{K+1}} - f_{n_1}$, tehát az $(f_{n_k})_{k\in\mathbb{N}}$ részsorozat is egyazon L_p-beli függvényhez tart L_p-ben és majdnem mindenütt pontonként. Figyelembe véve, hogy az L_p metrikus téren a sorozat határértéke egyértelműen meghatározott, ez a határfüggvény csak $f \in L_p$ lehetséges. $\square$

5.1.14. definíció. Tetszőleges $1 \le p < +\infty$ mellett jelölje

$$l_p = \left(L_p\left(X, \mathcal{M}, \mu\right), \|\cdot\|_{L_p}\right).$$

Világos, hogy l_p elemei azon $(a_n)_{n\in\mathbb{N}}$ sorozatok, melyekre $\sum_{n=1}^{\infty} |a_n|^p < \infty$, és egy ilyen $a = (a_n)_{n\in\mathbb{N}}$ sorozat normája:

$$\|a\|_{l_p} = \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{\frac{1}{p}}.$$

A $p = +\infty$ esetben l_∞ elemei azon $(a_n)_{n\in\mathbb{N}}$ sorozatok, melyekre $\sup_{n\in\mathbb{N}} |a_n| <$ $< \infty$. Egy ilyen $a = (a_n)_{n\in\mathbb{N}}$ sorozat normája:

$$\|a\|_{l_\infty} = \sup_{n\in\mathbb{N}} |a_n|.$$

Gondoljuk meg, hogy a fejezet elején konvex függvényekre bizonyított Jensen-egyenlőtlenség, valóban az elemi analízisből jól ismert diszkrét Jensen-egyenlőtlenség általánosítása!

5.2. Riesz-reprezentációs tétel Hilbert-térben

A fejezetben csak $\mathbb{R}$ feletti vektorereket vizsgálunk, de könnyű utánagondolni, $\mathbb{C}$ feletti vektorterekre valamennyi állítás igaz marad. Erre akkor lenne szükség, ha $X \to \mathbb{C}$ függvények integrál-elméletével is foglalkoznánk.

5.2.1. definíció (Hilbert-tér)**.** Egy $(X, \langle\cdot,\cdot\rangle)$ skaláris szorzatos teret *Hilbert-térnek* nevezünk, ha a $\|x\| = \sqrt{\langle x, x\rangle}$ normával az $(X, \|\cdot\|)$ normált tér teljes.

5.2.2.
Világos, hogy Az $\left(L_2\left(X, \mathcal{M}, \mu\right), \|\cdot\|_{L_2}\right)$ a Riesz–Fischer-tétel miatt Hilbert-tér, hiszen az

$$\langle f, g\rangle = \int_X f \cdot g \, d\lambda$$

olyan skaláris szorzat, hogy az általa indukált norma, megegyezik $\|\cdot\|_{L_2}$ normával.

5.2.3. állítás. *Legyen $(X, \langle\cdot,\cdot\rangle)$ Hilbert-tér, és $K \subseteq X$ egy nem üres konvex zárt halmaz. Ekkor létezik egyetlen $v \in K$ pont, amelyre minden $u \in K$ mellett $\|v\| \leq \|u\|$.*

Bizonyítás. Legyen $\alpha \in \mathbb{R}_+$, amelyre $\alpha = \inf\{\|u\| : u \in K\}$. Világos, hogy létezik $u_n \in K$ sorozat, amelyre

$$\|u_n\| \to \alpha. \tag{5.1}$$

Megmutatjuk, hogy u_n Cauchy-sorozat:

$$\|u_i - u_j\|^2 + \|u_i + u_j\|^2 = 2\|u_i\|^2 + 2\|u_j\|^2$$

a paralelogramma szabály szerint, így ezt $\frac{1}{2}u_i$ és $\frac{1}{2}u_j$ -re felírva

$$\frac{1}{4}\|u_i - u_j\|^2 = \frac{1}{2}\left(\|u_i\|^2 + \|u_j\|^2\right) - \left\|\frac{u_i + u_j}{2}\right\|^2,$$

ami K konvexitása miatt azt jelenti, hogy

$$\frac{1}{4}\|u_i - u_j\|^2 \leq \frac{1}{2}\left(\|u_i\|^2 + \|u_j\|^2\right) - \alpha^2, \tag{5.2}$$

amiből (5.1) miatt valóban következik, hogy u_n Cauchy-sorozat. A Hilbert-tér teljessége miatt viszont létezik $v \in X$, amelyre

$$u_n \to v,$$

de K zártsága szerint $v \in K$ is teljesül. A normaleképezés folytonossága miatt

$$\|u_n\| \to \|v\|, \quad \text{tehát} \quad \|v\| = \alpha,$$

ami azt jelenti, hogy v valóban minimális hosszúságú pontja K-nak. Az egyértelműség indoklásához legyenek $v_1, v_2 \in K$ olyan pontok, amelyekre

$$\|v_1\| = \alpha = \|v_2\|.$$

Az (5.2)-ot újra számolva azt kapjuk, hogy

$$\frac{1}{4}\|v_1 - v_2\|^2 \leq \frac{1}{2}\left(\|v_1\|^2 + \|v_2\|^2\right) - \alpha^2 = 0$$

tehát $v_1 = v_2$ azaz a legkisebb pont egyértelmű is. □

5.2.4.
Látható, hogy amennyiben M egy zárt altér a Hilbert-térben, akkor tetszőleges $x \in X$ mellett az $x + M$ halmaznak a 0-tól való távolsága azonos az x pontnak az M halmaztól vett távolságával. Ugyanis $u \in x + M$ mellett:

$$\begin{aligned}\|u\| &= \inf\{\|v\| : v \in x + M\} = \\ &= \inf\{\|x + m\| : m \in M\} = \inf\{\|x - m\| : m \in M\} = d(x, M).\end{aligned}$$

5.2.5. állítás. *Legyen* $(X, \langle\cdot,\cdot\rangle)$ *egy Hilbert-tér,* $M \subseteq X$ *pedig egy zárt altér,* $x \in X$ *egy tetszőleges pont. Az előző állítás miatt létezik egyetlen pontja az* M *halmaznak, amely* x*-hez legközelebb van. Ezt a pontot nevezzük az* x-nek az M-re vonatkozó merőleges vetületének *és* $p_M(x)$*-szel jelöljük. Magyarul:*

$$p_M(x) \in M, \quad \textit{amelyre} \ \|x - p_M(x)\| = \inf\{\|x - v\| : v \in M\}.$$

Minden $x \in X$ *mellett az* $x - p_M(x)$ *vektor merőleges* M *altér minden elemére.*

Bizonyítás. Legyen $z = x - p_M(x)$ és $y \in M$ olyan vektor, amelyre $\|y\| = 1$. Ekkor tetszőleges $\alpha \in \mathbb{R}$ valós szám esetén

$$\begin{aligned}\|z\|^2 = \|x - p_M(x)\|^2 &\leq \|x - p_M(x) - \alpha y\|^2 = \\ &= \|z - \alpha y\|^2 = \|z\|^2 - 2\alpha\langle z, y\rangle + \alpha^2.\end{aligned}$$

Válasszuk meg most az α számot $\alpha = \langle z, y\rangle$ módon. Így a fentiek szerint

$$0 \leq -2\alpha^2 + \alpha^2 = -\alpha^2$$

Azt kaptuk tehát, hogy minden $y \in X$ $\|y\| = 1$ esetén $\langle z, y\rangle = 0$. Ezt kellett belátni. □

5.2.6. tétel (Hilbert-tér felbontási tétele (Riesz))**.** *Legyen* $(X, \langle\cdot,\cdot\rangle)$ *egy Hilbert-tér,* $M \subseteq X$ *pedig egy zárt altér. Ekkor*

$$M \oplus M^\perp = X.$$

Bizonyítás. Világos, hogy $M^\perp$ is (zárt) altér, amely diszjunkt M-től. Amennyiben $x \in X$ egy rögzített vektor, akkor

$$x = p_M(x) + (x - p_M(x)).$$

No de $p_M(x) \in M$ a projekció operátor definíciója miatt és $x - p_M(x) \in M^\perp$ is teljesül az előző állítás szerint. □

Emlékezzünk arra, hogy egy normált térek között értelmezett lineáris operátor folytonossága ekvivalens a 0 pontbeli folytonosságával, és azzal, hogy létezzék olyan K szám, melyre $\|Ax\| \leq K\|x\|$ az értelmezési tartomány minden x elemére. Ebben az esetben egy A lineáris operátornak az *operátor normáját* az alábbi ekvivalens módon definiáltuk

$$\begin{aligned} \|A\| &= \sup\{\|Ax\| : x \in X, \|x\| \leq 1\} = \\ &= \min\{K \in \mathbb{R}_+ : \|Ax\| \leq K\|x\|\}. \end{aligned}$$

Láttuk azt is, hogy ezzel valóban normafüggvényt definiáltunk, a két normált tér között definiált lineáris operátorok vektorterén, amely az itt bevezetett normával teljes is, feltéve, hogy a képtérül szolgáló normált tér teljes.

5.2.7. állítás („Példa” Hilbert-téren értelmezett folytonos lineáris funkcionálra)**.** *Legyen $(X, \langle\cdot,\cdot\rangle)$ skaláris szorzatos tér, valamint $y \in X$ egy tetszőleges elem. Ekkor a $\Phi_y(x) = \langle x, y\rangle$ lineáris funkcionál folytonos, és az operátornormájára*

$$\|\Phi_y\| = \|y\|.$$

Bizonyítás. $|\Phi_y(x)| \leq \|y\|$ minden $x \in X$, $\|x\| \leq 1$ mellett fennáll, tehát $\|\Phi_y\| \leq \|y\|$, és $\Phi_y\left(\frac{y}{\|y\|}\right) = \|y\|$, ezért $\|\Phi_y\| = \|y\|$ valóban teljesül. □

Most megmutatjuk, hogy a fenti alakún kívül nincs is más folytonos lineáris funkcionál egy Hilbert-téren

5.2.8. tétel (Folytonos lineáris funkcionálok reprezentációja – Riesz)**.** *Legyen $(X, \|\cdot\|)$ egy Hilbert-tér, valamint Φ egy folytonos lineáris funkcionál X-en. Ekkor létezik egyetlen olyan $y \in X$ vektor, amelyre*

$$\Phi(x) = \langle x, y\rangle$$

minden $x \in X$ esetén.

Bizonyítás. Válasszunk egy $z \in (\ker\Phi)^\perp$ -beli pontot, amely különbözik a 0 vektortól. Mivel $\ker\Phi$ zárt altér ezért $(\ker\Phi)^\perp = \{0\}$ csak úgy lehetne, ha $\Phi = 0$. Tekintsük tetszőleges $x \in X$ mellett a

$$\Phi(x)\,z - \Phi(z)\,x \in \ker\Phi$$

vektort. Világos, hogy $\langle\Phi(x)\,z - \Phi(z)\,x, z\rangle = 0$, ezért $\Phi(x) = \left\langle x, \frac{\Phi(z)}{\|z\|^2} z\right\rangle$.

Az egyértelműséghez, ha y_1 és y_2 ilyen vektorok, akkor tetszőleges $x \in X$ mellett $\langle x, y_1 - y_2\rangle = 0$, ezért speciálisan $x = y_1 - y_2$ mellett is $\|y_1 - y_2\|^2 = 0$. □

5.3. Lebesgue-felbontás és a Radon–Nikodym-tétel

A következő állításokban fontos, hogy a szereplő mértékek negatív halmazt is felvehetnek.

5.3.1. definíció (halmazra koncentrált mérték)**.** Legyen λ egy előjeles mérték az $(X, \mathcal{M})$ mérhető téren és $A \in \mathcal{M}$ mérhető halmaz. Azt mondjuk, hogy λ az A-ra koncentrált mérték, ha

$$\lambda(E) = \lambda(E \cap A)$$

minden $E \in \mathcal{M}$ mellett.

5.3.2.
Amennyiben a λ előjeles mérték B_1-re és B_2-re is koncentrált akkor $B_1 \cap \cap B_2$-re is, hiszen $\lambda(E) = \lambda(E \cap B_1) = \lambda(E \cap B_1 \cap B_2)$. Továbbá ha λ a B halmazra koncentrált és $B \subseteq D$ mérhető halmaz, akkor λ előjeles mérték a D-re is koncentrált, hiszen $\lambda(E \cap D) = \lambda(E \cap D \cap B) = \lambda(E \cap B) = \lambda(E)$.

5.3.3. definíció (szinguláris mérték)**.** Legyen λ_1 és λ_2 előjeles mértékek az $(X, \mathcal{M})$ mérhető téren. Azt mondjuk, hogy λ_1 és λ_2 egymásra nézve szingulárisak, ha létezik B_1 és B_2 diszjunkt mérhető halmazok úgy, hogy λ_1 a B_1-re és λ_2 a B_2-re koncentrált. Ezt $\lambda_1 \perp \lambda_2$ módon jelöljük.

5.3.4. állítás. *Legyenek λ_1, λ_2 és μ előjeles mértékek. Ekkor fennállnak az alábbi implikációk:*
- $\lambda_1 \ll \mu$ *és* $\lambda_2 \ll \mu \Rightarrow (\lambda_1 + \lambda_2) \ll \mu$;
- $\lambda_1 \perp \mu$ *és* $\lambda_2 \perp \mu \Rightarrow (\lambda_1 + \lambda_2) \perp \mu$;
- $\lambda_1 \perp \mu$ *és* $\lambda_2 \ll \mu \Rightarrow \lambda_1 \perp \lambda_2$;
- $\lambda \perp \mu$ *és* $\lambda \ll \mu \Rightarrow \lambda = 0$.

Bizonyítás. Az első állítás nyilvánvaló. A másodikhoz legyen λ_1 a B_1-re, μ a C_1-re koncentrálva ahol $B_1 \cap C_1 = \varnothing$; továbbá λ_2 a B_2-re, μ a C_2-re koncentrálva valamint $B_2 \cap C_2 = \varnothing$. Legyen $B = B_1 \cup B_2$ és $C = C_1 \cap C_2$. Világos, hogy λ_1, λ_2 a B-re így $\lambda_1 + \lambda_2$ is B-re koncentrált, továbbá μ a C-re koncentrált. No de $B \cap C = (B_1 \cap C) \cup (B_2 \cap C) = \varnothing$.

A harmadik állítás indoklásához is legyen λ_1 a B-re és μ a C-re koncentrálva, ahol $B \cap C = \varnothing$. Eszerint $\mu(E \cap B) = \mu(E \cap B \cap C) = \mu(\varnothing) = 0$, így az abszolút folytonosság miatt $\lambda_2(E \cap B) = 0$. Ez viszont azt jelenti, hogy minden $E \in \mathcal{M}$ mérhető halmazra $\lambda_2(E) = \lambda_2(E \cap B^c)$, tehát λ_2 a B^c halmazra van koncentrálva.

Az utolsó állítás az előző következménye, mivel ekkor $\lambda \perp \lambda$, ami csak úgy lehet, hogy λ az $\varnothing$-ra van koncentrálva, tehát az azonos zéró mérték. $\square$

5.3.5. tétel (Lebesgue felbontás, Radon–Nikodym, von Neumann)**.** *Legyen λ és μ véges, nem negatív mértékek az $(X, \mathcal{M})$ mérhető téren. Ekkor*

· *létezik egyetlen λ_a és λ_s nem negatív mérték, amelyre*

$$\lambda_a \ll \mu, \ \lambda_s \perp \mu, \quad \textit{továbbá} \ \ \lambda_a + \lambda_s = \lambda;$$

· *létezik egyetlen $h \in L_1(X, \mathcal{M}, \mu)$ függvény, hogy minden $E \in \mathcal{M}$ mérhető halmaz esetén*

$$\lambda_a(E) = \int_E h \, d\mu.$$

Bizonyítás. Legyen $\varphi(E) = \lambda(E) + \mu(E)$ minden $E \in \mathcal{M}$ esetén. A monoton konvergenciatételt háromszor alkalmazva, kapjuk, hogy minden nem negatív mérhető f függvényre és minden E mérhető halmazra $\int_E f \, d\varphi = \int_E f \, d\lambda + + \int_E f \, d\mu$.

Az $(X, \mathcal{M}, \varphi)$ egy véges mértéktér. Definiáljuk tetszőleges $f \in L_2(X, \mathcal{M}, \varphi)$ esetén az alábbi $\Phi : L_2(X, \mathcal{M}, \varphi) \to \mathbb{R}$ függvényt:

$$\Phi(f) = \int_X f \, d\lambda.$$

A Hölder-egyenlőtlenség szerint

$$|\Phi(f)| = \left| \int_X f \cdot 1 \, d\lambda \right| \leq \int_X |f| \cdot 1 \, d\lambda \leq \left(\int_X |f|^2 \, d\lambda \right)^{\frac{1}{2}} \cdot \sqrt{\lambda(X)} \leq$$
$$\leq \left(\int_X |f|^2 \, d\varphi \right)^{\frac{1}{2}} \cdot \sqrt{\lambda(X)} = \sqrt{\lambda(X)} \cdot \|f\|_{L_2(\varphi)} .$$

Egyrészt a λ végessége miatt minden $f \in L_2(X, \mathcal{M}, \varphi)$ esetén $\Phi(f)$ véges, tehát valóban $\Phi : L_2(X, \mathcal{M}, \varphi) \to \mathbb{R}$ lineáris funkcionál, másrészt a fenti egyenlőtlenség azt jelenti, hogy Φ egy folytonos lineáris funkcionál az $L_2(X, \mathcal{M}, \varphi)$ Hilbert-téren (Riesz–Fischer-tétel). Tudjuk, hogy Hilbert-téren értelmezett folytonos lineáris funkcionálok, egy rögzített elemmel való skaláris szorzatként állnak elő. Létezik tehát egy $g \in L_2(X, \mathcal{M}, \varphi)$ függvény, melyre minden $f \in L_2(X, \mathcal{M}, \varphi)$ esetén

$$\int_X f \, d\lambda = \int_X fg \, d\varphi. \tag{†}$$

Most megmutatjuk, hogy olyan $g \in L_2(X, \mathcal{M}, \varphi)$ függvény is van, melyre a fenti (†) igaz, és $g(x) \in [0{,}1]$ minden $x \in X$ esetén. Ha ugyanis (†)-et felírjuk minden $f = \chi_E$ karakterisztikus függvényre, akkor azt kapjuk, hogy

$$\lambda(E) = \int_E g \, d\varphi,$$

ezért minden $E \in \mathcal{M}$ mérhető halmazra, amelyre $\varphi(E) \neq 0$

$$1 \geq \frac{1}{\varphi(E)} \lambda(E) = \frac{1}{\varphi(E)} \int_E g \, d\varphi \geq 0.$$

Ezek szerint $\varphi(\{x \in X : g(x) \notin [0{,}1]\}) = 0$, tehát egy φ-null-mértékű halmazon megváltoztathatjuk a g-t úgy, hogy $g(x) \in [0{,}1]$ már minden pontban teljesüljön és (†) is igaz maradjon.

Legyenek most

$$A = \{x \in X : 0 \leq g(x) < 1\} \quad \text{és} \quad B = \{x \in X : g(x) = 1\}.$$

Világos, hogy A és B diszjunkt mérhető halmazok, melyeknek egyesítése az egész X. Definiáljuk a következő λ_a és λ_s mértékeket:

$$\lambda_a(E) = \lambda(E \cap A) \text{ valamint } \lambda_s(E) = \lambda(E \cap B).$$

Világos, hogy λ_a és λ_s nem negatív mértékek, melyekre az $A \cup B = X$ miatt $\lambda = \lambda_a + \lambda_s$.

Megmutatjuk, hogy $\lambda_s \perp \mu$. A λ_s mérték a definíciója szerint B-re koncentrált. De

$$\lambda(E \cap B) = \int_{E \cap B} g \, d\varphi = \varphi(E \cap B) = \lambda(E \cap B) + \mu(E \cap B),$$

ezért $\mu(E \cap B) = 0$, minden $E \in \mathcal{M}$ mellett, ami azt jelenti, hogy μ a B-től diszjunkt $B^c = A$ halmazra van koncentrálva. Tehát $\lambda_s \perp \mu$ valóban fennáll.

Megmutatjuk, hogy $\lambda_a \ll \mu$. Ehhez találunk $h \in L_1(X, \mathcal{M}, \mu)$ függvényt, melyre $\lambda_a(E) = \int_E h \, d\mu$, ami bizonyítja is mindkét állítás egzisztenciára vonatkozó részét. Először átírjuk (†)-et: $\int_X f \, d\lambda = \int_X fg \, d\varphi = \int_X fg \, d\lambda +$ $+ \int_X fg \, d\mu$, azaz

$$\int_X f(1-g) \, d\lambda = \int_X fg \, d\mu.$$

Alkalmazzuk ezt az $f = \chi_E (1 + g + ... + g^n)$ nem negatív korlátos függvényre. Azt kapjuk így, hogy minden $E \in \mathcal{M}$ mérhető halmaz esetén és minden $n \in \mathbb{N}$ természetes számra

$$\int_E 1 - g^{n+1} \, d\lambda = \int_E g + g^2 + ... + g^{n+1} \, d\mu. \tag{‡}$$

A bal oldalt számolva: $\int_E 1 - g^{n+1} \, d\lambda = \int_{E \cap A} 1 - g^{n+1} \, d\lambda$, hiszen B pontjain g éppen 1. Az $x \in A$ esetén viszont $g(x) < 1$, ezért az $1 - g^{n+1}(x)$ sorozat monoton növekedőleg tart 1-hez, ami a monoton konvergenciatétel miatt azt jelenti, hogy (‡) bal oldala tart a $\lambda(E \cap A) = \lambda_a(A)$ számhoz.

A jobb oldalt számolva: legyen $h_n = g + g^2 + ... + g^{n+1}$ nem negatív korlátos függvény. Világos, hogy a $(h_n)_{n\in\mathbb{N}}$ függvénysorozat monoton növekedőleg tart valamely h mérhető nem negatív függvényhez, és a monoton konvergenciatételt újra használva azt kapjuk, hogy $\int_E h_n \, d\mu \to \int_E h \, d\mu$, ami azt jelenti, hogy $\lambda_a (E) = \int_E h \, d\mu$ valóban fennáll. Mivel $\int_X h \, d\mu = \lambda_a (X) = \lambda (B) < < \infty$, ezért $h \in L_1 (X, \mathcal{M}, \mu)$ is teljesül.

Csak az unicitások igazolása van hátra: Tegyük fel, hogy λ előáll $\lambda = \lambda_a^{'} + +\lambda_s^{'}$ alakban is. Ekkor átrendezve $\lambda_a^{'} - \lambda_a = \lambda_s - \lambda_s^{'}$ ahol a bal oldalon és a jobb oldalon is egy-egy előjeles mérték szerepel, amely egyrészt abszolút folytonos, másrészt szinguláris is μ-re, tehát $\lambda_a^{'} = \lambda_a$ és $\lambda_s^{'} = \lambda_s$. Amennyiben létezne $g \in L_1 (X, \mathcal{M}, \mu)$ nem negatív függvény, hogy $\lambda_a (E) = \int_E g \, d\mu$, akkor minden $E \in \mathcal{M}$ mérhető halmaz esetén $\int_E h - g \, d\mu = 0$, amiből valóban következik, hogy $h (x) = g (x)$ μ-majdnem minden $x \in X$ esetén. Ez azt jelenti, hogy h és g az $L_2 (X, \mathcal{M}, \mu)$ tér azonos ekvivalenciaosztályát képviselik. □

5.3.6.
Az előző tétel kiterjeszthető arra az esetre amikor mindkét mérték σ-véges, és arra az esetre is mikor λ komplex mérték, vagy λ előjeles mérték.

5.4. Jordan és Hahn felbontási tételei

Legyen λ valamely $\mathcal{M}$ σ-algebrán értelmezett komplex mérték, azaz $\lambda : \mathcal{M} \to \mathbb{C}$, melyre

$$\lambda \left(\cup_{n=}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \lambda (E_n)$$

minden megszámlálható diszjunkt $(E_n)_{n\in\mathbb{N}}$ halmaz-sorozat esetén. Világos, hogy a fenti definíció egyrészt általánosabb mint a szokásos nem negatív értéket felvevő mérték definíciója, másrészt pedig megszorítása annak, hiszen a $\lambda (E) = +\infty$ esetet ez nem tartalmazza. Azt is látni kell, hogy a fenti definíció szerint a $\sum_{n=1}^{\infty} \lambda (E_n)$ sor minden átrendezésének is konvergensnek kell lennie, ami valós vagy komplex számok lévén éppen az abszolút konvergenciával ekvivalens, tehát minden E_n diszjunkt halmaz-sorozatra $\sum_{n=1}^{\infty} |\mu (E_n)|$ is konvergens.

Az alábbiakban azt a technikát ismertetjük, hogy hogyan lehet egy ilyen komplex-speciális esetben valós, de véges-mérték szerinti integrál-elméletet visszavezetni nem negatív véges mérték szerinti integrál-elméletre. Az világos, hogy λ komplex mérték előáll $\lambda = \lambda_1 + i\lambda_2$ alakban, ahol λ_1 és λ_2 valós mértékek. Természetes gondolat lenne, hogy állítsuk elő a λ_1 valós mértéket a pozitív és negatív része különbségeként, ugyanúgy ahogyan ezt függvények esetén is tettük. Könnyen látható viszont, hogy egy mérték pozitív része nem feltétlen σ-additív, ezért ez a gondolat egyelőre zsákutcának bizonyul. Bevezetünk majd tetszőleges valós mértékhez két nem negatív mértéket, amelyeknek különbsége az adott előjeles mérték. Ez a Jordan-felbontási tétel.

5.4.1. definíció (mérték teljes változása)**.** Ha λ egy komplex mérték, akkor az

$$|\lambda|(E) = \sup\left\{\sum_{n=1}^{\infty} |\lambda(E_n)| : E = \cup_{n=1}^{\infty} E_n, E_i \cap E_j = \varnothing\right\}$$

halmazfüggvény a λ mérték *teljes változása*.

A fenti definíció az alábbi feladattal motiválható. Adott λ komplex mértékhez keressük meg azt a minimális μ nem negatív valós mértéket, amelyre

$$|\lambda(E)| \leq \mu(E). \tag{5.3}$$

Azt mondjuk, hogy $\hat{\mu}$ nem negatív mérték megoldása ennek a feladatnak, ha $|\lambda(E)| \leq \hat{\mu}(E)$ teljesül minden $E \in \mathcal{M}$ mellett és ha valamely μ nem negatív valós mértékre (5.3) fennáll, akkor $\hat{\mu} \leq \mu$.

Meg fogjuk mutatni, hogy a $|\lambda|$ teljes változás a fenti feladat megoldása.

5.4.2. állítás. *Komplex mérték teljes változása nem negatív mérték.*

Bizonyítás. Triviális, hogy $|\mu|(\varnothing) = 0$. Legyen $E = \cup_{i=1}^{\infty} E_i$, diszjunkt mérhető halmazok egyesítése.

$|\mu|(E) \leq \sum_i |\mu|(E_i)$. Ugyanis, ha $E = \cup A_j$ diszjunkt egyesítés, akkor

$$\sum_j |\mu(A_j)| = \sum_j \sum_i |\mu(A_j \cap E_i)| = \sum_i \sum_j |\mu(A_j \cap E_i)| \leq \sum_i |\mu|(E_i).$$

De $|\mu|(E)$ épp a fenti egyenlőtlenség bal oldalán lévő halmazok szuprémuma, ezért a $|\mu|(E) \leq \sum_i |\mu|(E_i)$ egyenlőtlenség valóban fennáll.

$\sum_i |\mu|(E_i) \leq |\mu|(E)$. Ugyanis legyenek a $t_i \in \mathbb{R}$ számokra $t_i < |\mu|(E_i)$. A szuprémum definíciója szerint léteznek $A_{i,j}$ $(j = 1,2,\ldots)$ egymástól diszjunkt halmazok, melyekre $\sum_j |\mu(A_{i,j})| > t_i$. De az E_i halmazok diszjunktsága miatt az összes $A_{i,j}$ halmaz is diszjunkt, ilyen módon egy partícióját kapjuk E-nek. Erre persze

$$\sum_i t_i < \sum_i \sum_j |\mu(A_{i,j})| \leq |\mu|(E).$$

De ez minden $t_i < |\mu|(E_i)$ valós szám mellett igaz, tehát $\sum_i |\mu|(E_i) \leq |\mu|(E)$ valóban fennáll. □

Világos, hogy ha adott egy nem negatív véges mérték, akkor ennek értékkészlete benne van egy korlátos és zárt intervallumban. Ehhez hasonlóan az is igaz, hogy amennyiben adott egy komplex mérték, úgy a mérték értékkészlete része egy zárt körnek. Ez talán meglepő, hiszen $A \subseteq B$ nem implikálja $|\mu(A)| \leq |\mu(B)|$ -t, ezért $\mu(X) \in \mathbb{C}$ feltétel nem tűnik elegendőnek. Azt fogjuk majd megmutatni, hogy minden $E \in \mathcal{M}$ esetén $|\mu(E)| \leq |\mu|(X) \in \mathbb{R}$. Ebből az állításból mindenesetre a $|\mu(E)| \leq |\mu|(E)$ egyenlőtlenség triviális, és mivel $|\mu|$ mérték így elegendő belátni, hogy $|\mu|(X) \in \mathbb{R}$.

Ennek igazolásához egy elemi lemma következik, melyhez csak a komplex számok fogalma szükséges:

5.4.3. lemma. *Legyenek a $z_1, \dots, z_n$ komplex számok rögzítve. Ekkor ezeknek van olyan $S \subseteq \{z_1, \dots, z_n\}$ részhalmaza, melyre $\left|\sum_{i \in S} z_i\right| > \frac{1}{6} \sum_{i=1}^{n} |z_i|$.*

Bizonyítás. Az $\mathbb{R}^2$ sík $|y| \geq x$ térnegyedébe eső z komplex számokra $\sqrt{2}\Re z \geq$ $\geq |z|$. Feltehető, hogy ebbe térnegyedbe esik az az S részhalmaz, melyre $\sum_{z_i \in S} |z_i| \geq \frac{1}{4} \sum_{i=1}^{n} |z_i|$. Erre az S részhalmazra $\left|\sum_{z_i \in S} z_i\right| \geq \Re\left(\sum_{z_i \in S} z_i\right) =$ $= \sum_{z_i \in S} \Re z_i \geq \frac{1}{4\sqrt{2}} \sum_{i=1}^{n} |z_i| > \frac{1}{6} \sum_{i=1}^{n} |z_i|$. □

5.4.4. állítás. *Ha μ egy komplex mérték, akkor $|\mu|(X) < \infty$, azaz komplex mérték teljes változása nem negatív véges mérték.*

Bizonyítás. Először azt mutatjuk meg, hogy amennyiben $Y \in \mathcal{M}$, $|\mu|(Y) =$ $= \infty$, úgy Y felbontható diszjunkt $Y = A \cup B$ mérhető halmazok egyesítésére, hogy $|\mu(A)| > 1$ és $|\mu|(B) = \infty$. Mivel $|\mu|$ teljes változás egy mérték, ezért ehhez elegendő belátni, hogy $|\mu(A)| > 1$ és $|\mu(B)| > 1$ is teljesül. Legyen $t =$ $= 6(1 + |\mu(Y)|)$. Világos, hogy létezik E_i $(i = 1, \dots, n)$ véges sok diszjunkt halmaz, melyre $\cup E_i \subseteq Y$ és

$$t < \sum_{i=1}^{n} |\mu(E_i)|.$$

Van tehát olyan S részhalmaz, hogy $\left|\sum_{i \in S} \mu(E_i)\right| > \frac{1}{6} \sum_{i=1}^{n} |\mu(E_i)|$. Legyen $A = \cup_{i \in S} E_i$. Ekkor

$$|\mu(A)| = \left|\sum_{i \in S} \mu(E_i)\right| > \frac{1}{6} \sum_{i=1}^{n} |\mu(E_i)| > \frac{t}{6}.$$

Ebből $\frac{t}{6} = 1 + |\mu(Y)| > 1$ miatt egyrészt $|\mu(A)| > 1$ következik, másrészt $B = Y \setminus A$ jelöléssel

$$|\mu(B)| = |\mu(Y) - \mu(A)| \geq |\mu(A)| - |\mu(Y)| > \frac{t}{6} - |\mu(Y)|,$$

amiből $\frac{t}{6} - |\mu(Y)| > 1$ miatt $|\mu(B)| > 1$ is fennáll.

Tegyük fel indirekt, hogy $|\mu|(X) = \infty$. Ekkor $X = A_1 \cup B_1$ diszjunkt egyesítésként, hogy $|\mu(A_1)| > 1$ és $|\mu|(B_1) = \infty$. Alkalmazva az előző gondolatot

B_1-re kapjuk, hogy $B_1 = A_2 \cup B_2$ diszjunkt egyesítés alakban és $|\mu(A_2)| > 1$, de $|\mu|(B_2) = \infty$ is fennáll. Az eljárást folytatva olyan diszjunkt A_n halmaz-sorozatot kapunk, hogy minden elemre $|\mu(A_n)| > 1$, ami ellentmond a $\sum \mu(A_n)$ sor konvergenciájának. □

Bebizonyítottuk tehát az alábbi állítást.

5.4.5. állítás. *Legyen μ egy komplex mérték. Ekkor μ értékkészlete része az origó középpontú $|\mu|(X)$ sugarú zárt körlemeznek.*

5.4.6. definíció (mérték pozitív és negatív változása)**.** Legyen λ egy előjeles mérték. Jelölje

$$\lambda^+ = \frac{1}{2}(|\lambda| + \lambda) \text{ valamint } \lambda^- = \frac{1}{2}(|\lambda| - \lambda)$$

a λ mérték *pozitív* illetve *negatív változását.*

5.4.7.
Mivel előjeles mértékek összege is előjeles mérték, ezért λ^+ és λ^- valóban előjeles mértékek. No de minden $E \in \mathcal{M}$ mellett $|\lambda(E)| \leq |\lambda|(E)$, ezért $-\lambda(E) \leq |\lambda|(E)$ és $\lambda(E) \leq |\lambda|(E)$ is fennáll. Eszerint $|\lambda| + \lambda$, ezért λ^+; és $|\lambda| - \lambda$, ezért λ^- is nem negatív mértékek. Amennyiben λ véges, úgy λ^+ és λ^- is az.

5.4.8. állítás (Jordan-féle felbontási tétel)**.** *Legyen λ egy véges előjeles mérték. Ekkor a λ^+ pozitív változás és a λ^- negatív változás halmazfüggvények véges nem negatív mértékek, továbbá*

$$\lambda = \lambda^+ - \lambda^- \textit{ valamint } |\lambda| = \lambda^+ + \lambda^-.$$

Bizonyítás. Az előzőek szerint csak a kiemelt egyenleteket kell indokolnunk. Nyilvánvalóan $\lambda^+ + \lambda^- = \frac{1}{2}(|\lambda| + \lambda + |\lambda| - \lambda) = |\lambda|$ és hasonlóan $\lambda^+ - \lambda^- = \frac{1}{2}(|\lambda| + \lambda - (|\lambda| - \lambda)) = \lambda$. Ezt kellett belátni. □

Most a Hahn-felbontási tételre térünk rá. Ehhez először általánosítanunk kell a Radon–Nikodym-tételt arra az esetre, mikor λ előjeles mérték:

5.4.9.
Tegyük fel, hogy $\lambda \ll \mu$. Ekkor $|\lambda| \ll \mu$ fennáll hiszen, ha $\mu(B) = 0$, akkor annak minden $A \in \mathcal{M}$, $A \subseteq B$ részhalmazára is $\mu(B) = 0$, amiből az abszolút folytonosság miatt már $\lambda(A) = 0$ is fennáll. Ebből persze $|\lambda|(B) = 0$ már következik. Ekkor viszont a pozitív változás és negatív változás definíciójára gondolva $\lambda^+ \ll \mu$ és $\lambda^- \ll \mu$. Alkalmazható tehát a Radon–Nikodym-tételnek a nem negatív mértékekre vonatkozó alakja. Kapunk tehát $h^+, h^- \in L_1(\mu)$ függvényeket, melyekre $\lambda^+(E) = \int_E h^+ \, d\mu$ és $\lambda^-(E) = \int_E h^- \, d\mu$. Ebből persze $\lambda(E) = \int h^+ - h^- \, d\mu$.

5.4.10. állítás. *Legyen μ egy véges előjeles mérték. Ekkor létezik egyetlen $h \in L_1(X, \mathcal{M}, |\mu|)$ mérhető függvény, amelyre*

$$\mu(E) = \int_E h \, d|\mu|, \quad \textit{valamint} \quad |h| = 1.$$

Bizonyítás. Triviális, hogy $\mu \ll |\mu|$, ezért a Radon–Nikodym-tétel miatt létezik h mérhető függvény, melyre $\mu(E) = \int_E h \, d|\mu|$ minden $E \in \mathcal{M}$ mérhető halmazra. Amennyiben $|\mu|(E) \neq 0$, akkor

$$\left| \frac{1}{|\mu|(E)} \int_E h \, d|\mu| \right| = \frac{|\mu(E)|}{|\mu|(E)} \leq 1,$$

amiből következik, hogy $|\mu|$ majdnem minden $x \in X$ mellett $-1 \leq h(x) \leq 1$.

Legyen most $r \in (0{,}1)$ tetszőleges szám és $A_r = \{x \in X : |h(x)| < r\}$. Elég megmutatnunk, hogy $|\mu|(A_r) = 0$, hiszen

$$\{x \in X : |h(x)| < 1\} = \cup_{n=1}^{\infty} \left\{ x \in X : |h(x)| < 1 - \frac{1}{n} \right\}$$

és $|\mu|$ mérték. Alkossák az E_j halmazok az A_r partícióját, tehát $A_r = \cup E_j$ és $E_j \cap E_i = \varnothing$ $(j \neq i)$. Minden E_j-re

$$|\mu(E_j)| = \left| \int_{E_j} h \, d|\mu| \right| \leq \int_{E_j} |h| \, d|\mu| < \int_{E_j} r \, d|\mu| = r|\mu|(E_j).$$

Ezért $\sum_j |\mu(E_j)| < r \sum_j |\mu|(E_j) = r|\mu|(A_r)$. Mivel épp a bal oldali szummák szuprémuma $|\mu|(A_r)$, ezért ebből $|\mu|(A_r) \leq r|\mu|(A_r)$ is következik, ami $r < 1$ miatt csak abban az esetben lehetséges, ha $|\mu|(A_r) = 0$. □

5.4.11. definíció (előjeles mérték szerinti integrál)**.** Legyen λ egy előjeles mérték az $(X, \mathcal{M})$ mérhető téren, és $f \in L_1(X, \mathcal{M}, |\lambda|)$ függvény. Az f függvény λ előjeles mérték szerinti integrálját, az

$$\int_X f \, d\lambda = \int_X f h \, d|\lambda|$$

formula definiálja, ahol h az egyetlen olyan $h \in L_1(X, \mathcal{M}, |\lambda|)$, függvény amelyre a $|h| = 1$ egyenlőség $|\lambda|$-majdnem mindenütt teljesül és $\lambda(E) = \int_E h \, d|\lambda|$ fennáll minden $E \in \mathcal{M}$ mellett.

5.4.12.
A fenti definíció értelmes, hiszen ha $f \in L_1(|\lambda|)$ és $|h| = 1$, akkor $fh \in L_1(|\lambda|)$.

Speciálisan, ha λ egy véges nem negatív mérték, akkor $h = 1$ és $|\lambda| = \lambda$, így ebben az esetben a fenti definíció egybeesik az az L_1-beli függvények integráljának definíciójával.

Az is látható, hogy a véges előjeles mérték szerinti integrál egy lineáris funkcionál az $L_1(X, \mathcal{M}, |\lambda|)$ normált téren.

5.4.13. állítás (előjeles integrál-mérték szerinti integrál)**.** *Legyen $(X, \mathcal{M}, \mu)$ egy nem negatív mértéktér, és $g \in L_1(X, \mathcal{M}, \mu)$ mérhető függvény. Láttuk, hogy*

$$\lambda(E) = \int_E g\, d\mu, \quad E \in \mathcal{M}$$

definícióval λ egy előjeles mérték az $\mathcal{M}$ σ-algebrán. Ekkor tetszőleges $f \in L_1(X, \mathcal{M}, |\lambda|)$ függvény esetén

$$\int_X f\, d\lambda = \int_X fg\, d\mu.$$

Bizonyítás. Világos, hogy $\lambda \ll \mu$, így $|\lambda| \ll \mu$. Ez utóbbira felírva a nem negatív véges mértékekre vonatkozó Radon–Nikodym-tételt, azt kapjuk, hogy létezik $u \in L_1(\mu)$ nem negatív függvény, amelyre

$$|\lambda|(E) = \int_E u\, d\mu.$$

Az előjeles mérték szerinti integrál definíciója és $fh \in L_1(|\lambda|)$ szerint

$$\int_X f\, d\lambda = \int_X fh\, d|\lambda| = \int_X fhu\, d\mu.$$

Ez tetszőleges $f \in L_1(|\lambda|)$ függvényre fennáll, speciálisan tehát $f = \chi_E$ karakterisztikus függvény mellett is tetszőleges $E \in \mathcal{M}$ mérhető halmazra is. Így

$$\lambda(E) = \int_E h\, d|\lambda| = \int_X \chi_E h\, d|\lambda| = \int_X \chi_E hu\, d\mu = \int_E hu\, d\mu.$$

No de λ definíciója szerint ez csak úgy lehet, ha a $g = hu$ egyenlőség μ-majdnem mindenütt fennáll. Ez viszont azt jelenti, hogy

$$\int_X f\, d\lambda = \int_X fg\, d\mu.$$

Ezt kellett belátni. □

5.4.14. állítás. *Legyen μ egy nem negatív mérték és $g \in L^1(\mu)$ függvény. Világos, hogy ekkor $\lambda(E) = \int_E g\, d\mu$ integrál véges előjeles mértéket definiál. E mérték teljes változására:*

$$|\lambda|(E) = \int_E |g|\, d\mu,$$

továbbá a pozitív és negatív változásokra:

$$\lambda^+(E) = \int_E g^+ \, d\mu \text{ valamint } \lambda^-(E) = \int_E g^- \, d\mu.$$

Bizonyítás. Alkalmazzuk az előző állítást λ mértékre. Ekkor létezik $|h| = 1$ függvény, hogy

$$\int_E g \, d\mu = \lambda(E) = \int_E h \, d|\lambda|.$$

Így minden $E \in \mathcal{M}$ mérhető halmazra

$$|\lambda|(E) = \int_E 1 \, d|\lambda| = \int_E h^2 \, d|\lambda| = \int_E hh \, d|\lambda| = \int_E h \, d\lambda = \int_E hg \, d\mu$$

De $|\lambda|$ nem negatív mérték, ezért $hg \geq 0$ teljesül μ-majdnem mindenütt. Így $hg = |hg| = |g|$ is fennáll, azaz $|\lambda|(E) = \int_E |g| \, d\mu$ valóban teljesül minden $E \in \mathcal{M}$ mellett.

A pozitív változásra vonatkozó állítás igazolásához

$$\lambda^+(E) = \frac{1}{2}(|\lambda| + \lambda)(E) = \int_E \frac{1}{2}(|g| + g) \, d\mu = \int_E g^+ \, d\mu,$$

és hasonlóan adódik a negatív változásra vonatkozó $\lambda^-(E) = \int_E g^- \, d\mu$ formula is. □

5.4.15. állítás (Hahn felbontási tétele)**.** *Legyen az $(X, \mathcal{M})$ mérhető téren μ egy véges előjeles mérték. Ekkor X felbontható diszjunkt A és B halmazok egyesítésére úgy, hogy a mérték μ^+ pozitív változása a μ-nek az A-ra koncentrált része, és a μ^- negatív változása a μ-nek B-re koncentrált része legyen. Pontosabban: létezik $A, B \in \mathcal{M}$, $A \cap B = \varnothing$, $A \cup B = X$ amelyekre*

$$\mu^+(E) = \mu(E \cap A) \text{ valamint } \mu^-(E) = -\mu(E \cap B).$$

Bizonyítás. Láttuk, hogy létezik $h \in L^1(|\mu|)$, $|h| = 1$ függvény, melyre $\mu(E) = \int_E h \, d|\mu|$ fennáll minden $E \in \mathcal{M}$ mellett. Legyen

$$A = \{x \in X : h(x) = 1\} \quad \text{és} \quad B = \{x \in X : h(x) = -1\}.$$

Világos, hogy $A \cup B = X$ és $A \cap B = \varnothing$, $A, B \in \mathcal{M}$, továbbá az is könnyen látszik, hogy $h^+ = \chi_A$. Tetszőleges $E \in \mathcal{M}$ esetén a az előző állítás pozitív változásra vonatkozó része szerint

$$\mu^+(E) = \int_E h^+ \, d|\mu| = \int_E \chi_A \, d|\mu| = \int_E \chi_A h \, d|\mu| = \int_E \chi_A \, d\mu = \mu(E \cap A).$$

Ezzel a pozitív részre vonatkozó állítást igazoltuk is. A negatív részre vonatkozó állítás már következik a fenti formulából hiszen tudjuk, hogy $\mu(E) = \mu^+(E) - \mu^-(E)$ és nyilván $\mu(E) = \mu(E \cap A) + \mu(E \cap B)$. □

6. fejezet

A Radon–Nikodym-tétel következményei

6.1. Az L_p terek duálisa

A fejezetben adott $1 < p < \infty$ mellett jelölje minden esetben $1 < q < \infty$ azt a számot, amelyre $\frac{1}{p} + \frac{1}{q} = 1$. Ha $p = 1$, akkor legyen $q = \infty$ és hasonlóan, ha $p = \infty$, akkor legyen $q = 1$.

Az alábbi kis lemmácskát sokat fogjuk használni.

6.1.1. lemma. *Legyen $1 \leq p < \infty$, $f \in L_p(X, \mathcal{M}, \mu)$, továbbá $A = \cup_{n=1}^{\infty} A_n$, ahol $A_n \in \mathcal{M}$, $A_n \subseteq A_{n+1}$ monoton bővülő halmaz-sorozat. Ekkor $f\chi_{A_n} \to f\chi_A$ az $L_p(X, \mathcal{M}, \mu)$ metrikája szerint.*

Bizonyítás. Definiálja $h_n = f\chi_{A_n}$. Az A_n halmaz-sorozat monoton növekedése miatt $h_n \to f\chi_A$ μ-m.m. Az $f^p \in L_1$ miatt f μ-m.m. véges. Ezért $|h_n - f\chi_A|^p \to 0$ μ-m.m. pontonként. No de $|h_n - f\chi_A| \leq |h_n| + |f| = 2|f|$, ezért $|h_n - f\chi_A|^p \leq 2^p |f|^p$. Eszerint találtunk $2^p |f|^p \in L_1$ majoráns függvényt. Alkalmazva a majorált konvergenciatételt, azt kapjuk, hogy

$$\|h_n - f\chi_A\|_p^p = \int_X |h_n - f\chi_A|^p \, d\mu \to 0,$$

azaz $h_n \to f\chi_A$ az L_p tér metrikája szerint. □

Vegyük észre, hogy a fenti állítás hamis $p = \infty$ esetben. Ha például $f = 1$ és $A_n = [-n, n]$, akkor χ_{A_n} nem tart egyenletesen a konstans 1 függvényhez.

6.1.2. állítás. *Legyen $1 \leq p \leq \infty$. A korlátos függvények L_p egy sűrű részhalmazát alkotják.*

Bizonyítás. Az állítás a $p = \infty$ esetben semmitmondó. Ha $p < \infty$, akkor legyen $A_n = X(f \leq n)$. Világos, hogy $A_n \subseteq A_{n+1}$ és $f^p \in L_1$ miatt f μ-m.m. véges, ezért null-mértékű halmaztó eltekintve $\cup A_n = X$. Alkalmazva az előző állítást azt kapjuk, hogy

$$f\chi_{A_n} \to f\chi_{\left(\cup_{n=1}^{\infty} A_n\right)} = f$$

az L_p-tér metrikájában. Persze az $f\chi_{A_n} \in L_p$ korlátos függvények, tehát készen is vagyunk.

6.1.3. állítás. *Legyen $1 \leq p < \infty$ és $g \in L_q(X, \mathcal{M}, \mu)$ rögzített. Ekkor $f \in L_p(X, \mathcal{M}, \mu)$ mellett a*

$$\Phi_g(f) = \int_X fg d\mu$$

módon definiált függvény egy folytonos lineáris funkcionálja $L_p(X, \mathcal{M}, \mu)$-nek. Ha $p > 1$; vagy a $p = 1$ esetben ha az $(X, \mathcal{M}, \mu)$ mértéktér σ-véges, akkor a

$$\|\Phi_g\|_{L_p^*} = \|g\|_q .$$

egyenlőség is fennáll.

□

Bizonyítás. Nyilván feltehető, hogy $\|g\|_{L_q} \neq 0$. A Hölder-egyenlőtlenség szerint minden $f \in L_p$ $(1 \leq p \leq \infty)$ függvény esetén

$$|\Phi_g(f)| = \left|\int fg d\mu\right| \leq \|f\|_p \|g\|_q ,$$

tehát Φ folytonos lineáris funkcionál, és $\|\Phi_g\|_{L_p^*} \leq \|g\|_q$.

A $\infty > p > 1$ mellett legyen $f = \alpha |g|^{q-1}$, ahol α olyan mérhető függvény melyre $\alpha g = |g|$. Ekkor

$$\|f\|_p^p = \int_X |f|^p \, d\mu = \int_X |g|^{p(q-1)} \, d\mu = \int_X |g|^q \, d\mu,$$

tehát $f \in L_p$ továbbá $\|f\|_p = \|g\|_q^{q/p}$, ezért

$$\Phi_g\left(\alpha |g|^{q-1}\right) = \int_X \alpha |g|^{q-1} g d\mu = \int_X |g|^q \, d\mu = \|g\|_q^{q(1/p+1/q)} = \|f\|_p \|g\|_q ,$$

ez azt jelenti, hogy a $h = f / \|f\|_p$ jelölést bevezetve találtunk $h \in L_p$, $\|h\|_p = 1$ függvényt, amelyre $\Phi(h) = \|g\|_q$. Ez épp azt jelenti, hogy $\|\Phi_g\|_{L_p^*} \geq \|g\|_q$ is fennáll.

Most tekintsük a $p=1$ esetet, amikor μ mérték σ-véges, azaz $X=\cup X_n$, ahol $\mu(X_n)<\infty$. Legyen rögzített $n\in\mathbb{N}$ mellett

$$A_n = X_n\left(|g|>\|\Phi_g\|_{L_1^*}\right).$$

Megmutatjuk, hogy $\mu(A_n)=0$. Ugyanis, $\mu(X_n)<\infty$ miatt $\chi_{A_n}\in L_1$. Ha $\mu(A_n)\neq 0$, akkor az $f_n=\frac{1}{\mu(A_n)}\alpha\chi_{A_n}$ függvényre $\|f_n\|_1=1$, ahol α olyan mérhető $|\alpha|=1$ függvény, amelyre $\alpha g=|g|$. No de

$$\|\Phi_g\|_{L_1^*}\geq \Phi_g(f_n)=\int_X f_n g d\mu=\frac{1}{\mu(A_n)}\int_X \chi_{A_n}\alpha g d\mu=\frac{1}{\mu(A_n)}\int_{A_n}|g|\,d\mu>$$
$$>\frac{1}{\mu(A_n)}\mu(A_n)\|\Phi_g\|_{L_1^*}=\|\Phi_g\|_{L_1^*},$$

ami nyilván ellentmondás, tehát $\mu(A_n)=0$ valóban fennáll. Világos, hogy

$$X\left(|g|>\|\Phi_g\|_{L_1^*}\right)=\cup_{n=1}^{\infty}A_n,$$

ezért $\mu\left(X\left(|g|>\|\Phi_g\|_{L_1^*}\right)\right)=0$, ami épp azt jelenti, hogy $\|g\|_\infty\leq\|\Phi_g\|_{L_1^*}$. □

6.1.4. állítás. *Legyen $1\leq p<\infty$ és $\Phi\in L_p^*(X,\mathcal{M},\mu)$ rögzített folytonos lineáris funkcionál, valamint $\mu(X)<\infty$ véges nem negatív mérték. Ekkor létezik egyetlen $g\in L_q(X,\mathcal{M},\mu)$ függvény, amelyre minden $f\in L_p(X,\mathcal{M},\mu)$ esetén*

$$\Phi(f)=\int_X fg d\mu.$$

Az előző állítás miatt ekkor $\|\Phi\|=\|\Phi_g\|=\|g\|_q$ is fennáll.

Bizonyítás. Az állítás unicitási része egyszerű, hiszen ha g_1 és g_2 két ilyen függvény, akkor $\mu(X)$ végessége miatt minden $E\in\mathcal{M}$ mellett $\chi_E\in L_p(X)$, ezért

$$\int_E (g_1-g_2)\,d\mu=0,$$

ami csak úgy lehet, hogy $g_1=g_2$ μ-m.m. teljesül.

Tetszőleges $E\in\mathcal{M}$ mellett $\chi_E\in L_p$, hiszen μ véges mérték. Értelmes tehát a

$$\lambda(E)=\Phi(\chi_E)$$

definíció. Világos, hogy $\lambda:\mathcal{M}\to\mathbb{R}$ halmaz függvény, amelyre $\lambda(\varnothing)=0$. A λ végesen additivitása a Φ linearitásának következménye. Megmutatjuk, hogy λ mérték. Ha $A=\cup_{n=1}^{\infty}A_n$ diszjunkt egyesítés, akkor $N\to\infty$ mellett

$$\|\chi_{\cup_{n=N}^{\infty}A_n}\|_p=\mu\left(\cup_{n=N}^{\infty}A_n\right)=\sum_{n=N}^{\infty}\mu(A_n)\to 0$$

a μ σ-additivitása miatt. Tehát a $\left(\chi_{\cup_{n=N}^{\infty} A_n}\right)_{N\in\mathbb{N}}$ függvénysorozat L_p-ben konvergál $0 \in L_p$-hez. A Φ folytonossága szerint:

$$\lambda\left(\cup_{n=N}^{\infty} A_n\right) = \Phi\left(\chi_{\cup_{n=N}^{\infty} A_n}\right) \to 0.$$

Így kihasználva λ végesen additivitását azt kapjuk, hogy

$$\lambda\left(\cup_{n=1}^{\infty} A_n\right) = \lambda\left(\cup_{n=1}^{N-1} A_n\right) + \lambda\left(\cup_{n=N}^{\infty} A_n\right) = \sum_{n=1}^{N-1} \lambda(A_n) + \lambda\left(\cup_{n=N}^{\infty} A_n\right),$$

amiből már látszik, hogy $N \to \infty$ esetén

$$\lambda\left(\cup_{n=1}^{\infty} A_n\right) - \sum_{n=1}^{N-1} \lambda(A_n) = \lambda\left(\cup_{n=N}^{\infty} A_n\right) \to 0.$$

Tehát λ valóban egy véges $\mathbb{R}$-értékű mérték.

Ha $E \in \mathcal{M}$ mellett $\mu(E) = 0$, akkor χ_E a zérus eleme L_q-nak, ezért

$$\lambda \ll \mu.$$

Alkalmazható tehát az 5.3.5. Radon–Nikodym-tétel, amely szerint létezik egyetlen olyan $g \in L^1(X, \mathcal{M}, \mu)$ függvény, amelyre

$$\Phi(\chi_E) = \lambda(E) = \int_E g d\mu = \int_X \chi_E g d\mu.$$

Most megmutatjuk, hogy az így kapott g jó is lesz az állítás bizonyításához. A fenti sor szerint ezzel készen is vagyunk abban az esetben, ha az $f \in L_p$ függvény valamely mérhető halmaz indikátor függvénye. Az egyszerű függvények indikátor függvények véges együtthatókkal képzett lineáris kombinációi, ezért ha $f = \sum_{i=1}^{N} \alpha_i \chi_{E_i}$ egyszerű függvény, akkor $f \in L_p$, továbbá

$$\Phi(f) = \sum_{i=1}^{N} \alpha_i \Phi(\chi_{E_i}) = \sum_{i=1}^{N} \alpha_i \int_X \chi_{E_i} g d\mu = \int_X \left(\sum_{i=1}^{N} \alpha_i \chi_{E_i}\right) g d\mu = \int_X f g d\mu.$$

Egyrészt, ha f korlátos függvény, akkor f előáll mint s_n egyszerű függvények egyenletes határértéke. De $\mu(X) < \infty$, így

$$\|f - s_n\|_p^p = \int_X |f - s_n|^p d\mu \le \mu(X) \|f - s_n\|_\infty^p \to 0.$$

Azaz $s_n \to f$ az L_p metrikája szerint is. Ezért $n \to \infty$ mellett

$$\Phi(s_n) \to \Phi(f). \tag{†}$$

Másrészt, $s_n g \to fg$ μ-m.m. pontonként, sőt elég nagy n-re

$$|s_n g| \leq 2 \|f\|_\infty |g|$$

az s_n egyenletes konvergenciája miatt. A μ mérték végességét újra kihasználva azt látjuk, hogy a $2\|f\|_\infty |g|$ függvény egy L_1-beli majoránsa az $s_n g$ függvénysorozatnak. Alkalmazva a majorált konvergenciatételt

$$\Phi(s_n) = \int_X s_n g d\mu \to \int_X f g d\mu. \tag{‡}$$

Összevetve: (†) és (‡) szerint tetszőleges f korlátos, ezért L_p-beli függvény mellett

$$\Phi(f) = \int_X f g d\mu.$$

Tudjuk viszont, hogy a korlátos függvények L_p-ben sűrű részhalmazt alkotnak, amelyen a Φ és a Φ_g folytonos függvények a fent bizonyítottak miatt egybeesnek. Ekkor

$$\Phi(f) = \Phi_g(f) = \int_X f g d\mu$$

minden $f \in L_p$ mellett is fennáll.

Már csak azt kell igazolnunk, hogy $g \in L_q$. Ehhez szét kell választanunk a $p = 1$ és a $1 < p$ esetet.

$p = 1$ esetben, meg kell mutatnunk, hogy g μ-m.m. korlátos. Ehhez tekintsük a

$$\frac{1}{\mu(E)} \int_E g d\mu, \quad \mu(E) > 0$$

integrál közepeket. No de $\mu(E) > 0$ mellett

$$\left|\frac{1}{\mu(E)} \int_E g d\mu\right| = \frac{|\lambda(E)|}{|\mu(E)|} = \frac{1}{|\mu(E)|} |\Phi(\chi_E)| \leq \frac{1}{|\mu(E)|} \|\Phi\|_{L_1^*} \|\chi_E\|_1 = \|\Phi\|_{L_1^*}.$$

Azt kaptuk tehát, hogy μ-m.m. $x \in X$ mellett $|g(x)| \leq \|\Phi\|_{L_1^*}$.

$p > 1$ esetben, meg kell mutatnunk, hogy $\int_X |g|^q d\mu < \infty$. Legyen $E_n = X(g \leq n)$, α olyan mérhető függvény, amelyre $\alpha g = |g|$, továbbá $f = \chi_{E_n} \alpha |g|^{q-1}$. Persze f korlátos ezért $\mu(X)$ végessége miatt $f \in L_p$. Így:

$$\Phi(f) = \int_X f g d\mu = \int_{E_n} |g|^q d\mu,$$

ezért

$$\int_{E_n} |g|^q d\mu = |\Phi(f)| \leq \|\Phi\| \|f\|_p =$$

$$= \|\Phi\| \left(\int_X |f|^p d\mu\right)^{1/p} = \|\Phi\| \left(\int_{E_n} |g|^q d\mu\right)^{1/p},$$

amiből azt kapjuk, hogy $\left(\int_{E_n} |g|^q \, d\mu\right)^{1-1/p} = \left(\int_{E_n} |g|^q \, d\mu\right)^{1/q} \leq \|\Phi\|$. No de $\chi_{E_n} |g|^q \to |g|^q$ monoton növekedőleg μ-m.m. pontonként a $g \in L_1$ miatt, ezért a monoton konvergenciatétel szerint

$$\int_X \chi_{E_n} |g|^q \, d\mu \to \int_X |g|^q \, d\mu.$$

Azt kaptuk tehát, hogy $\|g\|_q \leq \|\Phi\|$ is fennáll, azaz $g \in L_q$ valóban teljesül. □

6.1.5. állítás. *Tekintsük az $(X, \mathcal{M}, \mu)$ mértékteret, és az ehhez tartozó $L_p(X, \mathcal{M}, \mu)$ Banach-teret. Legyen $\Phi \in L_p^*$ folytonos lineáris funkcionál rögzítve.*

· Ha $p = 1$, és μ σ-véges mérték, akkor létezik egyetlen $g \in L_\infty$ függvény, amelyre

$$\Phi = \Phi_g, \ \textit{valamint} \ \|\Phi\| = \|g\|_\infty .$$

· Ha $1 < p < \infty$, akkor létezik egyetlen $g \in L_q$ függvény, amelyre

$$\Phi = \Phi_g, \ \textit{valamint} \ \|\Phi\| = \|g\|_q .$$

Bizonyítás. Mindkét állítást igazoltuk a $\mu(X) < \infty$ speciális esetben. Legyen most $A \in \mathcal{M}$ egy $\mu(A) < \infty$ véges mértékű mérhető halmaz. Tekintsük az ehhez tartozó $(A, \mathcal{M}_A, \mu_A)$ véges mértékteret, ahol

$$\mathcal{M}_A = \{M \cap A : M \in \mathcal{M}\}; \quad \mu_A = \mu_{|\mathcal{M}_A}.$$

Jelölje egyszerűen $L_p(A)$ az $L_p(A, \mathcal{M}_A, \mu_A)$ mértékteret és $L_p^*(A)$ ennek duális terét. Tekinthetjük az $L_p(A)$ Banach-teret mint az $L_p(X)$ alterét, hiszen egy $f \in L_p(A)$ függvényt A^c-re kiterjeszthetjük $\tilde{f}_{|A^c} = 0$ és $\tilde{f}_{|A} = f$ módon. Így $\|\tilde{f}\|_{L_p} = \|f\|_{L_p(A)}$. Ha $\Phi \in L_p^*(X)$ folytonos lineáris funkcionál, akkor az folytonos lineáris funkcionál $L_p(A)$ altéren is, hiszen

$$f \in L_p(A) \ \text{mellett} \ \left|\Phi\left(\tilde{f}\right)\right| \leq \|\Phi\|_{L_p^*(X)} \|\tilde{f}\|_{L_p(X)} = \|\Phi\|_{L_p^*(X)} \|f\|_{L_p(A)} .$$

Így Φ valóban folytonos, sőt $\|\Phi\|_{L_p(A)} \leq \|\Phi\|_{L_p^*(X)}$. A továbbiakban az egyszerűség kedvéért nem teszünk jelölésbeli különbséget f és $\tilde{f}$ között.

Minden $A \in \mathcal{M}$, $\mu(A) < \infty$ halmazra alkalmazhatjuk a már igazolt állítást. Így minden $A \in \mathcal{M}$, $\mu(A) < \infty$ halmazhoz létezik egyetlen $g_A \in L_q(A)$ függvény, melyre minden $f \in L_p(A)$ esetén

$$\Phi(f) = \int_A f g_A d\mu, \ \text{és} \ \|g_A\|_{L_q(A)} = \|\Phi\|_{L_p^*(A)} \leq \|\Phi\|_{L_p^*(X)} .$$

Az egyértelműség miatt az $A, B \in \mathcal{M}$ véges mértékű halmazokra $(g_A)_{|B} = (g_B)_{|A} = g_{A\cap B}$. Legyen

$$s = \sup \left\{ \|g_A\|_{L_q(A)} : A \in \mathcal{M},\ \mu(A) < \infty \right\} \leq \|\Phi\|_{L_p^*(X)}$$

Válasszunk olyan $A_n \in \mathcal{M}$, $\mu(A_n) < \infty$ halmaz-sorozatot, hogy $g_n = g_{A_n}$ jelöléssel

$$\|g_n\|_{L_q(A_n)} \to s.$$

Legyen $A_\infty = \cup_{n=1}^{\infty} A_n$, majd három megjegyzést teszünk az A_n halmazokra struktúrájára:

· Egyrészt feltehető, hogy $A_n \subseteq A_{n+1}$ minden $n \in \mathbb{N}$ mellett. Ugyanis ha $B_N = \cup_{n=1}^{N} A_n$, akkor $(g_{B_N})_{|A_N} = g_N$ az egyértelműség miatt. Így

$$\|g_N\|_{L_q(A_N)}^q = \int_{A_N} |g_N|^q \, d\mu = \int_{A_N} |g_{B_N}|^q \, d\mu \leq \int_{B_N} |g_{B_N}|^q \, d\mu = \|g_{B_N}\|_{L_q(B_N)}^q,$$

ezért $\|g_{B_N}\|_{L_q(B_N)} \to s$ is fennáll.

· Másrészt, ha az $(X, \mathcal{M}, \mu)$ mértéktér σ-véges akkor $A_\infty = X$ is feltehető. Ugyanis azzal az $E_n \in \mathcal{M}, \mu(E_n) < \infty$ halmaz-sorozattal, amelyre $\cup_{n=1}^{\infty} E_n = X$, bevezetve a $B_n = A_n \cup E_n$ jelölést, az előbbihez hasonló módon kapjuk, hogy

$$\|g_n\|_{L_q(A_n)} \leq \|g_{B_n}\|_{L_q(B_n)},$$

ezért $\|g_{B_n}\|_{L_q(B_n)} \to s$ is fennáll.

· Harmadrészt, $p > 1$, azaz $q < \infty$ esetén ha $B \in \mathcal{M}$, $\mu(B) < \infty$, olyan halmaz amelyre $B \cap A_\infty = \varnothing$, akkor $g_A = 0$, μ-m.m. Ugyanis ellenkező esetben $(g_{A_n \cup B})_{|A_n} = g_{A_n}$ és $(g_{A_n \cup B})_{|B} = g_B$ miatt

$$\begin{aligned}
\|g_{A_n \cup B}\|_q^q &= \int_X |g_{A_n \cup B}|^q \, d\mu = \int_{A_n} |g_{A_n \cup B}|^q \, gd\mu + \int_B |g_{A_n \cup B}|^q \, gd\mu = \\
&= \int_{A_n} |g_{A_n}|^q \, d\mu + \int_B |g_B|^q \, d\mu = \\
&= \int_X |g_{A_n}|^q \, d\mu + \int_X |g_B|^q \, d\mu = \|g_{A_n}\|_q^q + \|g_B\|_q^q > s^q
\end{aligned}$$

állna fenn elegendő nagy n mellett ellentmondva s szuprémum tulajdonságának. (Figyeljük meg, hogy $p = 1$ esetben mi történne!)

A tételben előírt $g \in L_q(X)$ függvény konstrukciója következik. A $g_n(x)$ sorozat kvázikonstans az A_n halmazok monoton növekedése miatt, tehát létezik a $g : X \to \mathbb{R}$ mérhető függvény, melyre $g_n \to g$ pontonként. Látható, hogy $g \in L_q(X)$, hiszen $|g_n|$ monoton növekedőleg tart $|g|$-hez, ezért $q < \infty$ esetben a monoton konvergenciatétel miatt $\int_X |g_n|^q \, d\mu \to \int_X |g|^q \, d\mu$. A bal oldali határérték minden tagja legfeljebb s^q, ezért $\|g\|_{L_q(X)} \leq s$. A $q = \infty$ esetben $|g(x)| = \lim_{n\to\infty} |g_n(x)| \leq \lim_{n\to\infty} \|g_n\|_\infty = s$ majdnem minden $x \in X$ mellett, tehát $\|g\|_{L_\infty(X)} \leq s$ most is teljesül.

Tekintsünk most egy olyan $f \in L_p(X)$ függvényt, amely az $A_\infty = \cup_{n=1}^\infty A_n$ halmazon kívül eltűnik. Az $f\chi_{A_n} \in L_p(A_n)$, ezért a g_n függvények definíciója szerint

$$\Phi(f\chi_{A_n}) = \int_{A_n} fg_n d\mu.$$

Ekkor a lemmát alkalmazva $f\chi_{A_n} \to f\chi_{A_\infty} = f$ az $L_p(X)$ topológiája szerint, ezért Φ folytonossága miatt

$$\Phi(f\chi_{A_n}) \to \Phi(f).$$

A Hölder-egyenlőtlenség szerint tehát $fg \in L_1(X)$. De az $(|fg_n|)_{n\in\mathbb{N}}$ sorozatnak fg majoránsa, ami ezek szerint $L_1(X)$-beli, ezért a majorált konvergenciatétel szerint

$$\int_{A_n} fg_n d\mu = \int_X fg_n d\mu \to \int_X fg d\mu.$$

A három utoljára kiemelt sort összevetve azt kapjuk, hogy

$$\Phi(f) = \int_X fg d\mu.$$

fennáll minden olyan $f \in L_p(X)$ függvény esetén, amelyre $f_{|A_\infty^c} = 0$.

Tekintsünk most egy olyan $f \in L_p(X)$ függvényt, amely az A_∞ halmazon konstans zéró. Megfogjuk mutatni, hogy $\Phi(f) = 0$.

Ha $p = 1$, akkor μ σ-végessége miatt, $A_\infty = X$, ezért ilyen függvény csak L_p zérus eleme lehet, tehát ebben az esetben $\Phi(f) = 0$ teljesül.

Ha $p > 1$, legyen $B_n = X(|f| \geq 1/n)$, és $B_\infty = \cup_{n=1}^\infty B_n$. Világos, hogy $X(f = 0) = B_\infty^c$ és $\mu(B_n) < \infty$ az $f \in L_p(X)$ miatt. A lemmát használva, kapjuk a B_n növő sorozatra, hogy $f\chi_{B_n} \to f$ az $L_p(X)$ topológiája szerint, ezért Φ folytonossága szerint

$$\Phi(f\chi_{B_n}) \to \Phi(f).$$

Másrészt $\mu(B_n) < \infty$ miatt

$$\Phi(f\chi_{B_n}) = \int_{B_n} fg_{B_n} d\mu.$$

No de a harmadik megjegyzésünk alkalmazható $p > 1$ miatt, így $(g_{B_n})_{|A_\infty^c} = 0$. A feltételünk szerint $f_{|A_\infty} = 0$ is fennáll, ezért

$$\int_{B_n} fg_{B_n} d\mu = \int_{B_n \setminus A_\infty} fg_{B_n} d\mu + \int_{B_n \cap A_\infty} fg_{B_n} d\mu = 0.$$

Összegzésként tetszőleges $f \in L_p(X)$ mellett $f\chi_{A_\infty}$ eltűnik A_∞^c-en, és $f\chi_{A_\infty^c}$ eltűnik A_∞-en. Így az előző két bekezdés szerint

$$\Phi(f) = \Phi\left(f\chi_{A_\infty} + f\chi_{A_\infty^c}\right) = \Phi\left(f\chi_{A_\infty}\right) + \Phi\left(f\chi_{A_\infty^c}\right) = \int_X fgd\mu + 0 = \int fgd\mu.$$

Ezt kellett a g egzisztenciáját követelő állításhoz belátni.

A bizonyításból már csak g unicitásának igazolása maradt hátra. Tegyük fel, hogy létezik g_1 és g_2 L_q-beli függvény, amelyre minden $f \in L_p$ mellett

$$\int_X fg_1 d\mu = \int_X fg_2 d\mu.$$

Legyen $A_n = X\left(|g_1| \geq \frac{1}{n}\right)$ és $B_n = X\left(|g_2| \geq \frac{1}{n}\right)$. Ekkor A_n és B_n véges mértékű halmazok, hiszen $g_1, g_2 \in L_q$. Így a véges halmazokon már meggondolt unicitás miatt $g_1 = g_2$ az $A_n \cup B_n$ halmazon, ezért $g_1 = g_2$ az $\cup_{n=1}^{\infty}(A_n \cup B_n)$ halmazon is. Ez utóbbi halmaz komplementerét véve kapjuk, hogy

$$\cap_{n=1}^{\infty}\left(A_n^c \cap B_n^c\right).$$

Ez utóbbi halmazon mind g_1, mind g_2 zérus. □

6.2. A sűrűségfüggvény és feltételes várható érték

A szakaszban a valószínűségszámítás legegyszerűbb fogalmait és definícióit szeretném formálisan is bevezetni. Felteszem, hogy az olvasó jól ismeri a naiv valószínűségszámítás alapfogalmait, ezért nem részletezem a valószínűségszámításbeli hátteret.

6.2.1. definíció (valószínűségi mező, valószínűségi változó, várhatóérték)**.** Egy $(\Omega, \mathcal{A}, \mu)$ mértékteret *valószínűségi mezőnek* mondunk, ha $\mu(\Omega) = 1$.

A $X : \Omega \to \mathbb{R}$ mérhető függvényt szokás *valószínűségi változónak* nevezni.

Egy X valószínűségi változóról azt mondjuk, hogy van várhatóértéke, ha integrálható, azaz $X \in L_1(\Omega, \mathcal{A}, \mu)$. A X *várható értékét* $E(X) = \int_\Omega X d\mu$ módon definiáljuk.

6.2.2. állítás (valószínűségi változó eloszlása)**.** *Tekintsünk egy $(\Omega, \mathcal{A}, \mu)$ valószínűségi mezőt, és legyen $X : \Omega \to \mathbb{R}$ egy valószínűségi változó. Minden $B \in \mathcal{B}$ Borel-halmaz esetén jelölje*

$$Q_X(B) = \mu(\Omega(X \in B)) = \mu(\{\omega \in \Omega : X(\omega) \in B\}) = \mu\left(X^{-1}(B)\right).$$

Látható, hogy az így definiált $Q_X : \mathcal{B} \to [0,1]$ egy valószínűségi mérték. Ezt a Q_X mértéket nevezzük az X valószínűségi változó eloszlásának.

6.2.3. állítás. *Legyen $g : \mathbb{R} \to \mathbb{R}$ egy nem negatív Borel-mérhető függvény, X egy valószínűségi változó, amelyre $g \geq 0$ vagy $g \circ X \in L_1(\mu)$. Ekkor*

$$\int_\Omega g(X(\omega))\,\mu(d\omega) = \int_\Omega g \circ X\, d\mu = \int_\mathbb{R} g\, dQ_X = \int_\mathbb{R} g(t)\, Q_X(dt)$$

egyenlőség.

Bizonyítás. Az első és utolsó egyenlőség egyszerűen egy a valószínűségi számítás körében szokásos jelölés, amely megadja a szóban forgó függvények változóját is. A középső egyenlőség éppen az eloszlás definíciója, ha g egy Borel-halmaz indikátor függvénye, hiszen

$$\chi_A \circ X = \chi_{X^{-1}(A)}$$

Ha g egy Borel-mérhető egyszerű függvény, akkor a μ szerinti és Q_X mértékek szerinti integrálás lineáris funkcionál volta adja az állítás, míg ha g egy nem negatív mérhető függvény, akkor a monoton konvergenciatétel szokásos alakja mutatja, hogy az állítás valóban igaz. □

Érdemes felírni az előző tétel néhány speciális esetét.

6.2.4.
Ha valamely rögzített $A \subseteq \mathbb{R}$ halmazra és $k \in \mathbb{N}$ számra $g = \chi_A \operatorname{id}^k$, akkor $g \circ X = \chi_{X^{-1}(A)} \cdot X^k$ Tehát

$$\int_{\Omega(X \in A)} X^k\, d\mu = \int_A t^k\, Q_X(dt)\,,$$

feltéve persze, hogy a fenti bal oldali integrál értelmes. Így az $A = \mathbb{R}$ speciális esetben kapjuk a várható érték és a k-adik momentumra vonatkozó formulát.

$$E\left(X^k\right) = \int_\Omega X^k\, d\mu = \int_\mathbb{R} t^k\, Q(dt)\,.$$

6.2.5. definíció (sűrűség függvény)**.** Azt mondjuk, hogy egy valószínűségi változó *abszolút folytonos* vagy *abszolút folytonos eloszlású*, ha annak eloszlása mint a számegyenes egy Borel-mértéke abszolút folytonos mérték a Lebesgue-mértéknek a Borel-halmazokra való leszűkítésére. Ha ez fennáll, akkor az eloszlás mérték Lebesgue-mérték szerinti Radon–Nikodym-deriváltját nevezzük az abszolút folytonos valószínűségi változó *sűrűség függvényének.*

6.2.6.
Jelölje $m : \mathcal{B} \to \mathbb{R}$ a Lebesgue-mérték megszorítását a számegyenes Borel-halmazaira. Ha X egy abszolút folytonos valószínűségi változó, akkor annak

f sűrűségfüggvénye azaz egyetlen $f \in L_1(\mathbb{R}, \mathcal{B}, m)$ függvény, amelyre minden $B \in \mathcal{B}$ Borel-halmaz esetén

$$\mu(\Omega(X \in B)) = \int_B f\,dm.$$

Ennek megfelelően

$$\int_{\Omega(X \in A)} X^k\,d\mu = \int_A t^k\,Q_X(dt) = \int_A t^k f(t)\,m(dt).$$

Speciálisan a várható érték és második momentum formulája:

$$E(X) = \int_{\mathbb{R}} t\,Q_X(dt) = \int_{\mathbb{R}} t f(t)\,m(dt),$$
$$E\left(X^2\right) = \int_{\mathbb{R}} t^2\,Q_X(dt) = \int_{\mathbb{R}} t^2 f(t)\,m(dt).$$

Most rátérünk a feltételes várható érték fogalmára. A lényeg itt is az lesz, mint fent, nevezetesen hogy a feltételes várható érték is egy Radon–Nikodym-derivált. Ehhez először felidézünk néhány egyszerű definíciót.

6.2.7.
Emlékezzünk vissza a feltételes valószínűségi mező fogalmára. Legyen $A \in \mathcal{A}$ egy olyan esemény, melyre $\mu(A) > 0$. Látható, hogy ekkor

$$\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}$$

egy mértéket definiál az $\mathcal{A}$ σ-algebrán, tehát $(\Omega, \mathcal{A}, \mu_A)$ is valószínűségi mező. Ezt a valószínűségi mezőt nevezzük az A pozitív valószínűségű esemény által generált *feltételes valószínűségi mezőnek*.

Persze a feltételes valószínűségi mező is csak egy valószínűségi mező, ezért az összes fent bevezetett fogalom feltételes formája is definiált mint a megfelelő fogalom a feltételes valószínűségi mezőre nézve.

6.2.8. definíció (feltételes várható érték egy eseményre nézve)**.** Legyen $(\Omega, \mathcal{A}, \mu)$ egy valószínűségi mező és $\mu(A) > 0$ egy pozitív valószínűségű esemény. Egy $X : \Omega \to \mathbb{R}, X \in L_1(\Omega, \mathcal{A}, \mu)$ valószínűségi változónak az A eseményre vonatkozó feltételes várható értéke nem más, mint az X valószínűségi változónak az $(\Omega, \mathcal{A}, \mu_A)$ téren vett várható értéke:

$$E(X|A) = \int_\Omega X\,d\mu_A.$$

6.2.9.
Látjuk, hogy $\mu_A \ll \mu$, és

$$\mu_A(B) = \frac{1}{\mu(A)} \int_B \chi_A \, d\mu,$$

azaz a μ_A mérték $\frac{d\mu_A}{d\mu}$ Radon–Nikodym-deriváltja éppen $\frac{1}{\mu(A)}\chi_A$. Ezek szerint

$$E(X|A) = \int_\Omega X \frac{1}{\mu(A)} \chi_A \, d\mu = \frac{1}{\mu(A)} \int_A X \, d\mu.$$

Ha tehát X még abszolút folytonos eloszlású is az f sűrűség függvénnyel, és az A feltételi halmaz $A = X^{-1}(B)$ alakú a számegyenes valamely $B \subseteq \mathbb{R}$ Borel-mérhető halmazára, akkor

$$E(X|\Omega(X \in B)) = \frac{1}{\mu(\Omega(X \in B))} \int_B t f(t) \; m(dt).$$

6.2.10. definíció (feltételes várható érték megszámlálható partícióra nézve)**.** Legyen $\{A_n : n \in \mathbb{N}\}$ egy pozitív valószínűségű eseményekből álló megszámlálható partíciója az $(\Omega, \mathcal{A}, \mu)$ valószínűségi mezőnek, és jelölje $\mathcal{B}$ e partíció generálta legszűkebb σ-algebrát. Adott $X \in L_1(\Omega, \mathcal{A}, \mu)$ valószínűségi változó mellett definiáljuk a

$$E(X|\mathcal{B}) := \sum_{n=1}^{\infty} E(X|A_n)\chi_{A_n}$$

függvényt. Vegyük észre, hogy a fent definiált $E(X|\mathcal{B}) : \Omega \to \mathbb{R}$ függvény rendelkezik az alábbi két fontos tulajdonsággal:

- $E(X|\mathcal{B})$ mérhető a $\mathcal{B}$ σ-algebrára nézve.
- Minden $C \in \mathcal{B}$ esemény esetén

$$\int_C E(X|\mathcal{B}) \; d\mu = \int_C X \, d\mu$$

A fenti $E(X|\mathcal{B})$ függvényt nevezzük a X valószínűségi változónak a $\mathcal{B}$ részszigma-algebrára vonatkozó *feltételes várható értékének.*

Bizonyítás. Igazolásként, azt kell meggondolnunk, hogy $\mathcal{B}$ elemei a partíció egyes elemeinek egyesítéseként állnak ellő. Mivel eleve csak megszámlálhatóan sok halmazunk van, ezért minden egyesítésben is csak megszámlálhatóan sok halmaz fordulhat elő. Ha $C = A_n$ valamelyik $n \in \mathcal{N}$ mellett, akkor $E(X|\mathcal{B}) = E(X|A_n)\,\chi_{A_n}$, így

$$\int_C E(X|\mathcal{B})\,d\mu = \int_{A_n} E(X|A_n)\,d\mu = \\ = \mu(A_n)\,E(X|A_n) = \int_{A_n} X\,d\mu = \int_C X\,d\mu.$$

Ha C a néhány partíció elem egyesítése, akkor az állítás abból következik, hogy rögzített L_1-beli függvény vagy rögzített nem negatív függvény melletti integrál-mérték valóban egy mérték. □

6.2.11. definíció (feltétles várható érték egy részszigmaalgebrára nézve)**.** Legyen $(\Omega, \mathcal{A}, \mu)$ egy valószínűségi mező, $\mathcal{B} \subseteq \mathcal{A}$ egy részszigma-algebra és $X \in L_1(\Omega, \mathcal{A}, \mu)$ egy valószínűségi változó. Ekkor létezik egyetlen olyan $E(X|\mathcal{B}) \in L_1(\Omega, \mathcal{B}, \mu)$ függvény, amelyre az alábbi két tulajdonság egyszerre teljesül:

- $E(X|\mathcal{B}) : \Omega \to \mathbb{R}$ mérhető a $\mathcal{B}$ σ-algebrára nézve.
- Minden $C \in \mathcal{B}$ esemény esetén

$$\int_C E(X|\mathcal{B})\,d\mu = \int_C X\,d\mu.$$

A fenti $E(X|\mathcal{B})$ függvényt nevezzük a X valószínűségi változónak a $\mathcal{B}$ részszigma-algebrára vonatkozó *feltételes várható értékének.*

Bizonyítás. Jelölje $\nu(C) = \int_C X\,d\mu$. Világos, hogy $\nu : \mathcal{B} \to [0,1]$ egy mérték, amely abszolút folytonos μ-re nézve. A Radon–Nikodym-tételt alkalmazzuk az $(X, \mathcal{B})$ mérhető téren a ν előjeles mértékre és a μ véges nem negatív mértékre. Eszerint létezik egyetlen $E(X|\mathcal{B})$-vel jelölt függvény, amelyre

$$E(X|\mathcal{B}) \in L_1(\Omega, \mathcal{B}, \mu) \quad \text{és} \quad \int_C X\,d\mu = \nu(C) = \int_C E(X|\mathcal{B})\,d\mu$$

minden $C \in \mathcal{B}$ mellett. Ezt kellett belátni. □

6.2.12.
A várható érték definíciójának $C = \Omega$ speciális esetére kapjuk az

$$E(E(X|\mathcal{B})) = E(X)$$

azonosságot. Persze $\mathcal{B} = \mathcal{A}$ esetben semmi nem történik, így $E(X|\mathcal{A}) = X$.

Fontos látnunk, hogy a Radon–Nikodym-tétel unicitása szerint a fent definiált feltételes várható érték fogalma egybeesik azzal a fogalommal, amit a megszámlálható partíció által generált σ-algebrára esetén vezettünk be. Természetesen mint L_1-beli függvénynek az unicitására kell gondolnunk, ami azt jelenti, hogy a feltételes várható érték is csak null-mértékű halmazoktól eltekintve egyértelműen meghatározott valószínűségi változó.

6.2.13.
Most nézzük azt az esetet, amikor egy tetszőleges számosságú Γ indexhalmazzal melletti $\{A_\gamma : \gamma \in \Gamma\}$ partícióból kiindulva $\mathcal{B}$ a generált σ-algebra. Gondoljuk meg, hogy $\mathcal{B}$-nek nincs olyan nem üres halmaza, amely valódi részhalmaza lehetne az egyik A_γ partíció elemnek. Ez azt jelenti, hogy egy $\mathcal{B}$ mérhető függvény csak egyetlen értéket vehet fel az egyik A_γ halmaz felett, tehát a partíció egyes részhalmazai felett konstans. Legyen X egy véges várható értékű valószínűségi változója az $(\Omega, \mathcal{A}, \mu)$ mezőnek, és $E(X|\mathcal{B})$ a fenti partícióra vonatkozó feltételes várható érték. Ha α_γ jelöli az $E(X|\mathcal{B})$ függvény A_γ halmaz felett felvett értékét, akkor

$$\alpha_\gamma \mu(A_\gamma) = \int_{A_\gamma} E(X|\mathcal{B})\, d\mu = \int_{A_\gamma} X\, d\mu = \mu(A_\gamma)\, E(X|A_\gamma)$$

a feltételes várható érték definíciója szerint. Azt gondoltuk meg tehát, hogy

$$E(X|\mathcal{B})(\omega) = E(X|A_\gamma)$$

minden $\omega \in A_\gamma$ mellett feltéve persze, hogy $\mu(A_\gamma) > 0$.

Szívünkben örömmel eltelve[1] látjuk tehát, hogy Γ megszámlálható esetben a Radon–Nikodym-tétel garantálta feltételes várható érték valóban azonos a $\sum_{n=1}^{\infty} E(X|A_n)\, \chi_{A_n}$ definícióval. Még speciálisabban, ha Γ összesen két elemű $\{A, A^c\}$, akkor a generált σ-algebra $\mathcal{B} = \{\varnothing, A, A^c, \Omega\}$ és az $E(X|\mathcal{B})$ feltételes várható érték függvénynek két értéke van. Az A esemény felett $E(X|A)$, az A^c esemény felett $E(X|A^c)$ feltéve, hogy $0 < \mu(A) < 1$. Amennyiben A és A^c események egyike null-valószínűségű, mondjuk A^c, akkor csak annyit lehet mondani, hogy $E(X|\mathcal{B})(\omega) = E(X|A) = E(X)$ az A esemény felett, és az értéke tetszőleges valós szám is lehet, tehát nem releváns az A^c null-mértékű halmaz felett, hiszen mint Radon–Nikodym-derivált ez egy L_1-beli függvényosztály egyik reprezentánsa.

[1] :)

II. rész

Dinamikus programozás

7. fejezet

Determinisztikus eset

7.1. Elnevezések, jelölések

Tegyük fel, hogy adott egy

állapothalmaz: egy X halmaz, amelynek elemeit állapotoknak mondjuk;

átmenetfüggvény: egy $\Gamma : X \multimap X$ halmazértékű leképezés, amely azt írja le, hogy az x állapotból a $\Gamma(x)$ állapotok egyikébe juthatunk. E függvény gráfja: $\operatorname{graph}(\Gamma) = \{(x, y) \in X \times X : y \in \Gamma(x)\}$;

hozamfüggvény: egy $F : \operatorname{graph}(\Gamma) \to \mathbb{R}$ függvény, amelyre $F(x, y) \in \mathbb{R}$ számot az x állapotból az y állapotba jutás hozadékának, profitjának, értékének, vagy pillanatnyi hasznosságának gondolunk;

diszkont tényező: egy $\beta \geq 0$ valós szám.

7.1.1. (SP feladat)
A fenti négy objektum határozza meg a dinamikus programozás alapfeladatát, amelyet *szuprémum feladatnak* fogunk mondani és röviden *(SP) feladatként* fogunk rá hivatkozni. Az (X, F, Γ, β) négyest *dinamikus programozási alapfeladatnak*, *dinamikus programozásnak* vagy rövidebben *programozásnak* nevezzük.

Adott $\Gamma : X \multimap X$ mellett

megengedett út: $\bar{x} : \mathbb{N} \to X$ olyan sorozat, amelyre $x_{k+1} \in \Gamma(x_k)$ teljesül minden $k \in \mathbb{N}$ mellett. Nevezzük ezt *megengedett pályának*, *megengedett tervnek*, vagy egyszerűbben csak *tervnek* vagy *útnak* is. A megengedett utak halmazát Π-vel jelöljük. Ha $\bar{x} \in \Pi$ egy út, akkor $\bar{x}_k \in \Pi$ jelöli az $(x_k, x_{k+1}, \cdots)$ utat. Rögzített $u \in X$ mellett $\Pi(u) = \{\bar{x} \in \Pi : x_0 = u\}$ az u állapotból induló megengedett pályák halmaza.

Út hasznossága: Az $u : \Pi \to \overline{\mathbb{R}}$ függvényt az *utak hasznossági függvényének*, vagy az *út értékének*, vagy az *út hasznának*, vagy az *út profitjának* mondjuk, ahol $\bar{x} \in \Pi$ útra

$$u(\bar{x}) = \sum_{k=0}^{\infty} \beta^k F(x_k, x_{k+1}).$$

Mint az (SP) feladat egy alapfeltevése, azt tesszük fel, hogy az (X, F, Γ, β) programozás úgy van megadva, hogy a fenti sor általános értelemben konvergens, azaz konvergál egy valós számhoz vagy tart $\pm\infty$-hez.

7.2. Alapfeladat és Bellman-egyenlet

A dinamikus programozás alapfeladatát szuprémum problémának is mondjuk. Szuprémum problémán egy feladatot értünk. Azt a feladatot, hogy valamely rögzített állapot esetén keressük az ebből az állapotból induló utak hasznosságának felső határát.

7.2.1. definíció (Szupremum probléma. (SP))**.** Az (X, F, Γ, β) által meghatározott szuprémum feladat azt jelenti, hogy rögzített $x_0 \in X$ állapot mellett keressük az u függvény szuprémumát a $\Pi(x_0)$ utak halmaza felett:

$$w(x_0) = \sup_{\bar{x} \in \Pi(x_0)} u(\bar{x}) = \sup_{\bar{x} \in \Pi(x_0)} \sum_{k=0}^{\infty} \beta^k F(x_k, x_{k+1}).$$

A szuprémum $w(x_0)$ értékét az x_0 pontból kiinduló *(SP) feladat értékének* fogjuk mondani, és a $w : X \to \overline{\mathbb{R}}$ függvényt pedig a *(SP) feladat értékfüggvényének* mondjuk.

Amennyiben a fenti sup helyett max írható, azaz ha létezik $\bar{x} \in \Pi(x_0)$, hogy $w(x_0) = u(\bar{x})$, akkor azt mondjuk, hogy $\bar{x}$ az x_0 *állapotból induló optimális út*.

7.2.2. definíció (Bellman-egyenlet)**.** Tekintsük az (X, F, Γ, β)-hoz tartozó dinamikus programozási alapfeladatot. Azt mondjuk, hogy a $v : X \to \mathbb{R}$ függvény kielégíti a feladat *Bellman-egyenletét*, ha minden $x_0 \in X$ mellett

$$v(x_0) = \sup_{y \in \Gamma(x_0)} \{F(x_0, y) + \beta v(y)\}$$

7.2.3. lemma (iterációs lemma)**.** *Legyen $\omega : X \to \overline{\mathbb{R}}$ olyan függvény, amely egy $\bar{x} \in \Pi$ útra minden $k \in \mathbb{N}$ mellett teljesíti az*

$$\omega(x_k) = F(x_k, x_{k+1}) + \beta\omega(x_{k+1}) + \delta_k$$

egyenleteket, valamely $\delta_0, \delta_1, \cdots$ *valós sorozat mellett. Ekkor minden* $n \in \mathbb{N}$*-re:*

$$\omega(x_0) = \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \beta^{n+1}\omega(x_{n+1}) + \sum_{k=0}^{n} \beta^k \delta_k.$$

Ha a feltételi egyenlőségben az $=$ *relációt a* $\leq$ *vagy* $\geq$ *relációkra cseréljük, akkor ugyanazt kell tennünk az implikáció egyenlőségében.*

Bizonyítás. Teljes indukció. $n = 0$-ra az állítás a feltételre egyszerűsödik $k = 0$ esetben. Tegyük fel, hogy igaz n-re. Ekkor a feltételt alkalmazva $k = n + 1$-re azt kapjuk, hogy

$$\begin{aligned}\omega(x_0) &= \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \beta^{n+1}\omega(x_{n+1}) + \sum_{k=0}^{n} \beta^k \delta_k = \\ &= \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \\ &\qquad + \beta^{n+1}\big(F(x_{n+1}, x_{n+2}) + \beta\omega(x_{n+2}) + \delta_{n+1}\big) + \sum_{k=0}^{n} \beta^k \delta_k = \\ &= \sum_{k=0}^{n+1} \beta^k F(x_k, x_{k+1}) + \beta^{n+2}\omega(x_{n+2}) + \sum_{k=0}^{n+1} \beta^k \delta_k.\end{aligned}$$

Ezt kellett belátni. □

7.2.4. lemma (indukciós lemma)**.** *Legyen* $\bar{x} \in \Pi$ *egy út. Ekkor az út* u *hasznossági függvényére:*

$$u(\bar{x}) = F(x_0, x_1) + \beta u(\bar{x}_1).$$

Sőt minden $n \in \mathbb{N}$ *mellett fennáll a következő egyenlőség:*

$$u(\bar{x}) = \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \beta^{n+1} u(\bar{x}_{n+1}).$$

Bizonyítás. Az egynél magasabb indexű tagokból a diszkont tényező kiemelésével, majd a szumma elcsúsztatásával:

$$\begin{aligned}u(\bar{x}) &= F(x_0, x_1) + \beta \sum_{k=1}^{\infty} \beta^{k-1} F(x_k, x_{k+1}) = \\ &= F(x_0, x_1) + \beta \sum_{k=0}^{\infty} \beta^k F(x_{k+1}, x_{k+2}) = F(x_0, x_1) + \beta u(\bar{x}_1).\end{aligned}$$

Persze minden $k \in \mathbb{N}$ mellett $\bar{x}_k$ is csak egy út, így a már bizonyított állítást $\bar{x}_k$-ra alkalmazva kapjuk, hogy $u(\bar{x}_k) = F(x_k, x_{k+1}) + \beta u(\bar{x}_{k+1})$. Alkalmazhatjuk tehát az $x_k \mapsto u(\bar{x}_k)$ függvényre az imént bizonyított iterációs lemmát $\delta_k = 0$ konstansok mellett. Ezt kellett belátni. □

7.2.5. állítás (optimális út jellemzése)**.** *Legyen* (X, F, Γ, β) *egy dinamikus programozás.*

1. *Tegyük fel, hogy* $\bar{x} \in \Pi(x_0)$ *a feladathoz tartozó optimális út. Ha* $\beta > 0$*, akkor* $\bar{x}_k \in \Pi(x_k)$ *az* x_k*-ból induló optimális út, továbbá az út minden* x_k *állapotára*
$$w(x_k) = F(x_k, x_{k+1}) + \beta w(x_{k+1}).$$
2. *Megfordítva, ha egy* $\bar{x} \in \Pi$ *útra fennáll a fenti egyenlőség minden* $k \in \mathbb{N}$ *mellett és*
$$\limsup_{k\to\infty} \beta^k w(x_k) \leq 0,$$
akkor $\bar{x}$ *az* x_0*-ból induló szuprémum feladatnak optimális útja.*

A 7.3.1. „fő a változatosság"-modellben látni fogjuk, hogy $\beta > 0$ feltétel az állítás első két pontjában nem hagyható el. Hasonlóan, a 7.3.3. „megtakarítás"-modell azt mutatja, hogy az elegendőséghez nélkülözhetetlen a fenti többlet feltevés.

Bizonyítás. Megmutatjuk, hogy $\bar{x}_1$ optimális x_1-ből induló út. Ha ugyanis – indirekt – $w(x_1) > u(\bar{x}_1)$, akkor létezik $\bar{y} \in \Pi(x_1)$, amelyre $u(\bar{y}) > u(\bar{x}_1)$. Legyen ekkor $\bar{z} = (x_0, \bar{y}), \bar{z} \in \Pi(x_0)$. A 7.2.4. indukciós lemma és $\beta > 0$ szerint

$$u(\bar{z}) = F(x_0, x_1) + \beta u(\bar{y}) > F(x_0, x_1) + \beta u(\bar{x}_1) = u(\bar{x}) = w(x_0),$$

ami ellentmond $w(x_0)$ felső korlát voltának. Megmutattuk tehát, hogy $\bar{x}_1$ az x_1-ből induló optimális út. Ezt alkalmazva $\bar{x}_1$-re kapjuk, hogy $\bar{x}_2$ az x_2-ből induló optimális út. Stb. Így az indukciós lemmát és az eddig igazolt állítást alkalmazva minden $k \in \mathbb{N}$ mellett:

$$w(x_k) = u(\bar{x}_k) = F(x_k, x_{k+1}) + \beta u(\bar{x}_{k+1}) = F(x_k, x_{k+1}) + \beta w(x_{k+1}).$$

Most nézzük a fordított állítást. A 7.2.3. iterációs lemma speciális esete szerint minden $n \in \mathbb{N}$-re

$$w(x_0) = \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \beta^{n+1} w(x_{n+1}).$$

Így ha $n \to \infty$, akkor $w(x_0) = u(\bar{x}) + \limsup_{n\to\infty} \beta^n w(x_n) \leq u(\bar{x}) \leq w(x_0)$. Ezt kellett belátni. □

7.2.6. definíció (a programozás op-függvénye)**.** Adott (X, F, Γ, β) programozás mellett egy $G(x) \subseteq \Gamma(x)$ halmazértékű függvényt *policy függvénynek* mondunk. Azt mondjuk, hogy az $\bar{x} \in \Pi$ megengedett út a *G policy függvény által generált*, ha $x_{k+1} \in G(x_k)$ teljesül minden k mellett. Ha w az (X, F, Γ, β) program értékfüggvénye, akkor a

$$G^*(x) = \{y \in \Gamma(x) : w(x) = F(x, y) + \beta w(y)\}$$

policy függvényt, *optimal policy* függvénynek, vagy[1] *op-függvénynek* mondjuk.

Ha többet nem teszünk fel, akkor lehetséges, hogy $G^*(x) = \varnothing$. Az alábbiakban csak az eddig bizonyított állításokat fogalmazom meg a bevezetett policy függvény szóhasználatára támaszkodva.

7.2.7.
Azt igazoltuk, hogy

1. ha $\bar{x} \in \Pi(x_0)$ terv egy az x_0 állapotból induló optimális pálya és $\beta > 0$, akkor a tervet a feladat op-függvénye generálja;
2. ha $\bar{x}$ egy az op-függvény által generált pálya, és $\limsup \beta^k w(x_k) \le 0$, akkor $\bar{x}$ az x_0 állapotból induló optimális út.
3. Ha tehát feltesszük, hogy az F pillanatnyi hasznosság korlátos, és $0 <$ $< \beta < 1$, úgy $\bar{x}$ akkor és csak akkor optimális pálya, ha azt a feladat op-függvénye generálja!

7.2.8. állítás ((SP) értékfüggvénye kielégíti a Bellman-egyenletet)**.** *Tekintsük az (X, F, β, Γ) feladatot és annak w értékfüggvényét. Ekkor minden $x_0 \in$ $\in X$ állapot esetén*

$$w(x_0) = \sup_{y \in \Gamma(x_0)} \{F(x_0, y) + \beta w(y)\}.$$

Bizonyítás. Ha $\beta = 0$, akkor minden $\bar{x} \in \Pi(x_0)$ mellett $u(\bar{x}) = F(x_0, x_1)$, ezért

$$w(x_0) = \sup_{\bar{x} \in \Pi(x_0)} u(\bar{x}) = \sup_{\bar{x} \in \Pi(x_0)} F(x_0, x_1) = \sup_{y \in \Gamma(x_0)} F(x_0, y).$$

A továbbiakban tehát $\beta \neq 0$ feltehető.

–1– Ha $w(x_0) = -\infty$, akkor minden $\bar{x} \in \Pi(x_0)$ pálya esetén $-\infty = u(\bar{x}) =$ $= F(x_0, x_1) + \beta u(\bar{x}_1)$, ami csak úgy lehet, ha minden $u(\bar{x}_1) = -\infty$, tehát $w(y) = -\infty$ minden $y \in \Gamma(x_0)$ mellett.

[1] hogy ne ennyire sértse nyelvérzékünket:)

–2– Ha $w(x_0) \in \mathbb{R}$, akkor adott $\varepsilon > 0$-hoz létezik $\bar{x} \in \Pi(x_0)$, amelyre

$$w(x_0) - \varepsilon < u(\bar{x}) = F(x_0, x_1) + \beta u(\bar{x}_1) \leq F(x_0, x_1) + \beta w(x_1) \leq$$
$$\leq \sup_{y \in \Gamma(x_0)} \{F(x_0, y) + \beta w(y)\}$$

az $x_1 \in \Gamma(x_0)$ szerint. De ε tetszőlegesen kicsi lehet, emiatt a bal oldal legfeljebb akkora, mint a jobb oldal.

Most rögzítsünk egy $y \in \Gamma(x_0)$ állapotot. Legyen $\varepsilon > 0$. Ekkor $w(y) - \varepsilon$-hoz létezik $\bar{y} \in \Pi(y)$ út, amelyre $w(y) - \varepsilon < u(\bar{y})$. Legyen $\bar{x} = (x_0, \bar{y}) \in \Pi(x_0)$. Ekkor az indukciós lemma szerint

$$F(x_0, y) + \beta w(y) \leq F(x_0, y) + \beta u(\bar{y}) + \beta\varepsilon = u(\bar{x}) + \beta\varepsilon \leq w(x_0) + \beta\varepsilon.$$

Újra kihasználva, hogy ε tetszőleges, kapjuk a fordított irányú egyenlőtlenséget.

–3– Ha $w(x_0) = +\infty$, akkor tegyük fel indirekt, hogy van olyan K konstans, amelyre minden $y \in \Gamma(x_0)$ mellett $F(x_0, y) + \beta w(y) < K$. Ekkor minden $\bar{x} \in \Pi(x_0)$-ra

$$u(\bar{x}) = F(x_0, x_1) + \beta u(\bar{x}_1) \leq F(x_0, x_1) + \beta w(x_1) \leq K,$$

ami ellentmond a $w(x_0) = \infty$ feltételnek.

Ezt kellett belátni. □

7.2.9. állítás (Bellman-egyenlet megoldása az (SP) értékfüggvénye)**.** *Tegyük fel, hogy a $v : X \to \overline{\mathbb{R}}$ függvény kielégíti az (X, F, β, Γ) alap feladatának Bellman-egyenletét és minden $\bar{x} \in \Pi$ terv esetén*

$$\lim_{n\to\infty} \beta^n v(x_n) = 0.$$

Ekkor a Bellman-egyenletnek ez az egyetlen megoldása és ez a v megoldás egyben a szuprémum feladat értékfüggvénye.

A 7.3.2. „befektetés-megtakarítás”-modellben példát adunk olyan problémára, ahol a Bellman-egyenletnek több megoldása is van és ezek persze eshetnek egybe a (SP) feladat egyértelműen meghatározott érték függvényével.

Bizonyítás. Legyen $x_0 \in X$. Ekkor minden $\bar{x} \in \Pi(x_0)$ mellett $x_1 \in \Gamma(x_0)$, $x_2 \in \Gamma(x_1), \cdots$ miatt, és mivel v kielégíti a Bellman-egyenletet

$$v(x_k) \geq F(x_k, x_{k+1}) + \beta v(x_{k+1})$$

minden k-ra. Így az a 7.2.3. iterációs lemma szerint

$$v(x_0) \geq \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \beta^{n+1} v(x_{n+1}).$$

Kihasználva tehát az állítás feltételét $v(x_0) \geq u(\bar{x})$ fennáll minden $\bar{x} \in \Pi(x_0)$ terv mellett, ami azt jelenti, hogy $v(x_0)$ szám a szuprémum feladat felső korlátja, ergo $v(x_0) \geq w(x_0)$, ahol $w(x_0)$ a korábbiakhoz hasonlóan az (SP) feladat értékfüggvénye.

–1– Ha $v(x_0) = -\infty$, akkor $-\infty = v(x_0) \geq w(x_0)$ miatt $w(x_0) = v(x_0)$ valóban fennáll.

–2– Ha $v(x_0) \in \mathbb{R}$, akkor legyen $\varepsilon > 0$. Most konstruálunk $\bar{x} \in \Pi(x_0)$ utat, amelyre $v(x_0) - \varepsilon < u(\bar{x})$, ami azt jelenti, hogy $v(x_0)$ a szuprémum feladat legkisebb felső korlátja. Legyenek a $\delta_k \leq 0$ számok olyanok, hogy $\sum_{k=0}^{\infty} \beta^k \delta_k < \varepsilon$.
δ_0-hoz és x_0-hoz létezik $x_1 \in \Gamma(x_1)$, hogy

$$v(x_0) < F(x_0, x_1) + \beta v(x_1) + \delta_0$$

δ_1-hez és x_1-hez létezik $x_2 \in \Gamma(x_1)$, hogy

$$v(x_1) < F(x_1, x_2) + \beta v(x_2) + \delta_1$$

δ_2-höz és x_2-höz létezik $x_3 \in \Gamma(x_1)$, hogy

$$v(x_2) < F(x_2, x_3) + \beta v(x_3) + \delta_2$$

stb. Így olyan $\bar{x} \in \Pi(x_0)$ sorozatot kapunk, amelyre teljesül a 7.2.3. iterációs lemma feltétele a $\leq$ reláció mellett. Ezért minden n-re

$$v(x_0) \leq \sum_{k=0}^{n} \beta^k F(x_k, x_{k+1}) + \beta^{n+1} v(x_{n+1}) + \sum_{k=0}^{n} \beta^k \delta_k \to u(\bar{x}) + \sum_{k=0}^{\infty} \beta^k \delta_k,$$

ami azt jelenti, hogy $v(x_0) < u(\bar{x}) + \varepsilon$. Ezt kellett belátni.

–3– Ha $v(x_0) = +\infty$, akkor először is azt kell látnunk, hogy van olyan $x_1, ..., x_N$ véges hosszúságú megengedett út ($x_{i+1} \in \Gamma(x_i)$), amelyre $v(y) \in \mathbb{R}$ minden $y \in \Gamma(x_N)$ mellett. Ha ugyanis ilyen nem lenne, akkor könnyen konstruálhatnánk olyan $\bar{x} \in \Pi(x_0)$ utat, amelyre $v(x_n) \notin \mathbb{R}$ minden n mellett, ami ellentmond a tétel $\beta^n v(x_n)$ konvergenciájára tett feltételének.
Rögzítsünk tehát egy ilyen $x_0, x_1, \ldots, x_N$ az x_0-ból induló véges sorozatot. Legyen $a \in \mathbb{R}$ egy valós szám. Most megmutatjuk azt, hogy a sorozat folytatható olyan módon, hogy a kapott $\bar{x} \in \Pi(x_0)$ útra $u(\bar{x}) > a$ teljesüljön, ami éppen azt jelenti, hogy $w(x_0) = +\infty$ is teljesül.
Ehhez az első észrevétel, hogy minden olyan megengedett útra, amely folytatása az $x_0, x_1, \cdots, x_N$-nak

$$\begin{aligned} u(\bar{x}) &= \sum_{k=0}^{N-1} \beta^k F(x_k, x_{k+1}) + \beta^N u(\bar{x}_N) = \\ &= \sum_{k=0}^{N-1} \beta^k F(x_k, x_{k+1}) + \beta^N \left(F(x_N, x_{N+1}) + \beta u(\bar{x}_{N+1}) \right). \end{aligned}$$

Válasszuk ezért az $x_{N+1} \in \Gamma(x_N)$ állapotot olyannak, hogy

$$a + 1 < \sum_{k=0}^{N-1} \beta^k F(x_k, x_{k+1}) + \beta^N \left(F\left(x_N, x_{N+1}\right) + \beta v(x_{N+1})\right)$$

teljesüljön. Ez $v(x_N) = \sup_{y\in\Gamma(x_N)} \{F(x_N, y) + \beta v(y)\} = \infty$ miatt megtehető. De $v(x_{N+1})$ valós szám, és azt már láttuk, hogy ekkor $v(x_{N+1}) = w(x_{N+1})$, így átrendezve kapjuk, hogy:

$$a < \sum_{k=0}^{N} \beta^k F(x_k, x_{k+1}) + \beta^{N+1} \left(w(x_{n+1}) - 1/\beta^{N+1}\right).$$

Itt $w(x_{N+1})$ a választott x_{N+1} pontbeli (SP) feladat megoldása, ami azt jelenti, hogy az eredeti $x_0, \cdots, x_N$ véges sorozatnak létezik olyan folytatása, amelyre $w(x_{N+1}) - \frac{1}{\beta^{N+1}} < u(\bar{x}_{N+1})$. A fent kiemelt egyenlőtlenséget tovább írva azt kapjuk, hogy

$$a < \sum_{k=0}^{N} \beta^k F(x_k, x_{k+1}) + \beta^{N+1} u\left(\bar{x}_{N+1}\right) = u(\bar{x}).$$

Ezt kellett belátni. □

7.2.10.
Ha v olyan megoldása az (X, F, Γ, β) program Bellman-egyenletének, amely az (SP) feladat megoldása is, akkor

$$G^*(x) = \operatorname{argmax}_{y\in\Gamma(x)} \{F(x, y) + \beta v(y)\}.$$

7.3. Illusztráció

A szakasz három egyszerű példát tartalmaz mint az eddig elhangozottak illusztrációját.

7.3.1. (fő a változatosság-modell)
Tegyük fel, hogy két állapot közt választhatunk. Az állapotváltás azonnali haszna 1 *egység, a helyben maradásnak nincs haszna. A diszkont faktor* $0 \leq \beta$.

Formalizáljuk a (SP) feladatot, írjuk fel a megoldás függvényt, azaz az egyes állapotok diszkontált hasznát, majd írjuk fel és oldjuk meg a Bellman-egyenletet.

Az X állapot tér két elemű: $X = \{u, v\}$. Mindkét állapotból mindkét állapot választható, ergo $\Gamma(x) = X$ minden $x \in X$ mellett. A pillanatnyi hasznosság függvényre $F(x, y) = 1$, ha $x \neq y$ és persze $F(x, x) = 0$.

A $\sum_{k=0}^{\infty} F(x_k, x_{k+1})\beta^k$ sor értéke akkor a legnagyobb, ha a β^k értékek együttható mind 1-ek. Ez azt jelenti, hogy bármely állapotból indulva az optimális terv az állapotváltogatás. Tehát a geometriai sor összegképlete szerint

$$w(x) = w(y) = \sum_{k=0}^{\infty} \beta^k = \frac{1}{1-\beta}.$$

Azt kaptuk, hogy a (SP) feladat megoldása az $w = \frac{1}{1-\beta}$ konstans függvény.

Most nézzük a Bellman-egyenletet: Az $f : X \to \mathbb{R}$ függvény megoldása a Bellman-egyenletnek, ha

$$f(v) = \max\{\beta f(v), 1 + \beta f(u)\} \quad \text{és} \quad f(u) = \max\{\beta f(u), 1 + \beta f(v)\}.$$

Ez azzal ekvivalens, hogy $f(v) = 1 + \beta f(u)$ és $f(u) = 1 + \beta f(v)$. Ennek megoldása $f(v) = 1 + \beta(1 + \beta f(v)) = 1 + \beta + \beta^2 f(v)$, amiből

$$f(v) = \frac{1}{1-\beta} \quad \text{és} \quad f(u) = 1 + \frac{\beta}{1-\beta} = \frac{1}{1-\beta}.$$

Nézzük azt az esetet, amikor $\beta = 0$. Ekkor minden olyan út optimális, amelyik állapotváltással kezd. Ez azt jelenti, hogy $\beta = 0$ esetben nem igaz az az állítás, hogy az optimális út annak minden állapotából indulva is optimális marad. Vessük ezt össze a 7.3.1. állítás feltételével. Láthatjuk, hogy $\beta > 0$ feltétel nem elhagyható!

A következő modell minden periódusában befektetés vagy megtakarítás nagyságáról kell döntenünk, és e döntés után a periódusban rendelkezésre álló vagyonunkat teljesen elfogyasztjuk. A 7.2.9. állítás feltételrendszerét fogjuk elemezni, amely szerint a Bellman-egyenletnek csak egyetlen függvény lehet a megoldása.

7.3.2. (befektetés-megtakarítás-modell)
Az induló állapot $x_0 \in \mathbb{R}$ kezdőkészlet mint vagyon megadását jelenti. Minden periódusban tetszőlegesen nagy kölcsön felvehető vagy akár az egész vagyon befektethető. A kölcsön felvételével növelt vagyon, vagy a vagyon egy részének vagy egészének befektetése utáni része minden periódusban teljesen elfogyasztásra kerül. A felvett kölcsönt vagy a befektetett vagyont α kamatláb alkalmazása mellett a következő periódusban vissza kell adni vagy kapni. Célunk a fogyasztás maximalizálása $\beta = \frac{1}{1+\alpha}$ diszkont faktor mellett.

Formalizáljuk a fenti modellt mint dinamikus programozási alapfeladatot, írjuk fel a (SP) feladatot és annak megoldásfüggvényét, írjuk fel a Bellman-egyenletet, és próbáljuk az összes megoldást megtalálni.

$X = \mathbb{R}$ tehát az állapot tér a peridusonkénti lehetséges vagyonok nagyságát modellezi. Legyen z_0 a kiinduló periódusban felvett kölcsön, mikor $z_0 > 0$;

vagy legyen z_0 a befektetett vagyon, amennyiben $z_0 < 0$. A kiinduló periódus fogyasztása $c_0 = x_0 + z_0 \geq 0$.

Az első periódusban a vagyon $x_1 = -(1+\alpha)z_0$ és a fogyasztás $c_1 = x_1 + z_1$, ahol olyan nagy z_1 kölcsönt kell felvenni vagy legfeljebb akkora befektetést lehet eszközölni, hogy $c_1 \geq 0$ teljesüljön.

Ha $x_0, x_1, \ldots, x_{t-1}$ vagyon, $z_0, z_1, \ldots, z_{t-1}$ kölcsön vagy befektetés és $c_0, c_1, \ldots, c_{t-1}$ fogyasztás már definiált, akkor $x_t = -(1+\alpha)z_{t-1}$, $c_t = x_t + z_t$, ahol z_t olyan, hogy $c_t \geq 0$ teljesül.

Célunk a

$$\sum_{k=0}^{\infty} c_k \beta^k$$

szám maximalizálása.

A következő

$$z_t = \frac{-1}{1+\alpha} x_{t+1} = -\beta x_{t+1}$$

egyenlőség szerint a vagyon értékek és a kölcsön-befektetés értékek egyértelműen meghatározzák egymást, elegendő tehát mondjuk a vagyonok X halmazán választani a következő periódusban szereplő vagyon változót. Ekkor

$$c_t = x_t + z_t = x_t - \beta x_{t+1},$$

tehát az x vagyon y-ra változtatásának pillanatnyi hasznossága

$$F(x, y) = x - \beta y.$$

Mivel legalább akkora kölcsönt fel kell venni minden periódusban, hogy a fogyasztás ne negatív legyen, ezért $c_t = x_t - \beta x_{t+1} \geq 0$. Ez azt jelenti, hogy $x_{t+1} \leq \frac{x_t}{\beta}$, ezért az átmenetfüggvényt érdemes

$$\Gamma(x) = \left(-\infty, \frac{x}{\beta}\right]$$

módon felírni.

Minél kisebb $x_{t+1} \in (-\infty, x_t/\beta]$ vagyon választása, annál nagyobb pillanatnyi fogyasztást jelent, hiszen $c_t = x_t - \beta x_{t+1}$. Pozitív x_{t+1} választása az adott periódusban $z_t = -\beta x_{t+1}$ szerint az adott periódusban a megtakarítás melletti döntést jelenti. A maximális $x_{t+1} = x_t/\beta$ választása az adott periódusban meglévő vagyon teljes megtakarítását jelenti, ami $c_t = x_t - \beta x_{t+1} = x_t - x_t = 0$ szerint a zérus fogyasztás választása.

A fent definiált (X, F, Γ, β) által meghatározott (SP) feladat jelenti a modellbeli fogyasztás maximális értékének megkeresését!

Látható, hogy tetszőleges $x_0 \in \mathbb{R}$ mellett $w(x_0) = \infty$ a (SP) megoldás függvénye, hiszen minden $c_t \geq 0$ és már az első tag $c_0 = x_0 + z_0$ is tetszőlegesen nagyra megválasztható. Minden olyan $\bar{x} \in \Pi(x_0)$ optimális pálya, amelyre a

$$\sum_{k=0}^{\infty} (x_k - \beta x_{k+1}) \beta^k$$

sor divergens. Írjuk fel az n-edik részletösszeget:

$$\begin{aligned} s_n = \sum_{k=0}^{n} (x_k - \beta x_{k+1}) \beta^k = \\ &= (x_0 - \beta x_1) + (x_1 - \beta x_2) \beta + (x_2 - \beta x_3) \beta^2 + \cdots + (x_n - \beta x_{n+1}) \beta^n = \\ &= x_0 - \beta^{n+1} x_{n+1} \end{aligned} \tag{7.1}$$

Ezek szerint az optimális pályát nagyon egyszerű megtalálni. Tetszőleges olyan $\bar{x} \in \Pi(x_0)$ optimális, melyre $\beta^{t+1} x_{t+1} \to -\infty$, ami azzal ekvivalens, hogy $\beta^t z_t \to \infty$ teljesüljön.

Nézzük a Bellman-egyenletet. A v függvény megoldása a Bellman-egyenletnek, ha

$$v(x) = \sup_{y \le x/\beta} \{x - \beta y + \beta v(y)\}$$

teljesül minden $x \in \mathbb{R}$ mellett. Próbáljuk megoldani ezt a függvényegyenletet. Ekvivalens átalakításokkal

$$v(x) - x = \beta \sup_{y \le x/\beta} \{v(y) - y\},$$

bevezetve tehát az $u = v - \mathrm{id}$ függvényt, keressük az

$$\frac{1}{\beta} u(x) = \sup_{y \le x/\beta} u(y) = \sup_{y \le x} u\left(\frac{y}{\beta}\right)$$

függvényegyenlet megoldását. Világos, hogy $u = \infty$ konstans függvény megoldás. Mik a valós értékű megoldások? A jobb oldali $x \mapsto \sup_{y \le x/\beta} u(y)$ függvény csak monoton növekedő lehet, hiszen bővebb halmazon a szuprémum is bővebb. Más oldalról, ha u egy megoldás, akkor $u(x/\beta) \le \frac{1}{\beta} u(x)$, és minden ε-hoz van olyan $y \le x$, hogy $\frac{1}{\beta} u(x) - \varepsilon < u(y/\beta) \le u(x/\beta)$ az imént igazolt monoton növekedés szerint.

Összevetve tehát, ha u olyan megoldása a fenti függvényegyenletnek, amely a β diszkont faktortól független, akkor u pozitív homogén és monoton növekvő. Márpedig ilyen függvény csak az $u(x) = cx$ lehetséges, ahol $c \ge 0$. Ugyanis az $\frac{u(t)}{t}$ arány t-től független konstans, tehát a függvény minden pontban differenciálható, és deriváltja azonos a 0-beli deriválttal. E monoton függvény deriváltja így konstans, ezért cx alakú, valamely $c \ge 0$ mellett.

A Bellman-egyenlet β-tól független megoldásai tehát a $v = c \cdot \mathrm{id}$ alakú függvények, ahol $c \geq 1$ valamint a konstans ∞ függvény.

A fentiekben példát mutattunk olyan feladatra, amelyhez tartozó Bellman-egyenletet végtelen sok függvény elégít ki. Persze nem teljesül a $c \cdot \mathrm{id}$ függvényekre a 7.2.9. állításban az összes útra vonatkozó $\lim_{k\to\infty} v(x_k)\beta^k = 0$ feltevés, hiszen megválasztható olyan út, amelyre $\beta^k x_k \to -\infty$.

A következő modellben fogyasztónk nem kap kölcsönt így számára csak a megtakarítás marad lehetőségként. Továbbra is fenntartjuk azt a feltételezésünket, hogy a vagyon nem menthető át egyik periódusról a másikra, mivel minden periódusban elfogyasztjuk.

7.3.3. (Megtakarítás-modell)
Tekintsük a 7.3.2. „befektetés-megtakarítás"-modellt, de azzal a megszorítással, hogy nincs kölcsön, azaz $z_t \leq 0$.

Formalizáljuk a modellt, adjuk meg az (SP) feladat értékfüggvényét, az optimális terv(ek)et, írjuk fel a Bellman-egyenlet megoldásait.

Fel kell persze tennünk, hogy az x_0 kiinduló vagyon nem negatív, hiszen ellenkező esetben már a kezdő időszaki fogyasztás nem negatívitását sem lehetne garantálni. Ekkor $x_{t+1} = (-1/\beta)\, z_t$ szerint $x_t \geq 0$ minden t mellett, tehát az állapotváltozók X tere most $X = \mathbb{R}_+$ és a Γ átmenetfüggvény

$$\Gamma(x) = [0, x/\beta]$$

módon változik. Itt $x_{t+1} = 0$ jelenti az x_t vagyon t-edik periódusbeli teljes elfogyasztását, hiszen $c_t = x_t + z_t = x_t - \beta x_{t+1} = x_t$. Hasonlóan $x_{t+1} = \frac{x_t}{\beta}$ a teljes befektetést jelenti, hiszen ekkor $c_t = 0$, ezért $z_t = -x_t$. A (7.1) részletösszegre kapott formula szerint egy $x_0 \geq 0$ pontból induló $\bar{x} \in \Pi(x_0)$ pályára

$$u(\bar{x}) = x_0 - \lim_{n\to\infty} \beta^n x_n.$$

Az x_n vagyonok nem negatívitása miatt rögtön látszik, hogy $w(x_0) \leq x_0$. De az is látszik, hogy ha valamelyik lépésben $x_k = 0$-t választjuk, tehát ha elfogyasztjuk a teljes vagyonunkat, akkor utána már minden vagyon érték zérus lesz, így $u(\bar{x}) = x_0 - \lim_{n\to\infty} \beta^n x_n = x_0$. No de ez minden $x_0 \geq 0$ kiinduló pontra igaz, tehát az identitásfüggvény a feladat értékfüggvénye.

Optimális viselkedés tehát, ha az egyik periódusban – akár a kiindulási periódusban – a teljes vagyont elfogyasztjuk, és utána persze nem fogyasztunk semmit. Látható, hogy az identitásfüggvény mint a feladat értékfüggvénye a Bellman-egyenlet egyik megoldása is. Például az $x \mapsto 2x$ függvény $x \geq 0$ mellett szintén megoldása a Bellman-egyenletnek, tehát továbbra sem egyértelmű a függvény egyenlet megoldhatósága.

A molière-i eset, mikor $x_k = \frac{x_0}{\beta^k}$, tehát minden periódusban a teljes vagyonunkat megtakarítjuk, ezért egyetlen periódusban sem fogyasztunk. Kárpótlásként a vagyonunk sorozatára $x_k \to \infty$, a $\beta < 1$ feltevés miatt. Ekkor persze $\beta^k x_k = x_0$, így a (7.1) részletösszegekre

$$u(\bar{x}) = x_0 - x_0 = 0,$$

tehát a molière-i pálya haszna zéró, ezért az $x_0 > 0$ feltétel mellett az nem optimális terve az (X, F, Γ, β) dinamikus programozási feladatnak. Ez azért érdekes, mert erre a pályára az optimalitás szükséges feltétele

$$w(x_k) = x_k = x_k - \beta x_{k+1} + \beta x_{k+1} = F(x_k, x_{k+1}) + \beta w(x_{k+1})$$

teljesül. Ez azt mutatja, hogy a 7.3.1. állításban a fenti szükséges feltétel mellett az optimalitást garantáló

$$\limsup_{k \to \infty} \beta^k w(x_k) \leq 0,$$

extra feltétel az elegendőség igazolásához valóban nem hagyható el.

7.4. Bellman-egyenlet megoldása mint fixpont

A szakaszban a Bellman-egyenlet megoldását mint egy bizonyos operátor fixpontját fogjuk előállítani. Olyan tételeket fogunk kapni, amelyek a Bellman-egyenlet unicitását biztosítják, és amellett feltételeket kapunk a megoldás szigorúan monotonitására, szigorúan konvexitására, és az op-függvény egyértékűségére.

Az egész szakaszban feltesszük, hogy olyan (X, F, Γ, β) programozási feladattal állunk szemben, amelynek pillanatnyi hasznosság függvénye korlátos és $\beta < 1$. Tegyük fel még, hogy $X \subseteq \mathbb{R}^d$ konvex és $\Gamma(x) \subseteq X$ nem üres, konvex, kompakt. Ami nincs feltéve, hogy X kompakt is lenne. Ezt majd csak a legerősebb 7.5.1. tételben tesszük fel. A szakaszban nagyban támaszkodunk Berge maximumtételére. A halmazértékű leképezések folytonosságával kapcsolatos legfontosabb gondolatokat, és a Berge-tételt a teljesség kedvéért a Függelék 11. fejezetében tárgyaljuk. A továbbiakban feltesszük, hogy az olvasó a Berge-körrel tisztában van.

Jelölje $\mathcal{B}(X)$ az X-en értelmezett korlátos függvények terét, és $f \in \mathcal{B}(X)$ mellett $\|f\| = \sup_X |f|$. Tudjuk, hogy $(\mathcal{B}(X), \|\cdot\|)$ Banach-tér, és $f_n \to f$ ebben a normában pontosan azt jelenti, hogy $f_n \to f$ egyenletesen. A Banach-fixponttételt fogjuk alkalmazni, mely szerint egy *teljes metrikus téren értelmezett kontrakciónak egyetlen fixpontja van, és tetszőleges x pontból kiindulva a $T^n x$ iteráltak a fixponthoz konvergálnak.*

Kiinduló észrevétel az alábbi lemma:

7.4.1. lemma (Blackwell). *Legyen $T : \mathcal{B}(X) \to \mathcal{B}(X)$ operátor és $\beta \geq 0$ valós szám. Tegyük fel, hogy*

– *T monoton növő, azaz minden $f, g \in \mathcal{B}(X)$ $f \leq g$ esetén $Tf \leq Tg$, és*
– *minden $\alpha > 0$ mellett $T(f + \alpha) \leq Tf + \alpha\beta$.*

Ekkor T kontrakció a β konstanssal, azaz $\|Tf - Tg\| \leq \beta\|f - g\|$ minden $f, g \in \mathcal{B}(X)$ mellett.

Bizonyítás. Világos, hogy $f - g \leq \|f - g\|$ pontonként, ezért $f \leq g + \|f - g\|$ is fennáll pontonként, így a monotonitás és a második feltétel miatt $T(f) \leq$ $\leq T(g) + \beta\|f - g\|$. Ebből először a $Tf - Tg \leq \beta\|f - g\|$ egyenlőtlenséget, majd f és g szerepét felcserélve a $|Tf - Tg| \leq \beta\|f - g\|$ egyenlőtlenséget kapjuk pontonként. Ekkor persze $\|Tf - Tg\| \leq \beta\|f - g\|$ is fennáll, ami éppen azt jelenti, hogy T valóban egy β konstansú kontrakció. $\square$

Most bevezetjük azt az operátort, amelyre majd alkalmazni szeretnénk a Blackwell-lemmát. Ez az operátor az X halmazon értelmezett összes folytonos és korlátos függvények $C_b(X)$ Banach-terén van értelmezve. Mivel folytonos függvények egyenletes határértéke folytonos függvény, ezért $C_b(X) \subseteq \mathcal{B}(X)$ egy zárt altér, tehát $(C_b(X), \|\cdot\|)$ is Banach-tér.

7.4.2. állítás (Bellman-operátor)**.** *A rögzített (X, F, Γ, β) programozás mellett tegyük fel, hogy a Γ átmenetfüggvénye nem üres, kompakt, konvex értékű, folytonos halmazértékű leképezés; az $F : \operatorname{graph}(\Gamma) \to \mathbb{R}$ pillanatnyi megtérülés-függvény folytonos. Ekkor $f : X \to X$ folytonos, korlátos függvény mellett jelölje*

$$(Tf)(x) = \sup_{y \in \Gamma(x)} \{F(x, y) + \beta f(y)\} = \max_{y \in \Gamma(x)} \{F(x, y) + \beta f(y)\}.$$

Ekkor $Tf \in C_b(X)$. Ezt a $T : C_b(X) \to C_b(X)$ operátort nevezzük az (X, F, Γ, β) programozás Bellman-operátorának.

Definiáljuk tetszőleges $f \in C_b(X)$ mellett az

$$\begin{aligned} G_f(x) &= \operatorname{argmax}_{y \in \Gamma(x)} \{F(x, y) + \beta f(y)\} = \\ &= \{y \in \Gamma(x) : F(x, y) + \beta f(y) = Tf(x)\} \end{aligned}$$

halmazértékű leképezést. Ekkor G_f nem üres, kompakt értékű felülről félig folytonos függvény.

Bizonyítás. Mivel $\Gamma(x)$ kompakt, és az $y \mapsto F(x, y) + \beta f(y)$ függvény folytonos, ezért sup helyett valóban max írható. A 11.2.3. Berge-tétel szerint Tf folytonos, de F, illetve f korlátossága miatt korlátos is. Így a fenti T operátorra $T : C_b(X) \to C_b(X)$ valóban fennáll.

A G_f függvényre vonatkozó állítás is a Berge-tétel következménye. $\square$

Amikor a Bellman-operátorról beszélünk, akkor a fenti definíció feltételeit mindig elfogadjuk. Gyűjtsük össze a Bellman-operátor legfontosabb tulajdonságait.

7.4.3. állítás (Bellman-operátor tulajdonságai)**.** *Legyen T az (X, F, Γ, β) programozás Bellman-operátora. Ekkor*
– *T monoton növő;*
– *minden $\alpha > 0$ konstans mellett $T(f+\alpha) = Tf + \alpha\beta$;*
– *tegyük fel, hogy az átmenetfüggvény monoton növő, azaz $\Gamma(x_1) \subseteq \Gamma(x_2)$, ha $x_1 \leq x_2$.*

Ha minden rögzített $y \in X$ mellett az $F(\cdot, y) : X \to \mathbb{R}$ függvény szigorúan monoton növő, akkor Tf is az.
– *Tegyük fel, hogy az átmenetfüggvény konkáv, azaz $\alpha\Gamma(x_1) + (1-\alpha)\Gamma(x_2) \subseteq$ $\subseteq \Gamma(\alpha x_1 + (1-\alpha)x_2)$ fennáll $\alpha \in (0{,}1)$ és $x_1, x_2 \in X$ mellett.*

Ha az F hozamfüggvény szigorúan konkáv és f konkáv, akkor Tf is szigorúan konkáv.

Bizonyítás. Az első két állítás a Bellman-operátor definíciója szerint nyilvánvaló.

Most tegyük fel, hogy a harmadik állítás feltételei teljesülnek és $x_1 \leqq x_2$, de $x_1 \neq x_2$. Legyen $y_1 \in \operatorname{argmax}_{y\in\Gamma(x_1)} \{F(x_1, y) + \beta f(y)\}$. Ekkor $y_1 \in$ $\in \Gamma(x_1) \subseteq \Gamma(x_2)$ szerint

$$\begin{aligned}(Tf)(x_1) = F(x_1, y_1) + \beta f(y_1) < F(x_2, y_1) + \beta f(y_1) \leq \\ \leq \max_{y\in\Gamma(x_2)} \{F(x_2, y) + \beta f(y)\} = Tf(x_2).\end{aligned}$$

Nézzük a negyedik állítást. Legyen most is

$$y_i \in \operatorname{argmax}_{y\in\Gamma(x_i)} \{F(x_i, y) + \beta f(y)\}.$$

Ekkor az átmenetfüggvényre tett konkavitási feltevés szerint $\alpha y_1 + (1-\alpha)y_2 \in$ $\in \alpha\Gamma(x_1) + (1-\alpha)\Gamma(x_2) \subseteq \Gamma(\alpha x_1 + (1-\alpha)x_2)$. Ezért

$$\begin{aligned}&Tf(\alpha x_1 + (1-\alpha)x_2) \geq \\ &\geq F(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2) + \beta f(\alpha y_1 + (1-\alpha)y_2) > \\ &> \alpha F(x_1, y_1) + (1-\alpha)F(x_2, y_2) + \beta(\alpha f(y_1) + (1-\alpha)f(y_2)) = \\ &= \alpha(F(x_1, y_1) + \beta f(y_1)) + (1-\alpha)(F(x_2, y_2) + \beta f(y_2)) = \\ &= \alpha Tf(x_1) + (1-\alpha)Tf(x_2).\end{aligned}$$

Ezt kellett belátni. □

7.4.4. tétel (Bellman-egyenlet egyetlen megoldásának folytonossága)**.** *Tekintsük az (X, F, Γ, β) programozást, ahol $X \subseteq \mathbb{R}^d$ konvex halmaz; F folytonos függvény; Γ nem üres, konvex, kompakt értékű folytonos függvény; és $\beta < 1$. Tudjuk, hogy ekkor a feladat Bellman-egyenletének egyetlen megoldása van, hiszen az értékfüggvény is korlátos.*

Ekkor tetszőleges $v_0 \in C_b(X)$ *függvényből kiindulva az* $v_n = T^n v_0$ *iteráltak is korlátos folytonos függvények; egyenletesen konvergálnak a feladat Bellman-egyenletének egyetlen megoldásához, amely korlátos és folytonos.*

A feladat optimal policy leképezése ekkor egy nem üres, kompakt értékű, felülről félig folytonos leképezés.

Bizonyítás. A tett feltevések szerint alkalmazható a 7.4.3. állítás első két pontja, ami éppen a 7.4.1. Blackwell-lemma feltételeinek teljesülését adja, tehát $T : C_b(X) \to C_b(X)$ egy kontrakció. A Banach-fixponttétel szerint létezik $Tv = v \in C_b(X)$ fixpont, ami nyilván a Bellman-egyenlet egyetlen megoldása. A $v_n = T^n f$ iteráltakra $v_n \to v$ egyenletesen, tetszőleges választott kiindulási v_0 mellett.

Ezzel a v megoldásfüggvény felírt G_v leképezésről láttuk, hogy nem üres, kompakt értékű, felülről félig folytonos függvény. No de v a Bellman-egyenlet korlátos megoldása, ezért G_v megegyezik a programozás op-függvényével, tehát az op-függvény nem üres, kompakt értékű, felülről félig folytonos függvény. Ezt kellett belátni. □

7.4.5. tétel (a Bellman-egyenlet egyetlen megoldásának szigorúan monoton növekedése)**.** *Most tegyük fel a Bellman-egyenlet folytonosságát garantáló 7.4.4. tétel feltételei mellé, hogy az átmenetfüggvény monoton növő, és a hozadékfüggvény szigorúan monoton növő az első változójában. Ekkor a Bellman-egyenlet egyetlen megoldása a fentiek mellett még szigorúan monoton növő is.*

Bizonyítás. Jelölje $C'(X)$ az X-en értelmezett folytonos, korlátos, monoton növő függvények metrikus alterét. Látható, hogy $C'(X)$ egy zárt részhalmaza a $C_b(X)$ teljes metrikus térnek, ezért maga is teljes metrikus tér. Az F-re és Γ-ra tett kiegészítő feltételek szerint $T : C'(X) \to C'(X)$ operátor, hiszen láttuk 7.4.3. állításban, hogy T monoton függvényt monoton függvénybe képez. Itt is alkalmazhatjuk tehát a Banach-fixponttételt. Az emiatt létező v fixpont, nyilván a Bellman-egyenlet egyetlen megoldása, és $Tv = v$ miatt szigorúan monoton növekedő is. □

Most nézzük a szigorúan konkáv megoldás létezését garantáló állítást:

7.4.6. tétel (A Bellman-egyenlet egyetlen megoldásának szigorúan konkavitása)**.** *Most tegyük fel a Bellman-egyenlet megoldásának folytonosságát garantáló 7.4.4. tétel feltételei, mellé hogy az átmenetfüggvény konkáv, és az hozadékfüggvény együttesen szigorúan konkáv. Ekkor a Bellman-egyenlet egyetlen megoldása a fentiek mellett még szigorúan konkáv is. Sőt ebben az esetben az optimális policy leképezés minden értéke egyelemű halmaz, és a függvény folytonos.*

Bizonyítás. Jelölje most $C'(X)$ az X-en értelmezett folytonos, korlátos, konkáv függvények metrikus alterét. Látható, hogy $C'(X)$ egy zárt részhalmaza a $C(X)$ Banach-térnek, ezért maga is teljes metrikus tér. Az F-re és Γ-ra tett kiegészítő feltételek szerint $T : C'(X) \to C'(X)$ operátor, hiszen konkáv f mellett Tf is konkáv (7.4.3). Itt is alkalmazhatjuk tehát a Banach-fixponttételt. Az emiatt létező v fixpont, nyilván a Bellman-egyenlet egyetlen megoldása, és $Tv = v$ miatt szigorúan konkáv is.

Mivel egy szigorúan konkáv és egy konkáv függvény összege szigorúan konkáv, ezért az optimal policy függvény egy szigorúan konkáv függvény maximum helyeit tartalmazza. Persze szigorúan konkáv függvény két különböző pontban nem tudja felvenni a maximális értékét, így a képek egyelemű halmazok. Ekkor persze a felülről félig folytonosság folytonosságot jelent. □

7.5. A megoldás függvény és az op-leképezés közelítése

Az előző szakaszban nem csak egy jól használható feltételt kaptunk a Bellman-egyenlet megoldhatóságára, tehát a szuprémum feladat megoldhatóságára, hanem a Banach-fixponttételt használva a megoldásfüggvény iteratív közelítésére egy akár numerikusan is jól használható eljáráshoz jutottunk. A döntéshozó számára viszont elsősorban nem az optimális pálya az érdekes, hanem az, hogy hogyan lehet optimális pályára jutni, vagy ott maradni. A dinamikus programozási feladat nyelvére lefordítva ez azt jelenti, hogy az optimal policy leképezés az amire igazán kíváncsiak vagyunk. Adott pozícióból, milyen pozíciót kell választani ahhoz, hogy választásunkkal az optimális pályára jussunk.

Természetes tehát az a kérdés, hogy ismerve az optimális úthoz konvergáló iterációt vajon tudjuk-e az iteráló függvények segítségével az optimal policy leképezést is közelíteni?

Az igenlő választ a következő tétel jelenti.

7.5.1. tétel (Az op-leképezés közelítése szigorúan konkáv esetben)**.** *Tegyük fel, hogy teljesülnek a Bellman-egyenlet egyetlen megoldásának szigorúan konkavitására vonatkozó 7.4.6. tétel feltételei. Láttuk, hogy tetszőleges $v_0 \in C_b(X)$ korlátos, konkáv, folytonos kiindulási függvény esetén a $v_{n+1} = Tv_n$ iteráció korlátos, szigorúan konkáv függvények olyan sorozata, amelyre $v_n \to v$ egyenletesen, ahol v a Bellman-egyenlet egyetlen és szigorúan konkáv, korlátos és folytonos megoldása. Jelölje g a megoldáshoz tartozó op-függvényt, azaz*

$$g(x) = \operatorname{argmax}\{F(x,y) + \beta v(y) : y \in \Gamma(x)\}$$

és hasonlóan az n-edik közelítő függvényre:

$$g_n(x) = \operatorname{argmax}\{F(x,y) + \beta v_n(y) : y \in \Gamma(x)\}.$$

Ekkor $g_n \to g$ pontonként. Abban az esetben, ha X kompakt a $g_n \to g$ konvergencia még egyenletes is.

A szakasz célja, hogy a fenti tétel indoklását megértsük. A tétel sokkal inkább tartozik a függelék 11. fejezetében tárgyalt Berge-tételhez, mint a dinamikus programozáshoz. Két lemmára van először is szükségünk.

7.5.2. lemma. *Legyen (Y, d) kompakt metrikus tér és $f : Y \to \mathbb{R}$ olyan folytonos függvény, amely csak egyetlen pontban veszi fel a globális maximumát, azaz minden $y \in K, y \neq y^*$ mellett $f(y) < f(y^*)$. Ekkor minden $\varepsilon > 0$-hoz létezik $\delta > 0$, hogy*

$$f(y^*) - f(y) < \delta \quad \textit{esetén} \quad d(y, y^*) < \varepsilon.$$

Bizonyítás. Adott $\varepsilon > 0$ mellett tekintsük az alábbi halmazt.

$$A = \{y \in Y : d(y, y^*) \geq \varepsilon\}.$$

Ha $A = \emptyset$, akkor akármilyen $\delta > 0$ jó lesz, hiszen minden $y \in Y$-ra $d(y, y^*) < \varepsilon$.
Ha $A \neq \emptyset$, akkor legyen

$$\delta = f(y^*) - \sup_{y \in A} f(y) = \inf_{y \in A} \{f(y^*) - f(y)\}.$$

Az A kompaktsága miatt létezik $u \in A$, amelyre $\delta = f(y^*) - f(u)$, amiből következik, hogy $\delta > 0$, hiszen f nem veszi fel globális maximumát A-n. Ez azt jelenti, hogy találtunk olyan $\delta > 0$ számot, amelyre minden $y \in A$ esetén $f(y^*) - f(y) \geq \delta$, vagy ami ugyanaz $f(y^*) - f(y) < \delta$ esetén $y \notin A$ azaz $d(y, y^*) < \varepsilon$. Ezt kellett belátni. □

7.5.3. lemma. *Legyenek X, Y metrikus terek, $f : X \times Y \to \mathbb{R}$ a második változójában folytonos függvény, legyen $\Gamma : X \multimap Y$ egy nem üres, kompakt értékű leképezés. Tegyük fel, hogy minden $x \in X$ mellett a*

$$g(x) = \operatorname{argmax}\{f(x, y) : y \in \Gamma(x)\}$$

halmaz egyelemű, azaz a fenti $g : X \to \mathbb{R}$ leképezés egyértékű függvény.

1. *Ekkor minden $\varepsilon > 0$-hoz és minden $x \in X$-hez létezik $\delta_x > 0$, hogy*

   $$f(x, g(x)) - f(x, y) < \delta_x \implies d(y, g(x)) < \varepsilon$$

   fennáll minden $y \in \Gamma(x)$ esetén.

2. *Ha még azt is feltesszük, hogy X kompakt; Γ folytonos; f folytonos, akkor a fenti δ univerzálisan is megválasztható, azaz minden $\varepsilon > 0$-hoz létezik $\delta > 0$, hogy*

   $$f(x, g(x)) - f(x, y) < \delta \implies d(y, g(x)) < \varepsilon$$

   fennáll minden $(x, y) \in \operatorname{graph}(\Gamma)$ esetén.

Bizonyítás. Rögzített x mellett alkalmazzuk az előző lemmát a $f(x,\cdot):\Gamma(x)\to\mathbb{R}$ függvényre, így akármilyen δ, vagy $A(x)\neq\varnothing$ esetben

$$\begin{aligned}\delta(x) &= f\left(x,g(x)\right)-\max\left\{f\left(x,y\right):y\in A(x)\right\}=\\ &=\min\left\{f\left(x,g(x)\right)-f\left(x,y\right):y\in A(x)\right\}\end{aligned}$$

jelöléssel az első állítás indoklásával készen is vagyunk, ahol

$$A(x)=\left\{y\in\Gamma(x):d\left(y,g(x)\right)\geq\varepsilon\right\}=\Gamma(x)\cap\left(B^0(g(x),\varepsilon)\right)^c.$$

A második állításhoz először is vegyük észre, hogy f és Γ folytonossága és A 11.2.3. Berge-tétel szerint g is folytonos függvény. Másodszor a Γ felülről félig folytonossága és g folytonossága miatt a

$$\left\{x\in X:\Gamma(x)\subseteq B^0\left(g(x),\varepsilon\right)\right\}$$

halmaz nyílt. Legyen ugyanis x_0 egy olyan pont, amelyre $\Gamma(x_0)\subseteq B^\circ(g(x_0),\varepsilon)$. Jelölje ρ a $\Gamma\left(x_0\right)$ kompakt halmaz és a $B^\circ\left(g\left(x_0\right),\varepsilon\right)^c$ zárt halmaz távolságát. Világos, hogy egy kompakt és egy attól diszjunkt zárt halmaz távolsága pozitív, tehát $\rho>0$. Kihasználva Γ felülről félig folytonosságát és g folytonosságát, választható $N\in\tau\left(x_0\right)$ környezet, hogy minden $x\in N$ mellett

$$\Gamma\left(x\right)\subseteq B^\circ\left(g\left(x_0\right),\varepsilon-\rho/2\right)\subseteq B^\circ\left(g\left(x\right),\varepsilon\right),$$

tehát a halmaz valóban nyílt, ami azt jelenti, hogy az

$$X^{'}=\left\{x\in X:A(x)\neq\varnothing\right\}$$

komplementer halmaz zárt. Lévén X kompakt, az $X^{'}$ is az.

Ha $X^{'}=\varnothing$, akkor minden $x\in X$-re $A(x)=\varnothing$, azaz akármilyen $\delta>0$ szám univerzálisan jó lesz.

Ha $X^{'}\neq\varnothing$, akkor $A:X^{'}\multimap Y$ nem üres, kompakt értékű, felülről félig folytonos halmazértékű függvény. A 11.1.3. állítás ismeretében ehhez csak az

$$x\mapsto B^\circ\left(g\left(x\right),\varepsilon\right)^c$$

zárt leképezés voltát kell látnunk. Legyen tehát $(x_n,y_n)\to(x,y)$ olyan konvergens sorozat, melynek tagjai e leképezés gráfjába esnek, azaz $d(g(x_n),y_n)\geq\geq\varepsilon$. A g függvény folytonossága és a metrikus tér távolság függvényének folytonossága szerint

$$d\left(g\left(x\right),y\right)=\lim_n d\left(g\left(x_n\right),y_n\right)\geq\varepsilon$$

is fennáll, ami éppen e halmazértékű függvény zárt leképezés voltát igazolja.

A Berge-tétel felülről félig folytonos 11.2.2. alakja értelmében a $\delta : X^{'} \to \mathbb{R}$ értékfüggvény egy alulról félig folytonos, és persze pozitív függvény. Egy alulról félig folytonos függvény felveszi legkisebb értékét egy kompakt halmazon, tehát létezik $\rho \in \mathbb{R}_{+}$ szám, amelyre $\rho = \min\left\{\delta(x) : x \in X^{'}\right\} > 0$. Ekkor

$$f(x, g(x)) - f(x, y) < \rho \leq \delta(x) \implies d(y, g(x)) < \varepsilon.$$

már univerzálisan teljesül minden $x \in X$ mellett. $\square$

A következő állítás speciális esetként tartalmazza a szakasz fő tételét. Azt mutatjuk meg, hogy ha egy jól viselkedő feltételes szélsőérték feladat célfüggvényét egyenletesen közelítjük, akkor a feladat maximalizáló függvényét is egyenletesen közelítik a közelítő célfüggvényekhez tartozó maximalizáló függvények.

7.5.4. állítás. *Legyenek X, Y metrikus terek; $\Gamma : X \multimap Y$ nem üres, kompakt értékű halmazértékű leképezés; $f, f_n : \operatorname{graph}\Gamma \to \mathbb{R}$ a második változójukban folytonos függvények, amelyekre $f_n \to f$ egyenletesen, és a*

$$g_n(x) = \operatorname{argmax}\{f_n(x, y) : y \in \Gamma(x)\} \text{ és } g(x) = \operatorname{argmax}\{f(x, y) : y \in \Gamma(x)\}$$

függvények egyértékűek minden $x \in X$ mellett. Ekkor

1. *minden $x \in X$ mellett $g_n(x) \to g(x)$ pontonként; továbbá*
2. *ha feltesszük, hogy X kompakt; Γ halmazértékű leképezés folytonos; f_n, f folytonosak, akkor a $g_n \to g$ konvergencia még egyenletes is.*

Bizonyítás. Az első észrevétel, hogy mivel $g(x)$ az $f(x, \cdot)$ függvényt, és $g_n(x)$ az $f_n(x, \cdot)$ függvényt maximalizálja a $\Gamma(x)$ halmaz felett, ezért

$$\begin{aligned}
&f(x, g(x)) - f(x, g_n(x)) = \\
&= f(x, g(x)) - f_n(x, g(x)) + f_n(x, g(x)) - f(x, g_n(x)) \leq \\
&= f(x, g(x)) - f_n(x, g(x)) + f_n(x, g_n(x)) - f(x, g_n(x)) \leq 2\|f - f_n\|_\infty.
\end{aligned}$$

Legyen $\varepsilon > 0$ rögzítve. A gyengébb feltételrendszer esetén ehhez az ε-hoz és egy rögzített $x \in X$-hez létezik $\delta_x > 0$ szám a fenti 7.5.3. lemma szerint, tehát minden $x \in X$ és $y \in \Gamma(x)$ mellett

$$f(x, g(x)) - f(x, y) < \delta_x \implies d(y, g(x)) < \varepsilon.$$

Válasszunk $N_x \in \mathbb{N}$ indexet úgy, hogy $2\|f - f_n\|_\infty < \delta_x$ minden $n \geq N_x$-re teljesüljön. Ekkor a fentiek szerint $f(x, g(x)) - f(x, g_n(x)) < \delta_x$, ezért a lemma miatt $d(g(x), g_n(x)) < \varepsilon$. Ez éppen azt jelenti, hogy $g_n(x) \to g(x)$.

Az erősebb feltételrendszer esetén a 7.5.3. lemmát is annak erősebb formájában alkalmazhatjuk, ezért az $\varepsilon > 0$ számhoz választhatunk olyan univerzális $\delta > 0$ számot, hogy

$$f(x, g(x)) - f(x, y) < \delta \implies d(y, g(x)) < \varepsilon$$

fennáll minden $y \in \Gamma(x)$ esetén. Ehhez a δ-hoz legyen N olyan, hogy minden $n \geq N$ mellett $2\|f - f_n\|_\infty < \delta$, így $\sup_{x \in X} d\left(g(x), g_n(x)\right) \leq \varepsilon$. Ez azt jelenti, hogy $g_n \to g$ egyenletesen. □

Talán nyilvánvaló, de a rend kedvéért összefoglaljuk a szakasz fő tételének igazolását. Természetesen a Berge-tétel fenti 7.5.4. kiegészítését használjuk.

A 7.5.1. tétel bizonyítása. Mivel v_n, v szigorúan konkáv függvények, ezért legfeljebb egy pontban vehetik fel egy konvex halmaz felett a maximális értéküket. A Banach-iteráció szerint $v_n \to v$ egyenletesen, ami bevezetve az

$$\begin{aligned} f(x, y) &= F(x, y) + \beta v(y), \\ f_n(x, y) &= F(x, y) + \beta v_n(y) \end{aligned}$$

jelöléseket az jelenti, hogy $f_n \to f$ egyenletesen. Világos, hogy

$$g_n(x) = \operatorname{argmax}\{f_n(x, y) \colon y \in \Gamma(x)\} \text{ és } g(x) = \operatorname{argmax}\{f(x, y) \colon y \in \Gamma(x)\}.$$

Teljesülnek tehát a Berge-tétel argmax leképezésének közelítésére igazolt 7.5.4. állítás feltételei. Ezt kellett belátni. □

7.6. Differenciálhatósági feltételek

A szakaszban három fontos tételt látunk be. Először a értékfüggvény deriválhatósági tételét gondoljuk meg, majd rátérünk az Euler-egyenlet szükséges és elégséges voltát garantáló tételek igazolására.

7.6.1. Értékfüggvény differenciálhatósága

Előkészítésként három lemmára van szükségünk.

7.6.1. lemma. *Egy konkáv halmazértékű leképezés grafikonja konvex halmaz.*

Bizonyítás. Legyen $(x_1, y_1), (x_2, y_2) \in \operatorname{graph} \Gamma$ és $\alpha \in (0,1)$. Ekkor

$$\alpha y_1 + (1 - \alpha) y_2 \in \alpha \Gamma(x_1) + (1 - \alpha) \Gamma(x_2) \subseteq \Gamma(\alpha x_1 + (1 - \alpha) x_2),$$

ami éppen azt jelenti, hogy $(\alpha x_1 + (1 - \alpha) x_2, \alpha y_1 + (1 - \alpha) y_2) \in \operatorname{graph} \Gamma$. Ezt kellett belátni. □

Emlékezzünk arra, hogy egy $f : X \to \mathbb{R}$ konvex függvény esetén a $d \in \mathbb{R}^n$ vektor az f függvény x_0 pontbeli $\partial f(x_0)$ *szubderivált* halmazának egy eleme, ha minden $x \in X$ mellett

$$f(x) \geq f(x_0) + \langle x - x_0; d\rangle.$$

Tudjuk, hogy egy konvex függvény pontosan akkor differenciálható, ha a szubderivált halmaza egyetlen pontból áll, és tudjuk, hogy egy konvex függvény szubderivált halmaza soha nem üres.

7.6.2. lemma (konvex függvény differenciálhatósága)**.** *Legyenek $f, g : X \to \mathbb{R}$ konvex függvények, amelyekre*
– *$f(x_0) = g(x_0)$, továbbá*
– *$f(x) \geq g(x)$ minden $x \in U$ mellett, ahol $U \in \tau(x_0)$ az x_0 pont egy környezete.*

Ekkor az f függvény x_0 pontbeli differenciálhatóságából következik a g függvény x_0 pontbeli differenciálhatósága, és $g'(x_0) = f'(x_0)$.

Bizonyítás. Ha $d \in \partial g(x_0)$, akkor

$$f(x) - f(x_0) \geq g(x) - g(x_0) \geq \langle d; x - x_0\rangle,$$

ami azt jelenti, hogy $\partial g(x_0) \subseteq \partial f(x_0)$. De $\partial g(x_0) \neq \varnothing$, ami azt jelenti, hogy $\partial g(x_0)$ halmaz egyelemű, és megegyezik az $\partial f(x_0)$ halmaz egyetlen elemével. □

A következő lemma az differenciálhatóságra vonatkozó állítások fontos alkotóeleme. A Hahn–Banach szeparációs tételt a következő formában használjuk:

Ha adott egy $K \subseteq \mathbb{R}^n$ konvex halmaz, és $z \notin \operatorname{int} K$ pont, akkor létezik olyan $p \in \mathbb{R}^n, p \neq 0$ vektor, amelyre minden $x \in K$ mellett $\langle p; x\rangle \leq \langle p; z\rangle$. Látható tehát, hogy a lineáris függvények közül a konstans zésus függvény az egyetlen, amely egy nyílt halmazon felveszi a legnagyobb értékét.

7.6.3. lemma. *Legyen $\Gamma : X \multimap X$ egy konkáv halmazértékű leképezés. Tegyük fel, hogy $x_0 \in \operatorname{int} X$ és $y_0 \in \operatorname{int} \Gamma(x_0)$. Ekkor $(x_0, y_0) \in \operatorname{int}(\operatorname{graph}(\Gamma))$.*

Bizonyítás. Tegyük fel – indirekt –, hogy $(x_0, y_0) \notin \operatorname{graph} \Gamma$. Láttuk, hogy konkáv leképezés grafikonja konvex, így $\operatorname{graph} \Gamma$ konvex – 7.6.1. lemma. Alkalmazhatjuk tehát a Hahn–Banach-tételt. Kapunk $(p_1, p_2) \neq 0$ vektort, amelyre teljesül, hogy minden $(x, y) \in \operatorname{graph} \Gamma$ esetén

$$\langle p_1; x\rangle + \langle p_2; y\rangle \leq \langle p_1; x_0\rangle + \langle p_2; y_0\rangle.$$

Először is megmutatjuk, hogy $p_2 = 0$. Az $y_0 \in \operatorname{int} \Gamma(x_0)$ szerint van olyan $V \in \tau(y_0)$ nyílt környezete az y_0 pontnak, amelyre $\{x_0\} \times V \subseteq \operatorname{graph} \Gamma$, ezért a fenti egyenlőtlenség alkalmazható minden olyan (x_0, y) párra, ahol $y \in V$. Így

$$\langle p_1; x_0\rangle + \langle p_2; y\rangle \leq \langle p_1; x_0\rangle + \langle p_2; y_0\rangle,$$

ami azt jelenti, hogy az $\langle p_2; \cdot\rangle$ lineáris funkcionál felveszi a legnagyobb értékét a V nyílt halmazon. Ez persze csak úgy lehetséges, ha $p_2 = 0$ valóban teljesül. Hasonlóan, $p_1 = 0$ is adódik, hiszen $x_0 \in \operatorname{int} X$ szerint létezik $U \in \tau(x_0)$ nyílt halmaz, amelyre minden $x \in U$ hoz létezik $y \in X$, hogy $(x, y) \in \operatorname{graph} \Gamma$. Ezért minden $x \in U$ mellett

$$\langle p_1; x\rangle \leq \langle p_1; x_0\rangle,$$

ami persze csak $p_1 = 0$ esetben lenne lehetséges, ami ellentmond a Hahn–Banach-tétel konklúziójának. □

7.6.4. tétel (az értékfüggvény deriválhatósága)**.** *Tegyük fel: az (X, F, Γ, β) programozás kielégíti a a Bellman-egyenlet egyetlen megoldásának konkavitására vonatkozó 7.4.6. tétel feltételeit. Tudjuk, hogy ekkor az op-leképezés függvény. Jelölje ezt g. Ha $x_0 \in \operatorname{int} X$ és $g(x_0) \in \operatorname{int} \Gamma(x_0)$, akkor az F függvény x_0-beli első változóbeli differenciálhatóságából következik a Bellman-egyenlet v megoldásának differenciálhatósága az x_0 pontban és*

$$v'(x_0) = D_1F(x_0, g(x_0))$$

Bizonyítás. A tett feltevések és a fenti 7.6.3. lemma szerint $(x_0, g(x_0)) \in \in \operatorname{int} \operatorname{graph}(\Gamma)$, ami azt jelenti, hogy létezik $U \in \tau(x_0)$ és $V \in \tau(g(x_0))$ nyílt halmaz, amelyekre $U \times V \subseteq \operatorname{graph} \Gamma$. Ekkor persze $U \times \{g(x_0)\} \subseteq \operatorname{graph} \Gamma$, azaz minden $x \in U$ mellett $g(x_0) \in \Gamma(x)$, ezért minden rögzített $x \in U$ mellett fennáll a következő egyenlőtlenség:

$$v(x) = \sup \{F(x, y) + \beta v(y) : y \in \Gamma(x)\} \geq F(x, g(x_0)) + \beta v(g(x_0)).$$

Tekintsük most az $f : U \to \mathbb{R}$

$$f(x) = F(x, g(x_0)) + \beta v(g(x_0))$$

függvényt. Az F konkavitása és differenciálhatósága miatt f konkáv és differenciálható; a g op-leképezés definíciója szerint $f(x_0) = v(x_0)$; az eddig igazoltak szerint $v \geq f$ az U nyílt halmazon. A Bellman-egyenlet megoldása (szigorúan) konkáv, így alkalmazható a konvex függvény differenciálhatóságra vonatkozó 7.6.2. lemma a $-v \leq -f$ szereposztással. Azt kapjuk tehát hogy v is differenciálható x_0-ban és $v'(x_0) = f'(x_0) = D_1F(x_0, g(x_0))$. Ezt kellett belátni. □

7.6.2. Az Euler-egyenlet és a transzverzalitási feltétel

7.6.5. definíció (Belső terv)**.** Az $\bar{x} \in \Pi$ tervet a Γ átmenetfüggvényhez tartozó *belső tervnek* vagy *belső pályának* mondjuk, ha az $(x_t, x_{t+1}) \in \operatorname{int}(\operatorname{graph}\Gamma)$ tartalmazás fennáll minden $t \in \mathbb{N}$ mellett.

7.6.6.
A 7.6.3. lemma voltaképpen úgy fogalmazható, hogy konkáv átmenetfüggvény; $x_0 \in \operatorname{int} X$; $x_{t+1} \in \operatorname{int}\Gamma(x_t)$ feltételek mellett az $\bar{x}$ pálya egyben belső pálya is.

7.6.7.
Ha a $\bar{x} \in \Pi$ megengedett út az átmenetfüggvény belső terve, akkor $(x_t, x_{t+1}) \in$ $\in \operatorname{int}\operatorname{graph}\Gamma$, ezért létezik $V \in \tau(x_{t+1})$ nyílt halmaz, amelyre $\{x_t\} \times V \subseteq$ $\subseteq \operatorname{graph}\Gamma$. Hasonlóan, az $(x_{t+1}, x_{t+2}) \in \operatorname{int}\operatorname{graph}\Gamma$ miatt létezik $U \in \tau(x_{t+1})$ nyílt környezet, amelyre meg az $U \times \{x_{t+2}\} \subseteq \operatorname{graph}\Gamma$ tartalmazás is teljesül. Így

$$x_{t+1} \in V \cap U \subseteq \{y \in \Gamma(x_t) : x_{t+2} \in \Gamma(y)\},$$

ami azt jelenti, hogy x_{t+1} az $\{y \in \Gamma(x_t) : x_{t+2} \in \Gamma(y)\}$ halmaz belső pontja.

7.6.8. állítás (Euler-egyenlet szükségessége)**.** *Tegyük fel, hogy $\bar{x} \in \Pi(x_0)$ terv optimális belső megoldása az (X, F, Γ, β) programozásnak, ahol az F hozamfüggvény differenciálható és $\beta > 0$.*

Ha $w(x_0)$ véges, akkor minden $t \in \mathbb{N}$ mellett fennáll az úgynevezett Euler-egyenlet*:*

$$D_2F(x_t, x_{t+1}) + \beta D_1F(x_{t+1}, x_{t+2}) = 0.$$

Bizonyítás. Rögzített t mellett

$$w(x_0) = \sum_{k=0}^{t-1} \beta^k F(x_k, x_{k+1}) + \beta^t \left(F(x_t, x_{t+1}) + \beta F(x_{t+1}, x_{t+2})\right) + \\ + \sum_{k=t+2}^{\infty} \beta^k F(x_k, x_{k+1}).$$

Mivel $\bar{x}$ optimális pálya; $w(x_0)$ véges; és $\beta \neq 0$, az

$$h(y) = F(x_t, y) + \beta F(y, x_{t+2})$$

függvény éppen x_{t+1}-ben veszi fel véges maximumát az

$$\{y \in \Gamma(x_t) : x_{t+2} \in \Gamma(y)\}$$

halmaz felett. De feltettük az $\bar{x}$ optimális pályáról, hogy egyben belső pálya is, ezért az iménti 7.6.7. megjegyzés szerint az x_{t+1} állapot belső pontja ennek a halmaznak, így F differenciálhatósága szerint a fenti h hozzárendelés is differenciálható, és $h'(x_{t+1}) = D_2F(x_t, x_{t+1}) + \beta D_1F(x_{t+1}, x_{t+2}) = 0$. Ezt kellett belátni. $\square$

7.6.9. definíció (transzverzalitási feltétel)**.** Azt mondjuk, hogy az (X, F, Γ, β) programozás $\bar{x} \in \Pi$ útjára teljesül a *transzverzalitási feltétel*, ha

$$\lim_{t\to\infty} \beta^t \langle D_1F\left(x_t, x_{t+1}\right); x_t \rangle = 0$$

Most nézzük az optimum létezésének elegendő feltételét.

7.6.10. tétel (Euler-egyenlet és transzverzalitási feltétel elégségessége)**.** *Tegyük fel, hogy az (X, F, Γ, β) programozás X állapot tere nem tartalmaz negatív koordinátájú vektort; F hozadék függvénye konkáv; differenciálható és az első változójában monoton növő; valamint a Γ átmenetfüggvénye konkáv. Ekkor ha $\bar{x}^* \in \Pi(z)$ olyan út, amelyre fennáll a transzverzalitási feltétel és az Euler-egyenlet, akkor az $\bar{x}^* \in \Pi(z)$ a z állapotból induló optimális megoldás is.*

Bizonyítás. Legyen $\bar{x} \in \Pi(z)$ egy másik út. Megmutatjuk, hogy

$$\sum_{k=0}^{\infty} \beta^k \left(F(x_k^*, x_{k+1}^*) - F(x_k, x_{k+1})\right) \geq 0.$$

Egy f differenciálható, konkáv függvény gráfja minden pontja az érintősík alatt fekszik, tehát $f(x) \leq f(x^*) + \langle f'(x^*); x - x^* \rangle$ ami azzal ekvivalens, hogy $f(x^*) - f(x) \geq \langle f'(x^*); x^* - x \rangle$. Ezt az egyenlőtlenséget használjuk a F függvényre.

Az F konkavitása és differenciálhatósága szerint tehát

$$\begin{aligned} F(x_k^*, x_{k+1}^*) - F(x_k, x_{k+1}) \geq \langle F'(x_k^*, x_{k+1}^*); \left(x_k^* - x_k, x_{k+1}^* - x_{k+1}\right) \rangle = \\ = \langle D_1F(x_k^*, x_{k+1}^*); x_k^* - x_k \rangle + \langle D_2F(x_k^*, x_{k+1}^*); x_{k+1}^* - x_{k+1} \rangle. \end{aligned}$$

Most írjuk fel a keresett összeg T-edik részletösszegét:

$$\begin{aligned} &\sum_{k=0}^{T} \beta^k \left(F(x_k^*, x_{k+1}^*) - F(x_k, x_{k+1})\right) \geq \\ &\quad \geq \sum_{k=0}^{T} \beta^t \left(\langle D_1F(x_k^*, x_{k+1}^*); x_k^* - x_k \rangle + \langle D_2F(x_k^*, x_{k+1}^*); x_{k+1}^* - x_{k+1} \rangle\right). \end{aligned}$$

A szummát kibontva azt kapjuk, hogy az nem más, mint

$$\begin{array}{rr}
\langle D_1F(x_0^*,x_1^*);z-z\rangle\ + & \langle D_2F(x_0^*,x_1^*);x_1^*-x_1\rangle +\\
+\langle\beta D_1F(x_1^*,x_2^*);x_1^*-x_1\rangle\ + & \beta\langle D_2F(x_1^*,x_2^*);x_2^*-x_2\rangle +\\
+\beta\langle\beta D_1F(x_2^*,x_3^*);x_2^*-x_2\rangle\ + & \beta^2\langle D_2F(x_2^*,x_3^*);x_3^*-x_3\rangle +\\
+\beta^2\langle\beta D_1F(x_3^*,x_4^*);x_3^*-x_3\rangle\ + & \beta^3\langle D_2F(x_3^*,x_4^*);x_4^*-x_4\rangle +\\
\vdots & \\
+\beta^{T-1}\langle\beta D_1F(x_T^*,x_{T+1}^*);x_T^*-x_T\rangle\ + & \beta^T\langle D_2F(x_T^*,x_{T+1}^*);x_{T+1}^*-x_{T+1}\rangle.
\end{array}$$

Persze az első skaláris szorzat zérus. Most vegyük észre, hogy itt akármelyik sor utolsó tagjában és a következő sor első tagjában azonos a skaláris szorzat második tényezője. Ennek figyelembevételével csoportosítva az összeget a

$$\sum_{k=1}^{T}\beta^{k-1}\left(\langle D_2F(x_{k-1}^*,x_k^*)+\beta D_1F(x_k^*,x_{k+1}^*);x_k^*-x_k\rangle\right)+$$
$$+\ \langle\beta^TD_2F(x_T^*,x_{T+1}^*);x_{T+1}^*-x_{T+1}\rangle$$

felírást kapjuk. Mivel az Euler-egyenlet szerint minden k mellett a fenti összeg minden tényezője zérus, ezért az eddig számolást

$$\sum_{k=0}^{T}\beta^k\left(F(x_k^*,x_{k+1}^*)-F(x_k,x_{k+1})\right)\geq\langle\beta^TD_2F(x_T^*,x_{T+1}^*);x_{T+1}^*-x_{T+1}\rangle$$

az egyenlőtlenség foglalja össze. A $D_2F(x_T^*,x_{T+1}^*)=-\beta D_1F(x_{T+1}^*,x_{T+2}^*)$ Euler-egyenletet használva folytatjuk a becslést:

$$\begin{aligned}
&\sum_{k=0}^{T}\beta^k\left(F(x_k^*,x_{k+1}^*)-F(x_k,x_{k+1})\right)\geq\\
&\geq\langle-\beta^{T+1}D_1F(x_{T+1}^*,x_{T+2}^*);x_{T+1}^*-x_{T+1}\rangle=\\
&=\langle-\beta^{T+1}D_1F(x_{T+1}^*,x_{T+2}^*);x_{T+1}^*\rangle-\langle-\beta^{T+1}D_1F(x_{T+1}^*,x_{T+2}^*);x_{T+1}\rangle\geq\\
&\geq-\beta^{T+1}\langle D_1F(x_{T+1}^*,x_{T+2}^*);x_{T+1}^*\rangle.
\end{aligned}$$

Itt a középső sorbeli utolsó tagot azért hagyhattuk el, mert F az első változójában monoton növő, és mert az x_{T+1} állapot vektor minden koordinátája nem negatív. A transzverzalitási feltétel szerint

$$\begin{aligned}
&\sum_{k=0}^{\infty}\beta^k\left(F(x_k^*,x_{k+1}^*)-F(x_k,x_{k+1})\right)\geq\\
&\geq\lim_{T\to\infty}-\beta^{T+1}\langle D_1F(x_{T+1}^*,x_{T+2}^*);x_{T+1}^*\rangle=0.
\end{aligned}$$

Ezt kellett belátni. □

Most nézzük a transzverzalitási feltétel szükséges voltát indokló állítást.

7.6.11. tétel (transzverzalitási feltétel szükségessége). *Tegyük fel, hogy* $\bar{x} \in \Pi(x_0)$ *optimális belső megoldása az* (X, F, Γ, β) *programozásnak, ahol* X *nem tartalmaz negatív koordinátájú vektort, az* F *hozamfüggvény differenciálható, konkáv és az első változójában monoton növő;* $\beta < 1$*; a* Γ *átmenetfüggvény gráfja konvex és* $(0{,}0) \in \operatorname{graph}\Gamma$.

Ha $w(x_0)$ *véges, akkor az optimális tervre fennáll a transzverzalitási feltétel is, azaz*

$$\lim_{t\to\infty} \beta^t \langle D_1 F\left(x_t^*, x_{t+1}^*\right); x_t^* \rangle = 0.$$

Bizonyítás. Először mutassuk meg, hogy minden rögzített T-hez, létezik $\lambda_T \in (0{,}1)$, hogy minden $\lambda \in (\lambda_T, 1)$ mellett

$$(x_0^*, x_1^*, \cdots, x_T^*, \lambda x_{T+1}^*, \lambda x_{T+2}^* \cdots) \in \Pi(x_0).$$

Ugyanis $\bar{x}^*$ belső terv, ezért $x_{T+1}^* \in \operatorname{int}\Gamma(x_T^*)$, ezért létezik $\lambda_T < 1$, hogy $\lambda x_{T+1}^* \in \Gamma(x_T^*)$ fennáll minden $\lambda \in (\lambda_T, 1)$ mellett. Viszont $\operatorname{graph}\Gamma$ konvex és $(0{,}0); (x_t^*, x_{t+1}^*) \in \operatorname{graph}\Gamma$, tehát

$$(\lambda x_t^*, \lambda x_{t+1}^*) = ((1-\lambda)0 + \lambda x_t^*, (1-\lambda)0 + \lambda x_{t+1}^*) \in \operatorname{graph}\Gamma$$

is fennáll minden $t \geq T+1$ mellett. Az $\bar{x}^*$ optimalitása szerint

$$\sum_{t=0}^{T-1} \beta^t F(x_t^*, x_{t+1}^*) + \beta^T F(x_T^*, \lambda x_{T+1}^*) + \sum_{t=T+1}^{\infty} \beta^t F(\lambda x_t^*, \lambda x_{t+1}^*) \leq$$
$$\leq \sum_{t=0}^{T-1} \beta^t F(x_t^*, x_{t+1}^*) + \beta^T F(x_T^*, x_{T+1}^*) \sum_{t=T+1}^{\infty} \beta^t F(x_t^*, x_{t+1}^*),$$

ami azzal ekvivalens, hogy

$$\beta^T \left(F(x_T^*, \lambda x_{T+1}^*) - F(x_T^*, x_{T+1}^*)\right) \leq \sum_{t=T+1}^{\infty} \beta^t \left(F(x_t^*, x_{t+1}^*) - F(\lambda x_t^*, \lambda x_{t+1}^*)\right).$$

No de a $-F(\lambda(x_t^*, x_{t+1}^*)) \leq -(1-\lambda)F(0{,}0) - \lambda F(x_t^*, x_{t+1}^*)$ egyenlőtlenség is fennáll a $-F$ függvény konvexitása miatt, így

$$F(x_t^*, x_{t+1}^*) - F(\lambda x_t^*, \lambda x_{t+1}^*) \leq (1-\lambda)\left(F(x_t^*, x_{t+1}^*) - F(0{,}0)\right).$$

Azt kapjuk tehát, hogy

$$\beta^T \frac{F(x_T^*, \lambda x_{T+1}^*) - F(x_T^*, x_{T+1}^*)}{1-\lambda} \leq \sum_{t=T+1}^{\infty} \beta^t \left(F(x_t^*, x_{t+1}^*) - F(0{,}0)\right)$$

fennáll minden rögzített T és $\lambda_T < \lambda < 1$ esetén.

Most nézzük a bal oldali kifejezés számlálóját: Ha $h = 1 - \lambda$, akkor

$$F(x_T^*, \lambda x_{T+1}^*) = F(x_T^*, x_{T+1}^* - h x_{T+1}^*),$$

ezért bevezetve az $f(x) = F(x_T^*, x)$ jelölést, a $-x_{T+1}^*$ irány menti derivált definíciója szerint

$$\frac{F(x_T^*, \lambda x_{T+1}^*) - F(x_T^*, x_{T+1}^*)}{1-\lambda} = \frac{f(x_{T+1}^* + h(-x_{T+1}^*)) - f(x_{T+1}^*)}{h} \to$$
$$\to D_{-x_{T+1}^*} f(x_{T+1}^*) = \langle f'(x_{T+1}^*); -x_{T+1}^* \rangle = -\langle D_2 F(x_T^*, x_{T+1}^*); x_{T+1}^* \rangle,$$

midőn $\lambda \to 1$.

Az Euler-egyenletet szerint $D_2 F(x_T^*, x_{T+1}^*) = -\beta D_1 F(x_{T+1}^*, x_{T+2}^*)$ az optimális úton, ezért

$$\beta^{T+1} \langle D_1 F(x_{T+1}^*, x_{T+2}^*); x_{T+1}^* \rangle \le \sum_{t=T+1}^{\infty} \beta^t \left(F(x_t^*, x_{t+1}^*) - F(0{,}0) \right).$$

Tételünk feltétele szerint $u(\bar{x}^*) = \sum_{t=0}^{\infty} \beta^t F(x_t^*, x_{t+1}^*)$ sor konvergens, a $\beta < 1$ feltétel szerint $\sum_{t=0}^{\infty} \beta^t F(0{,}0)$ sor egy konvergens geometriai sor konstans szorosa, ezért a fenti kiemelt sor jobb oldalán lévő összeg egy konvergens sor T-edik végszelete. Másrészt a hozamfüggvény első változóbeli monotonitása, és az állapot vektorok nem negatívitása szerint a fent kiemelt sor bal oldala nem negatív, így $T \to \infty$ esetén

$$0 \le \beta^{T+1} \langle D_1 F(x_{T+1}^*, x_{T+2}^*); x_{T+1}^* \rangle \le \sum_{t=T+1}^{\infty} \beta^t \left(F(x_t^*, x_{t+1}^*) - F(0{,}0) \right) \to 0.$$

Ezt kellett belátni. □

7.7. Stabilitás

7.7.1. definíció (globális stabilitás)**.** Legyen (X, d) metrikus tér, $g : X \to X$ függvény, $\bar{x} \in X$ ennek fixpontja. Azt mondjuk, hogy az $\bar{x}$ fixpont globálisan stabil, ha minden $x_0 \in X$ mellett az $x_{t+1} = g(x_t)$ iterációra $x_t \to \bar{x}$.

7.7.1. Ljapunov-függvény

7.7.2. definíció (Ljapunov-függvény)**.** Legyen (X, d) egy kompakt metrikus tér, $g : X \to X$ folytonos függvény, és $\bar{x}$ ennek fixpontja. Azt mondjuk, hogy az $L : X \to \mathbb{R}$ folytonos függvény a g *Ljapunov-függvénye*, ha

– $L(x) \le L(\bar{x})$ minden $x \in X$ mellett.

– $L(x) < L(g(x))$ minden $x \ne \bar{x}$ esetén.

7.7.3. állítás (Ljapunov-függvény létezése a globális stabilitás elegendő feltétele)**.** *Legyen (X, d) kompakt metrikus tér, $g : X \to X$ folytonos függvény, $\bar{x} \in X$ a g fixpontja. Tegyük fel, hogy $\bar{x}$-nak létezik Ljapunov-függvénye. Ekkor $\bar{x}$ globálisan stabil.*

7.7.4.
Bizonyítás előtt emlékeztetni szeretnénk arra, hogy ha egy kompakt térnek van egy olyan sorozata, amelynek egyetlen torlódási pontja van, akkor az konvergál is a torlódási pontjához.

Bizonyítás. Legyen $x_0 \in X$ tetszőlegesen rögzítve, és $x_{t+1} = g(x_t)$ minden $t \in \mathbb{N}$ mellett. Megmutatjuk, hogy a sorozatnak csak $\bar{x}$ a torlódási pontja. Világos, hogy $L(x_t) \leq L(g(x_t)) = L(x_{t+1}) \leq L(\bar{x})$, tehát az $L(x_t)$ olyan valós sorozat, amely monoton és korlátos, így konvergens is. Legyen $\lambda \in \mathbb{R}$ a határérték, tehát $L(x_t) \to \lambda$. Ha $x \in X$ egy torlódási pontja a sorozatnak, akkor létezik egy $x_{t_j} \to x$ konvergens részsorozat. Így L folytonossága miatt $L(x_{t_j}) \to L(x) = \lambda$.

$$L(x) - L(g(x)) =$$
$$L(x) - L(x_{t_j+1}) + L(x_{t_j+1}) - L(g(x)) = \lambda - L(x_{t_j+1}) + L(g(x_{t_j})) - L(g(x)) \to 0,$$

hiszen $L \circ g$ folytonos, és az $L(x_t)$ sorozat $L(x_{t_j+1})$ részsorozatának is λ a határértéke. Megmutattuk tehát, hogy $x = \bar{x}$, tehát a sorozatnak csak egy torlódási pontja lehet, mégpedig $\bar{x}$. A tér kompaktsága miatt az x_t iteráció konvergens és $x_t \to \bar{x}$. Ezt kellett belátni. □

Persze a kérdés az, hogy adott $(g, \tilde{x})$ párhoz, g folytonos függvényhez és annak $\tilde{x}$ fixpontjához, milyen módon találhatunk Ljapunov-függvényt, amely biztosítja a globális stabilitást. A Ljapunov-függvénynek sokszor a konkrét feladathoz kötődő jelentése lehet. Fizikai példákban szerencsés lehet a Ljapunov-függvényt megválasztani, olyan módon amely a rendszer energiaszintjét fejezi ki. Egyszerű bolyongási példánkban a Ljapunov-függvény például a bolyongó pont tengerszint feletti magassága volt. Természetesen merül fel tehát, hogy akkor is adjunk a Ljapunov-függvény természetére valamilyen iránymutatást, amikor az (X, F, Γ, β) dinamikus programozási feladat g egyértékű és folytonos op-függvényének stabilitási tulajdonságait vizsgáljuk.

Az alábbi gondolat kísérletben tegyük fel, hogy fennáll a Bellman-egyenlet megoldásának szigorúan konkavitására vonatkozó, és az értékfüggvény deriválhatóságára vonatkozó tétel, azaz F szigorúan konkáv az értékfüggvény konkáv és differenciálható. Tudjuk, hogy ilyenkor a g op-leképezés egyértékű és folytonos. Tegyük fel még azt is, hogy X állapot tér egy kompakt, konvex halmaz. Ekkor g-nek létezik $\tilde{x}$ fixpontja (Brouwer-fixponttétel). Azt vizsgáljuk, hogy az így kapott $(g, \tilde{x})$ rendszerhez van-e valamilyen értelemben természetes Ljapunov-függvény.

Egy $v : X \to \mathbb{R}$ függvény szigorú konkavitásának szükséges és elegendő feltétele, hogy a $v' : X \to X$ derivált függvény szigorúan monoton fogyó legyen, azaz definíció szerint, minden $x, y \in X, x \neq y$ mellett

$$\langle v'(x) - v'(y); x - y \rangle < 0$$

teljesüljön.

Alkalmazzuk először ezt a fenti (X, F, Γ, β) programozás F hozamfüggvényére. Tekintsük az $(x, g(x))$ és $(\tilde{x}, \tilde{x})$ szereposztást. Így, ha $x \neq \tilde{x}$, akkor

$$\langle D_1F(x, g(x)) - D_1F(\tilde{x}, \tilde{x}); x - \tilde{x} \rangle + \langle D_2F(x, g(x)) - D_2F(\tilde{x}, \tilde{x}); g(x) - \tilde{x} \rangle < 0.$$

Ha g az egyértékű op-függvény, akkor a g iterálásával egy a kezdőpontból induló optimális pályát kapunk. Ez azt jelenti, hogy az Euler-egyenlet szerint $D_2F(x, g(x)) = -\beta D_1F(g(x), g(g(x)))$. A fent kiemelt egyenlőtlenség tehát

$$\begin{aligned} &\langle D_1F(x, g(x)) - D_1F(\tilde{x}, \tilde{x}); x - \tilde{x} \rangle < \\ &< \beta \langle D_1F(g(x), g(g(x))) - D_1F(\tilde{x}, \tilde{x}); g(x) - \tilde{x} \rangle \end{aligned}$$

alakú. Így $v'(x) = D_1F(x, g(x))$ szerint azt kapjuk, hogy

$$\langle v'(x) - v'(\tilde{x}); x - \tilde{x} \rangle < \beta \langle v'(g(x)) - v'(\tilde{x}); g(x) - \tilde{x} \rangle$$

Ha tehát bevezetjük a $L(x) = \langle v'(x) - v'(\tilde{x}); x - \tilde{x} \rangle$ jelölést, akkor $L(x) < \beta L(g(x))$ fennáll minden $x \neq \tilde{x}$ mellett. De a v értékfüggvény még konkáv is, ezért $L(x) \leq 0$ is fennáll minden $x \in X$ mellett. Ez azt jelenti, hogy L kis híján Ljapunov-függvény!

Ha a v értékfüggvény olyannyira szigorúan konkáv, hogy még az

$$\beta \langle v'(x) - v'(\tilde{x}); x - \tilde{x} \rangle < \beta \langle v'(g(x)) - v'(\tilde{x}); g(x) - \tilde{x} \rangle$$

egyenlőtlenség is fennáll, akkor $L(x) = \beta \langle v'(x) - v'(\tilde{x}); x - \tilde{x} \rangle$ valóban Ljapunov-függvény. Konklúzióként kapjuk tehát, hogy az

$$L(x) = \beta \langle v'(x) - v'(\tilde{x}); x - \tilde{x} \rangle$$

függvény egy természetes jelölt egy Ljapunov-féle próba függvénynek.

8. fejezet

Sztochasztikus eset

8.1. Előzetes példa

A sztochasztikus dinamikus programozási feladat alapvetően abban különbözik a determinisztikus dinamikus programozástól, hogy az út egyik x_t pontjáról a következő x_{t+1} pontra lépve az átlépés azonnali hozadékát meghatározó $F(x_t, x_{t+1})$ érték előre nem ismert. Arról van szó, hogy ez az érték függhet még egy véletlennek tekintett előre biztosan nem tudható kimenetelű eseménytől is. Magától értetődik, hogy az $F(x_t, x_{t+1})$ érték helyett számoljunk, annak várható értékét. Ezt szemlélteti az alábbi példa.

8.1.1.
Két vállalatunk van A és B, amelyek tökéletes helyettesítő terméket gyártanak, azaz a termékük iránti együttes kereslet teljes egészében az alacsonyabb árat kínáló vállalat felé fordul, vagy azonos árak esetében, a kereseten azonos arányban osztoznak. A kereslet függ egy előre nem megfigyelhető tényezőtől, amit most esőnek mondunk. Ha adott napon esik az eső, akkor $D = 0$ kereslet jelentkezik a termékek iránt. Mindkét cég 1 vagy 2 árat választhat, és a kereslet adott p_A és p_B árak mellett esőmentes napon

$$D = 10 - \min\{p_A, p_B\}.$$

Az eső minden nap $\frac{1}{2}$ valószínűséggel esik. A cégek minden napra más és más árat határozhatnak meg, a naponkénti diszkontráta $1 > \beta \geq 0$. Az adott napi árakat előző este kell meghatározniuk, tehát az árképzéskor nem tudják esni fog-e vagy sem.

A B vállalat nagyon egyszerű árstratégiát követ, amelyet A vállalat is ismer. A B által előállított termék ára $p_B = 2$ kezdetben, amelyet addig a napig tart B mígnem érzékeli, hogy terméke iránt nincs kereslet. Ekkor B változtat

nem esik	0	1	...	N	$N+1$
p_A	1	1	...	1	?
p_B	2	1	...	1	2
$\pi(A)$	9	$\frac{9}{4}$	...	$\frac{9}{4}$	v

esik	0	1	...	N	$N+1$
p_A	1	1	...	1	?
p_B	2	1	...	1	2
$\pi(A)$	0	$\frac{9}{4}$	...	$\frac{9}{4}$	v

8.1. táblázat. Árak és A várható bevételei esőmentes és esős kezdeti nappal, $p_A = 1$.

nem esik	0	1
p_A	2	?
p_B	2	2
$\pi(A)$	8	v

esik	0	1	...	N	$N+1$
p_A	2	1	...	1	?
p_B	2	1	...	1	2
$\pi(A)$	0	$\frac{9}{4}$	...	$\frac{9}{4}$	v

8.2. táblázat. Árak és A várható bevétele esőmentes és esős kezdeti nappal, $p_A = 2$.

és $p_B = 1$ árat alkalmaz N egymás utáni napon át, majd az $N+1$-edik napon újra $p_B = 2$, és így tovább.

Hogyan kell az A vállalat árait megválasztani, a bevétele diszkontált összegének maximalizálása céljából?

Az olyan napokra, amikor A tudja, hogy $p_B = 1$ lesz másnap, $p_A = 1$ árat kell előírnia, hiszen $p_A = 2$ a B iránti keresletet 0-ra nem változtathatja, és A nyilván jobban jár, ha valamennyi terméket elad, mintha semmit nem adna el. Az induló napon 4 lehetséges eset van: $p_A = 1$, $p_A = 2$ és mindkét esetben esik az eső vagy nem esik. Írjuk fel ennek megfelelően az A várható bevételeit!

A 8.1. táblázat bal oldalán a $p(A) = 1$ eset esőmentes 0. nap melletti árstratégiát találjuk. Az első sor jelöli a napok számát. A-nak a 0. nap ára felől kell döntenie. A második és a harmadik sor az A, illetve B vállalat árát tartalmazza a megfelelő napokra. Az utolsó sorban az egyes napokon keletkező várható bevétel értéke szerepel diszkontálás nélkül. A 0. napon ez 9 mivel $\pi(A) = D \cdot p_A = 9 \cdot 1$. Mivel B-re nem jut kereslet, ezért B stratégiája szerint N egymás utáni napon $p_B = 1$. Emiatt ezeken a napokon $p_A = 1$ is fennáll. Az ilyen napokon $\pi(A) = \frac{1}{2} \cdot D \cdot p_A = \frac{1}{2} \cdot \frac{9}{2} \cdot 1 = \frac{9}{4}$, hiszen a kereslet $\frac{1}{2}$ valószínűséggel 0 vagy $\frac{9}{2}$. Az $N+1$-edik napon B újra 2-re állítja az árát, ami azt jelenti, hogy az A vállalat éppen abba a döntési helyzetben kerül az $N+1$-edik napon, mint amiben az első napon volt. Ha v jelöli az optimális pálya értékét, akkor az $N+1$-edik naptól a diszkontált értékek összege éppen v.

A 8.1. táblázat jobb oldalán az esős 0. nappal kezdődő árstratégiát rögzítjük. Világos, hogy a 0. nap bevételétől eltekintve a táblázat megegyezik az esőmentes esettel, hiszen B most is N napra $p_B = 1$ árat határoz meg.

Az A vállalat 0. napi profitja a 0, hiszen a keresletet a 0. napi eső zérusra változtatta.

Összefoglalva a 8.1. táblázatot, az A vállalat $p_A = 1$ kezdő ár melletti diszkontált várható bevétele

$$\pi_1(v) = \frac{1}{2}\left(9 + \sum_{k=1}^{N} \frac{9}{4}\beta^k + \beta^{N+1}v\right) + \frac{1}{2}\left(\sum_{k=1}^{N} \frac{9}{4}\beta^k + \beta^{N+1}v\right) = \frac{9}{2} + r + \beta^{N+1}v,$$

ahol az r konstans jelöli az $r = \frac{9}{4}\beta\frac{\beta^N - 1}{\beta - 1}$ értéket a jobb áttekinthetőség kedvéért.

Most nézzük a $p_A = 2$ induló árat tartalmazó két esetet a 8.2. táblázatok segítségével. Ha nem esik az eső a kezdő napon, akkor osztozik a két cég a kereseten, és A már 0. nap estéjén ugyanabba a döntési helyzetbe kerül, amiben előző nap volt. Ha viszont esik a 0. napon, akkor ugyanazt a stratégiát kell követnie A-nak, amit $p_A = 1$ mellett is követett. Így a 8.2. táblázat összefoglalása a $p_A = 2$ kezdőár melletti diszkontált várható bevétel:

$$\pi_2(v) = \frac{1}{2}(8 + \beta v) + \frac{1}{2}\left(\sum_{k=1}^{N} \frac{9}{4}\beta^k + \beta^{N+1}v\right) = 4 + \frac{1}{2}r + \frac{1}{2}v\left(\beta + \beta^{N+1}\right).$$

A feladat Bellman-egyenlete tehát

$$\begin{aligned} v &= \max\{\pi_1(v), \pi_2(v)\} = \\ &= \max\left\{\frac{9}{2} + r + \beta^{N+1}v, 4 + \frac{1}{2}r + \frac{1}{2}\left(\beta + \beta^{N+1}\right)v\right\}. \end{aligned}$$

Oldjuk meg a Bellman-egyenletet.

Először is azt vegyük észre, hogy $\beta = 0$ esetben $v = \frac{9}{2}$ a megoldás, ami azt jelenti, hogy $p_A = 1$-et választ az A vállalat.

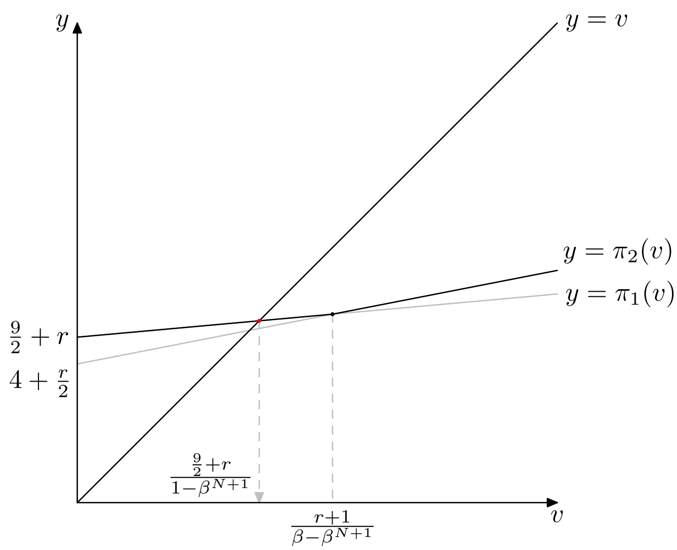
Ha $\beta > 0$, akkor a $\pi_1(v)$-t és $\pi_2(v)$ értékek v elsőfokú polinomjai. Ha $\pi_1(v)$ meredeksége m_1 és $\pi_2(v)$ meredeksége m_2, akkor

$$m_1 = \beta^{N+1} < \frac{1}{2}\left(\beta + \beta^{N+1}\right) = m_2 < 1.$$

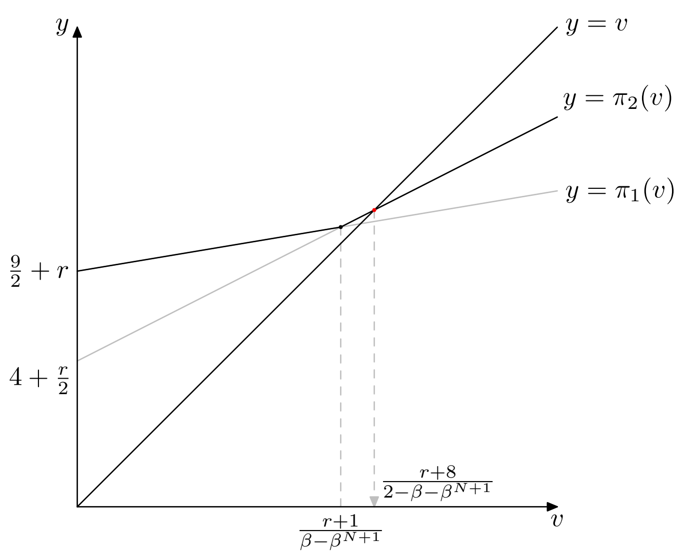
Innen azonnal látszik, hogy a Bellman-egyenletnek pontosan egy megoldása van.

Kis számolgatással kapjuk, hogy $\pi_1(v) = \pi_2(v)$ pontosan akkor, ha $v = \frac{r+1}{\beta - \beta^{N+1}}$, és $\pi_1(v) = v$ pontosan akkor, ha $v = \frac{\frac{9}{2} + r}{1 - \beta^{N+1}}$. Így a $p_A = 1$ választás szükséges és elegendő feltétele az

$$\frac{r + \frac{9}{2}}{1 - \beta^{N+1}} \le \frac{1 + r}{\beta - \beta^{N+1}}$$



8.1. ábra. $p_A = 1$ az optimális ár $N = 1$, $\beta = 0.3$ esetben



8.2. ábra. $p_A = 2$ az optimális ár $N = 10$, $\beta = 0.85$ mellett egyenlőtlenség teljesülése. Ekkor a feladat megoldásra $v = \frac{r+\frac{9}{2}}{1-\beta^{N+1}}$. Egy ilyen helyzetet látunk a 8.1. ábrán.

Hasonlóan, a $\pi_2(v) = v$ egyenlet megoldása $v = \frac{r+8}{2-\beta-\beta^{N+1}}$, emiatt a $p_A = 2$ választásának szükséges és elegendő feltétele az

$$\frac{1+r}{\beta - \beta^{N+1}} \leq \frac{r+8}{2-\beta-\beta^{N+1}}$$

egyenlőtlenség teljesülése. Ekkor a feladat értéke $v = \frac{r+8}{2-\beta-\beta^{N+1}}$. Egy ilyen eset illusztrációját a 8.2. ábrán követhetjük.

Összegezve az eddigieket azt kaptuk, hogy egyszerű eldönteni a β és az N paraméterek függvényében az A vállalat optimális árát. Ha az id függvény gráfja a törött vonalat annak második, meredekebbik részében metszi, akkor $p_A = 2$ az optimális választás, egyébként $p_A = 1$.

8.2. Sztochasztikus magok szorzata

A Markov-transzformáció elméletének felépítése során sokszor használjuk a Dynkin-tételt (1.1.18), annak következő kicsit egyszerűsített formájában:

8.2.1. tétel. *Legyen $\mathcal{H}$ egy monoton osztály, $\mathcal{P}$ egy félgyűrű, amelyekre $X \in \mathrm{r}(\mathcal{P}) \subseteq \mathcal{H}$. Ekkor $\sigma(\mathcal{P}) \subseteq \mathcal{H}$ is teljesül.*

Bizonyítás. Tudjuk, hogy egy algebra monoton osztály burka is algebra, ezért $\mathrm{m}(\mathrm{r}(\mathcal{P}))$ egy a $\mathcal{P}$-t tartalmazó σ-algebra. Így $\sigma(\mathcal{P}) \subseteq \mathrm{m}(\mathrm{r}(\mathcal{P})) \subseteq \mathcal{H}$ valóban fennáll. □

8.2.2. definíció (sztochasztikus mag)**.** Legyenek $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető terek. Azt mondjuk, hogy a

$$P : X \times \mathcal{N} \to [0,1]$$

függvény sztochasztikus mag e két mérhető téren, ha

1) $P(\cdot, B) : X \to [0,1]$ minden rögzített $B \in \mathcal{N}$ mellett $\mathcal{M}$-mérhető függvény;

2) $P(x, \cdot) : \mathcal{N} \to [0,1]$ minden rögzített $x \in X$ mellett egy valószínűségi mérték az $\mathcal{N}$ σ-algebrán.

8.2.3. definíció (Markov-transzformáció)**.** Legyen P egy sztochasztikus mag az $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető tereken, valamint $F : X \times Y \to \overline{\mathbb{R}}_+$ egy nem negatív függvény, amely az $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N})$ szorzat σ-algebrára nézve mérhető. Ekkor a $T_P F : X \to \mathbb{R}$ függvényt, az F függvény P szerinti *Markov-transzformáltjának* mondjuk, ahol

$$T_P F(x) = \int_Y F(x, y) P(x, dy).$$

8.2.4.
Időnként előfordul, hogy a Markov-transzformációt a fentitől kicsit eltérő környezetben használjuk. Ha $f : Y \to \mathbb{R}$ egy nem negatív mérhető függvény, akkor az természetes módon tekinthető egy $F : X \times Y \to \mathbb{R}$ mérhető függvénynek is az

$$F(x, y) = \chi_X(x) \cdot f(y)$$

definícióval. Világos, hogy tetszőleges $(x, y) \in X \times Y$ mellett $F(x, y) = f(y)$ és F is nem negatív mérhető függvény. Az ilyen tipusú f függvény Markov-transzformáltján egyszerűen a fenti F függvény

$$T_P f(x) = T_P F(x) = \int_Y f(y) P(x, dy)$$

Markov-transzformáltját értjük, és ezt a konvenciót a továbbiakban megjegyzés nélkül használjuk.

8.2.5.
Példaként nézzük meg azt a speciális esetet, amikor $F = \chi_{A \times B}$ alakú, ahol $A \times B \in \mathcal{A} \times \mathcal{B}$ egy mérhető tégla. Ha $x \in A$, akkor

$$\chi_{A \times B}(x, y) = \begin{cases} \chi_B(y), & \text{ha } x \in A; \\ 0 & \text{egyébként.} \end{cases}$$

Ezért

$$T_P F(x, y) = \int_Y \chi_{A \times B}(x, y) P(x, dy) = \chi_A(x) P(x, B).$$

A monoton konvergenciatételből könnyen látszanak a Markov-transzformáció következő tulajdonságai.

8.2.6. lemma. *Legyen $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető tér, valamint P ezen mérhető terek átmenetfüggvénye. Legyenek $F_n, F, G : X \times Y \to \mathbb{R}$ a szorzat σ-algebrára mérhető nem negatív függvények, $\alpha, \beta \geq 0$ valós számok. Ekkor*

1. $T_p(\alpha F + \beta G) = \alpha T_p F + \beta T_P G$*;*
2. $F \leq G$ *esetén* $T_P F \leq T_P G$.
3. *Ha* $F_n \to F$ *pontonként, és ha*
   a) $F_n \leq F_{n+1}$ *minden* n *mellett, akkor akkor* $T_P F_n \to T_P F$ *pontonként monoton növőleg;*
   b) $F_n \geq F_{n+1}$ *minden* n *mellett, és minden* $x \in X$*-re*

   $$\int_X F_1(x, y) P(x, dy) < \infty,$$

   akkor $T_P F_n \to T_P F$ *pontonként monoton fogyólag.*

4. $T_P(\sum_{n=1}^{\infty} F_n) = \sum_{n=1}^{\infty} T_P F_n$.

8.2.7.
Emlékezzünk arra, hogy egy nem negatív f függvény pontosan akkor mérhető, ha előáll

$$f = \sum_{n=1}^{\infty} \alpha_n \chi_{A_n}$$

alakban, ahol $\alpha_n \geq 0$ valós számok, és A_n mérhető halmazok.

8.2.8. állítás (mérhető függvény Markov-transzformáltja is mérhető)**.** *Legyenek $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető terek, P egy sztochasztikus mag a fenti mérhető tereken, T_P pedig e sztochasztikus mag által generált Markov-transzformáció. Ekkor $T_P F$ mérhető minden $F : X \times Y \to \mathbb{R}_+$ nem negatív szorzat-mérhető függvény esetén.*

Bizonyítás. Definiáljuk azon szorzat-mérhető halmazok halmazát, amelyek karakterisztikus függvényére az állítás igaz.

$$\mathcal{H} = \{Q \in \mathcal{M} \otimes \mathcal{N} : T_P \chi_Q \text{ mérhető}\}$$

· Láttuk, hogy $\mathcal{M} \times \mathcal{N} \subseteq \mathcal{H}$, hiszen $T_P \chi_{A \times B} = \chi_A P(\cdot, B)$, ami mérhető függvények összegeként maga is mérhető.
· Ha $A_n \times B_n \in \mathcal{H}$ diszjunkt mérhető téglák, és $Q = \cup_{n=1}^{N} A_n \times B_n$, akkor az egyesítés karaktersiztikus függvényére $\chi_Q = \sum_{n=1}^{N} \chi_{A_n \times B_n}$, ezért

$$T_P \chi_Q = \sum_{n=1}^{N} T_P \chi_{A_n \times B_n}.$$

Persze véges sok mérhető függvény összege is mérhető, ami azt jelenti, hogy $Q \in \mathcal{H}$, tehát $X \times Y \in \mathrm{r}(\mathcal{M} \times \mathcal{N}) \subseteq \mathcal{H}$.
· Látható, hogy $\mathcal{H}$ egy monoton osztály. Ugyanis, ha $Q_n \in \mathcal{H}$, $Q_n \subseteq Q_{n+1}$ bővülő halmazok sorozata, és $Q = \cup_{n=1}^{\infty} Q_n$, akkor $\chi_{Q_n} \to \chi_Q$ monoton növő módon, ezért

$$T_P \chi_{Q_n} \to T_P \chi_Q.$$

Látjuk tehát, hogy $T_P \chi_Q$ mérhető függvények pontonkénti hatáértékeként maga is mérhető, tehát $T_P \chi_Q \in \mathcal{H}$.
Analóg módon kapjuk, hogy $T_P \chi_Q \in \mathcal{H}$ olyan $Q_n \in \mathcal{H}$ sorozatra is, ahol a Q_n halmazok szűkülnek.
· Alkalmazható tehát az $\mathcal{M} \times \mathcal{N}$ félgyűrűre és a $\mathcal{H}$ monoton osztályra a Dynkin-tétel, amely szerint

$$\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N}) \subseteq \mathcal{H},$$

ami azt jelenti hogy a tétel tetszőleg $Q \in \mathcal{M} \otimes \mathcal{N}$ mérhető függvény χ_Q karakterisztikus függvényére igazolt.

· Ha most $F : X \times Y \to \mathbb{R}$ nem negatív mérhető függvény, akkor $F = \sum_{n=1}^{\infty} \alpha_n \chi_{Q_n}$ alakú, ahol $\alpha_n \geq 0$ és $Q_n \in \mathcal{M} \otimes \mathcal{N}$ mérhető halmazok. Láttuk, hogy

$$T_P F = \sum_{n=1}^{\infty} \alpha_n T_P \chi_{Q_n},$$

így $T_P F$ mérhető függvények határértékeként maga is mérhető. Ezt kellett belátni. □

8.2.9. definíció (sztochasztikus magok szorzata)**.** Tegyük fel, hogy adott két sztochasztikus mag. P_1 a $(W, \mathcal{W})$ és $(X, \mathcal{M})$ mérhető terek feletti sztochasztikus mag, és P_2 az $(X, \mathcal{M})$ és $(Y, \mathcal{N})$ mérhető terek feletti sztochasztikus mag. Definiáljuk a

$$P_3 : W \times \mathcal{M} \otimes \mathcal{N} \to [0,1]$$

függvényt az alábbi módon

$$P_3(w, Q) = \left(T_{P_1} \left(T_{P_2} \chi_Q\right)\right)(w) = \int_X \int_Y \chi_Q(x, y) P_2(x, dy) P_1(w, dx).$$

Ekkor P_3 sztochasztikus mag a $(W, \mathcal{W})$ és $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ mérhető tereken. Ezt nevezzük $P_3 = P_1 \cdot P_2$ *sztochasztikus szorzatnak.*

Bizonyítás. Először is tisztázzuk a definíció pontos jelentését. Legyen $F : W \times X \to \mathbb{R}$ függvény

$$F(w, x) = \left(T_{P_2} \chi_Q\right)(x) = \int_Y \chi_Q(x, y) P_2(x, dy)$$

módon definiálva, minden rögzített $w \in W$ mellett. Láttuk, hogy $F(w, \cdot) : X \to \mathbb{R}$ egy $\mathcal{M}$-mérhető függvény, amely nem függ az első változójától, így F persze $\mathcal{W} \otimes \mathcal{M}$ mérhető is. A $T_{P_1} F$ ezért definiált, és $T_{P_1} F : W \to [0,1]$ egy $\mathcal{W}$-mérhető függvény, ami azt jelenti, hogy minden rögzített $Q \in \mathcal{M} \otimes \mathcal{N}$ mellett $P_3(\cdot, Q) : W \to [0,1]$ függvény mérhető a $\mathcal{W}$ σ-algebrára nézve.

Most rögzített $w \in W$ mellett, ha $Q = \cup_{n=1}^{\infty} Q_n$ diszjunkt egyesítés, akkor kétszer felhasználva a Markov-transzformáltnak a végtelen összeggel való felcserélhetőségét, azt kapjuk, hogy

$$P_3(w, Q) = \left(T_{P_1} \left(T_{P_2} \chi_{\cup_{n=1}^{\infty} Q_n}\right)\right)(w) = \left(T_{P_1} \left(T_{P_2} \sum_{n=1}^{\infty} \chi_{Q_n}\right)\right)(w) =$$

$$= \left(T_{P_1} \left(\sum_{n=1}^{\infty} T_{P_2} \chi_{Q_n}\right)\right)(w) = \sum_{n=1}^{\infty} \left(T_{P_1} \left(T_{P_2} \chi_{Q_n}\right)\right)(w) = \sum_{n=1}^{\infty} P_3(w, Q_n),$$

ami éppen azt jelenti, hogy $P_3(w,\cdot)$ egy mérték az $\mathcal{M}\otimes\mathcal{N}$ szorzat σ-algebrán. □

8.2.10.
Most is érdemes kiszámolnunk egy mérhető tégla mértékét. Ha $A\times B\in\mathcal{M}\times\times\mathcal{N}$, akkor

$$P_3(w,A\times B)=T_1(T_2\chi_{A\times B})(w)=T_1(\chi_A P_2(\cdot,B))(w)=$$
$$=\int_X \chi_A(x)P_2(x,B)P_1(w,dx)=\int_A P_2(x,B)P_1(w,dx).$$

Emlékezzünk rá, hogy egy félgyűrűn értelmezett véges mérték egyféleképpen terjeszthető ki a generált σ-algebrára. Így ha rögzített $w\in W$ mellett

$$\mu_w(A\times B)=\int_A P_2(x,B)P_1(w,dx)$$

módon definiáljuk az $\mathcal{M}\times\mathcal{N}$ felgyűrűn értelmezett mértéket, akkor $\tilde{\mu}_w$-gal jelölve a Caratheodory-kiterjesztéssel kapott mértéket $\tilde{\mu}_w(Q)=P_3(w,Q)$ minden $Q\in\mathcal{M}\otimes\mathcal{N}$ mellett.

8.2.11. tétel (Fubini)**.** *Legyenek $(W,\mathcal{W})$, $(X,\mathcal{M})$ és $(Y,\mathcal{N})$ mérhető terek, továbbá P_1 és P_2 sztochasztikus magok a $(W,\mathcal{W})$, $(X,\mathcal{M})$ mérhető terek és az $(X,\mathcal{M})$ és $(Y,\mathcal{N})$ mérhető terek közt. Legyen $P_3=P_1\cdot P_2$ ezek szorzata.*

Ekkor $T_{P_3}=T_{P_1}\circ T_{P_2}$, azaz minden $F:X\times Y\to\mathbb{R}$ nem negatív mérhető függvény esetén $T_{P_3}F=T_{P_1}\left(T_{P_2}F\right)$, tehát minden $w\in W$ mellett

$$\left(T_{P_3}F\right)(w)=\int_{X\times Y}F(x,y)P_3(w,d(x,y))=$$
$$=\int_X\int_Y F(x,y)P_2(x,dy)P_1(w,dx)=\left(T_{P_1}\left(T_{P_2}F\right)\right)(w).$$

Bizonyítás. Először karakterisztikus függvényre, majd nem negatív mérhető függvényre igazoljuk az állítást.

· Ha $F=\chi_Q$ egy $Q\in\mathcal{M}\otimes\mathcal{N}$ mérhető halmaz karakterisztikus függvénye, akkor a szorzat sztochasztikus mag definíciója szerint minden $w\in W$ mellett

$$T_{P_3}\chi_Q(w)=\int_{X\times Y}\chi_Q(x,y)P_3(w,d(x,y))=P_3(w,Q)=T_{P_1}\left(T_{P_2}\chi_Q\right)(w),$$

ami azt jelenti, hogy karakterisztikus függvény esetére beláttuk az állítást.

· Ha most F nem negatív mérhető függvény, akkor $F = \sum_{n=1}^{\infty} \alpha_n \chi_{Q_n}$ alakú, ahol $Q_n \in \mathcal{M} \otimes \mathcal{N}$ és $\alpha_n \geq 0$. Ezért

$$\begin{aligned}
(T_{P_3} F)(w) &= \left(T_{P_3} \sum_{n=1}^{\infty} \alpha_n \chi_{Q_n} \right)(w) = \sum_{n=1}^{\infty} \alpha_n \left(T_{P_3} \chi_{Q_n} \right)(w) = \\
&= \sum_{n=1}^{\infty} \alpha_n \left(T_{P_1} \left(T_{P_2} \chi_{Q_n} \right) \right)(w) = \sum_{n=1}^{\infty} \left(T_{P_1} \left(T_{P_2} \alpha_n \chi_{Q_n} \right) \right)(w) = \\
&= \left(T_{P_1} \sum_{n=1}^{\infty} \left(T_{P_2} \alpha_n \chi_{Q_n} \right) \right)(w) = \\
&= \left(T_{P_1} \left(T_{P_2} \sum_{n=1}^{\infty} \alpha_n \chi_{Q_n} \right) \right)(w) = \left(T_{P_1} \left(T_{P_2} F \right) \right)(w).
\end{aligned}$$

Ezt kellett belátni. □

8.3. Átmenetfüggvény sztochasztikus mag szorzata

Az átmenetfüggvény a sztochasztikus magok speciális esete, amikor a két mérhető tér egybeesik.

8.3.1. definíció (átmenetfüggvény)**.** Legyen $(Z, \mathcal{Z})$ egy mérhető tér. Azt mondjuk, hogy a

$$Q : Z \times \mathcal{Z} \to [0{,}1]$$

függvény egy *átmenetfüggvény* a $(Z, \mathcal{Z})$ mérhető téren, ha Q egy sztochasztikus mag a $(Z, \mathcal{Z})$ és $(Z, \mathcal{Z})$ mérhető terek felett.

A sztochasztikus mag definícióját ismételve, a Q átmenetfüggvény, ha $Q(z, \cdot) : \mathcal{Z} \to [0{,}1]$ mérték a $\mathcal{Z}$ σ-algebrán, és a $Q(\cdot, B) : Z \to [0{,}1]$ függvény mérhető a $\mathcal{Z}$ σ-algebrára nézve.

Most írjuk fel a sztochasztikus szorzat definícióját ebben a speciális esetben.

8.3.2. definíció. Legyen Q egy átmenetfüggvény a $(Z, \mathcal{Z})$ mérhető téren. Jelölje $\mu = Q$. Ekkor az átmenetfüggvény sztochasztikus szorzata $\mu^2 = \mu \cdot \mu$, ahol $\mu^2 : Z \times \mathcal{Z}^2 \to [0{,}1]$

$$\mu^2(z_0, B) = (T_\mu(T_\mu \chi_B))(z_0) = \int_Z \int_Z \chi_B(z_1, z_2) Q(z_1, dz_2) Q(z_0, dz_1)$$

tetszőleges $z_0 \in Z$ és $B \in \mathcal{Z}^2 = \mathcal{Z} \otimes \mathcal{Z}$ mellett.

8.3.3.
Nézzük meg most is a mérhető tégla esetét. Ha $B = A_1 \times A_2$ ahol $A_1, A_2 \in \mathcal{Z}$, akkor

$$\mu^2(z_0, A_1 \times A_2) = \int_{A_1} \mu(z_1, A_2)\mu(z_0, dz_1) = \int_{A_1}\int_{A_2} \mu(z_1, dz_2)\mu(z_0, dz_1)$$

8.3.4.
Ha $F : Z \times Z \to \mathbb{R}$ egy nem negatív függvény, akkor a Fubini-tétel szerint

$$\begin{aligned}\left(T_{\mu^2}F\right)(z_0) &= \int_{Z^2} F(z_1, z_2)\mu^2(z_0, d(z_1, z_2)) = \\ &= \int_Z\int_Z F(z_1, z_2)\mu(z_1, dz_2)\mu(z_0, dz_1) = \left(T_\mu\left(T_\mu F\right)\right)(z_0).\end{aligned}$$

8.3.5. definíció. Legyen Q átmenetfüggvény a $(Z, \mathcal{Z})$ mérhető téren. Definiáljuk a $\mu = Q$ sztochasztikus magot a $(Z, \mathcal{Z})$ és $(Z, \mathcal{Z})$ tér, valamint a μ^2 sztochasztikus magot a $(Z, \mathcal{Z})$ és $(Z, \mathcal{Z}^2)$ terek felett. Ekkor a $\mu^3 = \mu \cdot \mu^2$ sztochasztikus szorzat értelmes, és μ^3 a $(Z, \mathcal{Z})$ és $(Z^3, \mathcal{Z}^3)$ mérhető terek sztochasztikus magja. Ez azt jelenti, hogy $\mu^3 : Z \times \mathcal{Z}^3 \to [0,1]$ függvény, ahol minden $z_0 \in Z$ és minden $B \in \mathcal{Z}^3$ esetén

$$\mu^3(z_0, B) = \left(T_\mu(T_{\mu^2}\chi_B)\right)(z_0) = \int_Z\int_{Z^2} \chi_B(z_1, (z_2, z_3))\mu^2(z_1, d(z_2, z_3)\mu(z_0, dz_1).$$

8.3.6.
Persze a μ^2 szerint a Fubini-tételt felhasználva a fenti definíció a

$$\mu^3(z_0, B) = \int_Z\int_Z\int_Z \chi_B(z_1, z_2, z_3)\mu(z_2, dz_3)\mu(z_1, dz_2)\mu(z_0, dz_1)$$

alakban is írható.

8.3.7.
Nézzük meg most is a mérhető tégla esetét. Ha $B = A_1 \times A_2 \times A_3$, ahol $A_1, A_2, A_3 \in \mathcal{Z}$, akkor

$$\begin{aligned}\mu^3(z_0, A_1 \times A_2 \times A_3) &= \mu^3(z_0, A_1 \times (A_2 \times A_3)) = \\ &= \int_{A_1} \mu^2(z_1, A_2 \times A_3)\mu(z_0, dz_1) = \\ &= \int_{A_1}\int_{A_2}\int_{A_3} \mu(z_2, dz_3)\mu(z_1, dz_2)\mu(z_0, dz_1).\end{aligned}$$

8.3.8.
Ha $F : Z^3 \to \mathbb{R}$ egy nem negatív függvény, akkor a Fubini-tétel szerint

$$\begin{aligned}\left(T_{\mu^3} F\right)(z_0) &= \int_{Z^3} F(z_1, z_2, z_3) \mu^3(z_0, d(z_1, z_2, z_3)) = \\ &= \int_Z \int_{Z^2} F(z_1, z_2, z_3) \mu^2(z_1, d(z_2, z_3)) \mu(z_0, dz_1) = \\ &= \left(T_\mu \left(T_{\mu^2} F\right)\right)(z_0).\end{aligned}$$

Felhasználva a μ^2-re vonatkozó Fubini-tételt azt kapjuk, hogy

$$\begin{aligned}&\int_{Z^3} F(z_1, z_2, z_3) \mu^3(z_0, d(z_1, z_2, z_3)) = \\ &= \int_Z \int_Z \int_Z F(z_1, z_2, z_3) \mu(z_2, dz_3) \mu(z_1, dz_2) \mu(z_0, dz_1).\end{aligned}$$

8.3.9. definíció. Legyen Q átmenetfüggvény a $(Z, \mathcal{Z})$ mérhető téren. Rögzített $k \geq 1$ egész szám mellett definiáljuk a $\mu = Q$ sztochasztikus magot a $(Z, \mathcal{Z})$ és $(Z, \mathcal{Z})$ tér, valamint a μ^k sztochasztikus magot a $(Z, \mathcal{Z})$ és $(Z, \mathcal{Z}^k)$ terek felett. Ekkor a $\mu^{k+1} = \mu \cdot \mu^k$ sztochasztikus szorzat értelmes, és μ^{k+1} a $(Z, \mathcal{Z})$ és $(Z^{k+1}, \mathcal{Z}^{k+1})$ mérhető terek sztochasztikus magja. Ez azt jelenti, hogy $\mu^{k+1} : Z \times \mathcal{Z}^{k+1} \to [0,1]$ függvény, ahol minden $z_0 \in Z$ és minden $B \in \mathcal{Z}^{k+1}$ esetén

$$\mu^{k+1}(z_0, B) = \left(T_\mu(T_{\mu^k} \chi_B)\right)(z_0) = \int_Z \int_{Z^k} \chi_B(z_1, z_2^{k+1}) \mu^k(z_1, dz_2^{k+1}) \mu(z_0, dz_1).$$

Itt $z \in Z^{k+1}$ mellett $z = (z_1, \overbrace{z_2, z_3, ..., z_{k+1}}^{z_2^{k+1}}) = (z_1, z_2^{k+1})$, azaz z_2^{k+1} a $k+1$-dimenziós z vektornak a hátsó, tehát a másodiktól a $k+1$.-ig számított, k koordinátájából álló vektort jelöli.

8.3.10.
Persze a μ^k szerint Fubini-tételt felhasználva a fenti definíció

$$\mu^{k+1}(z_0, B) = \int_Z \cdots \int_Z \chi_B(z_1, \cdots, z_{k+1}) \mu(z_k, dz_{k+1}) \cdots \mu(z_0, dz_1)$$

alakban is írható.

8.3.11.
A $B = A_1 \times A_2 \times \cdots \times A_{k+1}$ – ahol $A_1, \ldots, A_{k+1} \in \mathcal{Z}$ –, mérhető tégla

esetében

$$\begin{aligned}\mu^{k+1}(z_0, A_1\times\cdots\times A_{k+1}) &= \mu^{k+1}(z_0, A_1 \times (A_2 \times \cdots A_{k+1})) = \\ &= \int_{A_1} \mu^k(z_1, A_2 \times \cdots \times A_{k+1})\mu(z_0, dz_1) = \\ &= \int_{A_1}\int_{A_2}\cdots\int_{A_{k+1}} \mu(z_k, dz_{k+1})\cdots\mu(z_1, dz_2)\mu(z_0, dz_1).\end{aligned}$$

8.3.12.
Ha $F : Z^{k+1} \to \mathbb{R}$ egy nem negatív függvény, akkor a Fubini-tétel szerint

$$\begin{aligned}\big(T_{\mu^{k+1}}F\big)(z_0) &= \int_{Z^{k+1}} F(z^{k+1})\mu^{k+1}(z_0, dz^{k+1}) = \\ &= \int_Z\int_{Z^k} F(z_1, z_2^{k+1})\mu^k(z_1, dz_2^{k+1})\mu(z_0, dz_1) = \big(T_\mu\left(T_{\mu^k}F\right)\big)(z_0).\end{aligned}$$

Felhasználva a μ^k-ra vonatkozó Fubini-tételt azt kapjuk, hogy

$$\begin{aligned}&\int_{Z^{k+1}} F(z^{k+1})\mu^{k+1}(z_0, dz^{k+1}) = \\ &\quad = \int_Z\int_Z\cdots\int_Z F(z_1, \ldots, z_{k+1})\mu(z_k, dz_{k+1})\cdots\mu(z_1, dz_2)\mu(z_0, dz_1).\end{aligned}$$

8.4. Átmenetfüggvénynek átmenetfüggvény szorzata

8.4.1. definíció (Átmenenetfüggvény szorzat)**.** Legyen Q egy átmenetfüggvény a $(Z, \mathcal{Z})$ mérhető téren. Jelölje μ^k ennek mint sztochasztikus magnak a k-szoros szorzatát. Láttuk, hogy μ^k egy sztochasztikus mag a $(Z, \mathcal{Z})$ és $(Z^k, \mathcal{Z}^k)$ mérhető terek felett. Legyen $Q^k : Z \times \mathcal{Z} \to [0,1]$ az alábbi függvény. Tetszőleges $z \in Z$ és $A \in \mathcal{Z}$ mellett

$$Q^k(z, A) = \mu^k(z, Z \times Z \cdots \times Z \times A).$$

Látható, hogy Q^k valóban egy átmenetfüggvény a $(Z, \mathcal{Z})$ mérhető téren.

8.4.2.
A fentiek speciális esete szerint az átmenetfüggvény szorzatra az alábbi rekurzió áll fenn

$$\begin{aligned}Q^{k+1}(z_0, A) &= \mu^{k+1}(z_0, Z \times \cdots \times Z \times A) = \mu^{k+1}(z_0, Z \times (Z \times \cdots \times A)) = \\ &= \int_Z \mu^k(z_1, Z \times \cdots \times A)\mu(z_0, dz_1) = \int_Z Q^k(z_1, A)Q(z_0, dz_1).\end{aligned}$$

8.4.1. Markov-operátor

A Markov-transzformációnak azt a speciális esetét, amikor egy átmenetfüggvény által generált Markov-transzformációt valamely egy változós függvényre alkalmazzuk, Markov-operátornak nevezzük. Formálisan is átismételve a következő definícióhoz jutunk:

8.4.3. definíció (Markov-operáció)**.** Legyen $(Z, \mathcal{Z})$ egy mérhető tér, és $Q : Z \times \mathcal{Z} \to [0,1]$ átmenetfüggvény. Legyen $F : Z \to \mathbb{R}$ nem negatív mérhető függvény. Ekkor

$$(M_Q F)(z_0) = \int_Z F(z) Q(z_0, dz).$$

8.4.4.
Mivel a Markov-operáció egy speciális Markov-transzformáció, azért ha $F : Z \to \mathbb{R}$ nem negatív mérhető függvény, akkor $MF : Z \to \mathbb{R}$ szintén nem negatív mérhető függvény. Az is látható, hogy M_Q korlátos, mérhető, nem negatív függvényekhez, korlátos, mérhető, nem nem negatív függvényt rendel.

8.4.5.
Írjuk fel M_Q operátor hatványait egy mérhető halmaz karakterisztikus függvényén. Látni fogjuk, hogy

$$M^n \chi_A(z_0) = Q^n(z_0, A)$$

teljesül minden n természetes számra: Legyen tehát $A \in \mathcal{Z}$.

$n = 1$ mellett:

$$M_Q(\chi_A)(z_0) = \int_Z \chi_A(z) Q(z_0, dz) = Q(z_0, A).$$

Ha igaz n-re az állítás, akkor $n+1$-re:

$$(M^{n+1}\chi_A)(z_0) = (MQ^n(\cdot, A))(z_0) = \int_Z Q^n(z, A) Q(z_0, dz) = Q^{n+1}(z_0, A),$$

a 8.4.2. szerint.

8.4.6. definíció (adjungált Markov-operáció)**.** Legyen $(Z, \mathcal{Z})$ egy mérhető tér, és $Q : Z \times \mathcal{Z} \to [0,1]$ átmenetfüggvény. Legyen $\mu : \mathcal{Z} \to [0,1]$ valószínűségi mérték. Ekkor

$$(M_Q^* \mu)(B) = \int_Z Q(z, B) \mu(dz).$$

8.4.7.
Az adjungált Markov-operáció egy valószínűségi mértékhez egy valószínűségi mértéket rendel. Ugyanis, ha $B = \cup_{n=1}^{\infty} B_n$ diszjunkt egyesítés, akkor a

monoton konvergenciatétel szerint

$$(M^*\mu)(B) = (M^*\mu)(\cup_{n=1}^\infty B_n) = \int_Z Q(z, \cup_{n=1}^\infty B_n)\mu(dz) =$$
$$= \int_Z \sum_{n=1}^\infty Q(z, B_n)\mu(dz) = \sum_{n=1}^\infty \int_Z Q(z, B_n)\mu(dz) = \sum_{n=1}^\infty (M^*\mu)(B_n).$$

8.4.8.
Példaképpen nézzük meg hogyan hat az adjungált Markov-operáció a Dirac-mértékre. Ha δ_{z_0} jelöli a z_0 pontra koncentrált Dirac-mértéket, akkor

$$(M^*\delta_{z_0})(B) = \int_Z Q(z, B)\delta_{z_0}(dz) = Q(z_0, B).$$

8.4.9.
Hangsúlyozandó az integrálnak a mértéktől való pozitív homogenitását, vezessük be az alábbi alternatív jelöléseket. Ha μ egy mérték és f egy nem negatív mérhető függvény, akkor

$$\int_Z f d\mu = \int_Z f(z)\mu(dz) = \mu(f) = \mu f = \langle \mu; f\rangle.$$

Ha tehát μ, ν valószínűségi mértékek és f nem negatív mérhető függvény, akkor minden $\alpha, \beta \geq 0$ mellett

$$\langle \alpha\mu + \beta\nu; f\rangle = \alpha\langle \mu; f\rangle + \beta\langle \nu; f\rangle$$

továbbá, ha $f = \sum_{n=1}^\infty f_n$ nem negatív mérhető függvények pontonkénti összege, akkor a monoton konvergenciatétel szerint

$$\langle \mu; f\rangle = \sum_{n=1}^\infty \langle \mu; f_n\rangle.$$

8.4.10.
A fenti jelölést használva a Q átmenetfüggvény által generált Markov- és adjungált Markov-operáció definíciójára:

$$(Mf)(z_0) = Q(z_0, f) = \langle Q(z_0, \cdot); f\rangle,$$

valamint

$$(M^*\mu)(B) = \mu(Q(\cdot, B)) = \langle \mu; Q(\cdot, B)\rangle.$$

A következő szabály az adjungált Markov- és a Markov-operáció kapcsolatát írja le.

8.4.11. állítás (asszociatív szabály)**.** *Legyen $(Z, \mathcal{Z})$ mérhető tér, és Q rajta értelmezett átmenetfüggvény. Jelölje M, illetve M^* a Q által generált Markov- és adjungált Markov-operátorokat. Ekkor minden f nem negatív mérhető függvényre, és minden a $\mathcal{Z}$ σ-algebrán értelmezett valószínűségi mértékre*

$$\langle \mu; Mf\rangle = \langle M^*\mu; f\rangle, \quad \textit{azaz} \quad \int_Z \int_Z f(z)Q(z_1, dz)\mu(dz_1) = \int_Z f(z)M^*\mu(dz).$$

Bizonyítás. A szokásos mértékelméleti technika. Először belátjuk az állítás tetszőleges mérhető halmaz karakterisztikus függvényére, majd nem negatív mérhető függvényre.

- Legyen először $f = \chi_B$ egy mérhető halmaz karakterisztikus függvénye. Ekkor
$$\langle \mu; M\chi_B\rangle = \langle \mu; Q(\cdot, B)\rangle = (M^*\mu)(B) = \langle M^*\mu; \chi_B\rangle.$$
- Ha most f egy nem negatív mérhető függvény, akkor az előáll $f = \sum_{n=1}^{\infty} \alpha_n f_n$ alakban, ahol az f_n függvényekre már igazoltuk az állítást. Így a monoton konvergenciatétel miatt

$$\langle \mu; Mf\rangle = \left\langle \mu; M\left(\sum_{n=1}^{\infty} \alpha_n f_n\right)\right\rangle = \left\langle \mu; \sum_{n=1}^{\infty} \alpha_n M f_n\right\rangle = \sum_{n=1}^{\infty} \alpha_n \langle \mu; Mf_n\rangle =$$
$$= \sum_{n=1}^{\infty} \alpha_n \langle M^*\mu; f_n\rangle = \left\langle M^*\mu; \sum_{n=1}^{\infty} \alpha_n f_n\right\rangle = \langle M^*\mu; f\rangle.$$

Ezt kellett belátni. □

8.4.12. állítás. *A fenti feltételek mellett minden $n \in \mathbb{N}$ természetes számra*

$$\langle \mu; M^n f\rangle = \langle (M^*)^n \mu; f\rangle.$$

Bizonyítás. Az $n = 1$ esetet az imént láttuk. Tegyük fel, hogy igaz az állítás n-re, és tetszőleges nem negatív mérhető függvényre. Most lássuk be $n+1$-re.

$$\langle \mu; M^{n+1} f\rangle = \langle \mu; M(M^n f)\rangle = \langle M^*\mu; M^n f\rangle =$$
$$= \langle (M^*)^n (M^*\mu); f\rangle = \langle (M^*)^{n+1}\mu; f\rangle.$$ □

8.4.13.

Írjuk fel az adjungált Markov-operátor hatványait, egy Dirac-mértéken. Azt kapjuk, hogy

$$(M^*)^n \delta_{z_0} = Q^n(z_0, \cdot),$$

ugyanis alkalmazzuk az n-edik hatványokra vonatkozó asszociatív szabályt a δ_{z_0} Dirac-mértékre és a χ_B karakterisztikus függvényre. Ekkor

$$(M^*)^n \delta_{z_0}(B) = \langle (M^*)^n \delta_{z_0}; \chi_B\rangle = \langle \delta_{z_0}; M^n \chi_B\rangle = \langle \delta_{z_0}; Q^n(\cdot, B)\rangle = Q^n(z_0, B).$$

8.4.14. állítás (Chapman–Kolmogorov-azonosság)**.** *Legyen Q egy átmenetfüggvény a $(Z, \mathcal{Z})$ mérhető téren. Ekkor minden $z_0 \in Z$ és minden $B \in \mathcal{Z}$ mérhető halmazra*

$$\int_Z Q^n(z, B)Q^m(z_0, dz) = Q^{n+m}(z_0, B).$$

Bizonyítás. Láttuk, hogy $Q^m(z_0, \cdot) = (M^*)^m \delta_{z_0}$ és $Q^n(\cdot, B) = M^n \chi_B$. Így

$$\begin{aligned}\langle Q^m(z_0, \cdot); Q^n(\cdot, B)\rangle &= \langle (M^*)^m \delta_{z_0}; M^n \chi_B\rangle = \langle \delta_{z_0}; (M^m(M^n \chi_B)\rangle = \\ &= \langle \delta_{z_0}; M^{m+n}\chi_B\rangle = \langle \delta_{z_0}; Q^{m+n}(\cdot, B)\rangle = Q^{m+n}(z_0, B).\end{aligned}$$

Ezt kellett belátni. □

8.4.15. állítás. *Ha M a Q átmenetfüggvény által generált Markov-operátor, akkor M^n a Q^n átmenetfüggvény által generált Markov-operátor, azaz*

$$(M^n f)(z_0) = \int_Z f(z)Q^n(z_0, dz), \text{ továbbá } ((M^*)^n \mu)(B) = \int_Z Q^n(z, B)\mu(dz).$$

Bizonyítás. Az eddig kialakított kalkulus szerint

$$(M^n f)(z_0) = \langle \delta_{z_0}; M^n f\rangle = \langle (M^*)^n \delta_{z_0}; f\rangle = \langle Q^n(z_0, \cdot); f\rangle = \int_Z f(z)Q^n(z_0, dz).$$

Hasonlóan számolva

$$\begin{aligned}((M^*)^n \mu)(B) &= \langle (M^*)^n \mu; \chi_B\rangle = \langle \mu; M^n \chi_B\rangle = \\ &= \langle \mu; Q^n(\cdot, B)\rangle = \int_Z Q^n(z, B)\mu(dz).\end{aligned}$$ □

8.5. A sztochasztikus dinamikus programozási feladat

8.5.1. (SP feladat)
A lenti hat objektum határozza meg a dinamikus programozás alapfeladatát, amelyet *szuprémum feladatnak* fogunk mondani és röviden *(SP) feladatként* fogunk rá hivatkozni. Az

$$((X, \mathcal{M}), (Z, \mathcal{Z}), \mu, F, \Gamma, \beta)$$

hatost *sztochasztikus dinamikus programozási alapfeladatnak*, vagy rövidebben *sztochasztikus dinamikus programozásnak*, vagy még rövidebben *sztochasztikus programozásnak* nevezzük.

Állapottér: egy $(X, \mathcal{M})$ mérhető tér, amelynek elemeit állapotoknak mondjuk.

Sokktér: egy $(Z, \mathcal{Z})$ mérhető tér, amelynek elemeit véletlen sokkoknak nevezzük.

Átmenetfüggvény: $\mu : Z \times \mathcal{Z} \to [0{,}1]$ sztochasztikus mag a $(Z, \mathcal{Z})$ és $(Z, \mathcal{Z})$ mérhető terek felett. Adott $B \in \mathcal{Z}$ mellett $(\mu(z_0, B))$ annak valószínűsége, hogy a z_0 sokkot egy B-be eső sokk követ.

Transzformáció függvény: egy $\Gamma : X \times Z \multimap X$ halmazértékű leképezés, amely azt írja le, hogy az x állapotból a z sokk bekövetkeztekor a $\Gamma(x, z) \subseteq X$ állapotok egyikébe juthatunk. E függvény gráfja:

$$\operatorname{graph}(\Gamma) = \{(x, y, z) \in X \times X \times Z : y \in \Gamma(x, z)\}.$$

Hozamfüggvény: egy $F : \operatorname{graph}(\Gamma) \to \mathbb{R}$ függvény, amelyre az $F(x, y, z) \in \in \mathbb{R}$ számot az x állapotból az y állapotba jutás hozadékának, profitjának, értékének, vagy pillanatnyi hasznosságának gondolunk a z sokk bekövetkeztekor.

Diszkont tényező: egy $1 \geq \beta \geq 0$ valós szám.

8.5.1. Megengedett út vagy pálya

Fontos jelölésbeli konvenció, hogy az alsó indexszel a koordinátákat jelöljük és a felső index a koordináták számának jelölésére szolgál. Ha Z^t a Z halmaz t-szeres Descartes-szorzata, akkor $z^t = (z_1, z_2, \ldots, z_{t-1}, z_t) \in Z^t$ mellett jelölje $z^{t-1} = (z_1, \ldots, z_{t-1}) \in Z^{t-1}$, tehát z^t-ből z^{t-1} a t-edik koordináta elhagyásával keletkezik.

8.5.2. definíció (út)**.** Adott $(X, \mathcal{M})$ állapottér, $(Z, \mathcal{Z})$ sokkok és $\Gamma : X \times \times Z \multimap X$ átmenetfüggvény mellett a

$$\pi = (\pi_0, \pi_1, \pi_2, \cdots, \pi_t, \cdots)$$

sorozatot *útnak*, vagy *pályának* mondjuk, ha $\pi_t : Z^t \to X$ mérhető függvény, amelyekre

$$\pi_t(z^t) \in \Gamma(\pi_{t-1}(z^{t-1}), z_t)$$

minden $t = 1{,}2, \cdots$ mellett. A megengedett utak halmazát Π jelöli.

Ha $(x_0, z_0) \in X \times Z$ egy rögzített pont, akkor azt mondjuk, hogy a $\pi \in \in \Pi$ terv az (x_0, z_0) pontból indul, vagy ebből a pontból megvalósítható, ha $\pi_0 \in \Gamma(x_0, z_0)$. Az (x_0, z_0) pontból megvalósítható tervek halmazát jelölje $\Pi(x_0, z_0)$.

A szokásos konvencióknak megfelelően, a nulla-változós függvényt egy konstansnak tekintjük, mert nincs argumentuma. Így a π terv megadása azt jelenti, adott a

– $\pi_0 \in X$ konstans;
– $\pi_1 : Z \to X$ mérhető függvény, amelyre $\pi_1(z_1) \in \Gamma(\pi_0, z_1)$;
– $\pi_2 : Z^2 \to X$ mérhető függvény, amelyre $\pi_2(z_1, z_2) \in \Gamma(\pi_1(z_1), z_2)$;
– $\pi_3 : Z^3 \to X$ mérhető függvény, amelyre $\pi_3(z_1, z_2, z_3) \in \Gamma(\pi_2(z_1, z_2), z_3)$;
$\vdots$
– $\pi_t : Z^t \to X$ mérhető függvény, amelyre

$$\pi_t(z_1, z_2, \cdots, z_t) \in \Gamma(\pi_{t-1}(z_1, \cdots, z_{t-1}), z_t).$$

Ez azt jelenti, hogy a terv megadásával elhatározzuk, hogy ha a $z_1, z_2, \cdots, z_t$ sokkok érnek minket, akkor az x_t állapotból, az $x_{t+1} = \pi_t(z_1, \cdots, z_t)$ állapotba lépünk. Persze az átmenetfüggvény intuíciója szerint x_{t+1} csak $\Gamma(x_t, z_t)$-ből választható, így a t-edik függvénytől valóban azt kell elvárnunk, hogy az

$$x_{t+1} = \pi_t(z^t) \in \Gamma(\pi_{t-1}(z^{t-1}), z_t) = \Gamma(x_t, z_t)$$

tartalmazás teljesüljön. A kérdés, hogy ilyen tervet el lehet-e készíteni egyáltalán. Olyan feltételeket kell tennünk, amelyek biztosítják, hogy a tervek halmaza ne legyen üres!

8.5.3. állítás (út létezése)**.** *Tegyük fel, hogy a $\Gamma : X \times Z \multimap X$ átmenetfüggvénynek létezik mérhető szelekciója, azaz létezik $h : X \times Z \to X$ mérhető függvény, amelyre $h(x, z) \in \Gamma(x, z)$ minden $(x, z) \in X \times Z$ esetén. Jelöljön h_k ezen a mérhető szelekciók közül megszámlálhatóan sokat, tehát $k \in \mathbb{N}$. Ekkor tetszőleges $\pi_0 \in X$ állapot mellett*

$$\pi_t(z^t) = h_t(\pi_{t-1}(z^{t-1}), z_t)$$

rekurzióval megadott függvényekre $\pi = (\pi_0, \pi_1, \cdots, \pi_t, \cdots)$ egy a Γ átmenetfüggvény által generált út.

Bizonyítás. A rekurzió első lépése szerint

$$\pi_1(z_1) = h_1(\pi_0, z_1) \in \Gamma(\pi_0, z_1)$$

Itt az első változó rögzített és h_1 mérhető függvény. Tudjuk mérhető függvény szelete is mérhető, emiatt $\pi_1 : Z \to X$ is mérhető. Hasonlóan, ha $\pi_{t-1} : Z^{t-1} \to X$ már mérhető, akkor

$$\pi_t(z^t) = h_t(\pi_{t-1}(z^{t-1}), z_t) \in \Gamma(\pi_{t-1}(z^{t-1}), z_t)$$

szerint π_t mérhető függvény kompozíciójaként maga is mérhető. Így valamennyi π_t függvény mérhetőségét igazoltuk. $\square$

8.5.4. definíció (út folytatása)**.** Tekintsünk egy $\pi \in \Pi$ tervet, és legyen $w^T \in Z^T$ egy rögzített T-sokk. E π útnak a w^T T-sokk utáni folytatásán a következő utat értjük:

$$c(\pi, w^T) = (c_0^{\pi,w^T}, c_1^{\pi,w^T}, c_2^{\pi,w^T}, \cdots, c_t^{\pi,w^T}, \cdots),$$

ahol

$$c_0^{\pi,w^T} = \pi_T(w^T) \quad \text{és} \quad c_t^{\pi,w^T}(z^t) = \pi_{t+T}(w^T, z^t) \qquad \text{ha } t = 1,2,\ldots$$

8.5.5.
Világos, hogy $c_0^{\pi,w^T} = \pi_T(w^T) \in \Gamma(\pi_{T-1}(w^{T-1}), w_T)$ a π út definíciója miatt

$$c_t^{\pi,w^T}(z^t) = \pi_{t+T}(w^T, z^t) \in \Gamma(\pi_{t-1+T}(w^T, z^{t-1}), z_t) = \Gamma(c_{t-1}^{\pi,w^T}(z^{t-1}), z_t)$$

ami azt jelenti, hogy $x_T = \pi_{T-1}(w^{T-1})$ állapot mellett $c(\pi, w^T) \in \Pi(x_T, w_T)$ az (x_T, w_T) pontból megvalósítható út.

8.5.6.
Érdemes átgondolni speciális esetként egy 1-sokk utáni folytatást. Ekkor
– $c_0^{\pi,z_1} = \pi_1(z_1) \in \Gamma(\pi_0, z_1)$;
– $c_1^{\pi,z_1}(z_2) = \pi_2(z_1, z_2) \in \Gamma(\pi_1(z_1), z_2)$;
– $c_2^{\pi,z_1}(z_2, z_3) = \pi_3(z_1, z_2, z_3) \in \Gamma(\pi_2(z_1, z_2), z_3)$;
$\vdots$
– $c_t^{\pi,z_1}(z_2, \ldots, z_t, z_{t+1}) = \pi_{t+1}(z_1, z_2, \ldots, z_t, z_{t+1}) \in \Gamma(\pi_t(z_1, \cdots, z_t), z_{t+1})$.

Kicsit tömörebb jelöléssel tehát azt írhatjuk, hogy

$$c_0^{\pi,z_1} = \pi_1(z_1) \quad \text{és} \quad c_t^{\pi,z_1}(z_2^{t+1}) = \pi_{t+1}(z^{t+1}),$$

valamint

$$c(\pi, z_1) = (c_0^{\pi,z_1}, c_1^{\pi,z_1}, c_2^{\pi,z_1}, \cdots, c_t^{\pi,z_1}, \cdots).$$

8.5.2. A szuprémum feladat

8.5.7. definíció (SP feladat)**.** Tegyük fel, hogy adott az

$$((X, \mathcal{M}), (Z, \mathcal{Z}), \mu, F, \Gamma, \beta)$$

sztochasztikus programozás.
· Legyen $\pi \in \Pi(x_0, z_0)$ egy az $(x_0, z_0) \in X \times Z$ pontból megvalósítható terv. E terv *azonnali hasznossága*

$$u_0(\pi, (x_0, z_0)) = F(x_0, \pi_0, z_0).$$

A terv *egylépéses hasznossága*:

$$u_1(\pi,(x_0,z_0) = F(x_0,\pi_0,z_0) + \beta \int_Z F(\pi_0,\pi_1(z_1),z_1)\mu(z_0,dz_1).$$

A terv *kétlépéses hasznossága*:

$$u_2(\pi,(x_0,z_0) = u_1(\pi,(x_0,z_0) + \beta^2 \int_{Z^2} F(\pi_1(z_1),\pi_2(z_1,z_2),z_2)\mu^2(z_0,d(z_1,z_2)).$$

A t*-edik lépésben számolt hasznosság*:

$$u_t(\pi,(x_0,z_0)) = u_{t-1}(\pi,(x_0,z_0)) + \beta^t \int_{Z^t} F(\pi_{t-1}(z^{t-1}),\pi_t(z^t),z_t)\mu^t(z_0,dz^t).$$

· A π útnak az (x_0,z_0) pontból számított *hasznossága*:

$$u(\pi,(x_0,z_0)) = \lim_{t\to\infty} u_t(\pi,(x_0,z_0)).$$

· Az (x_0,z_0) *induló helyzet hasznossága*, a sztochasztikus programozási feladat SP feladatának kiszámolását jelenti, azaz

$$w(x_0,z_0) = \sup\{u(\pi,(x_0,z_0)) : \pi \in \Pi(x_0,z_0)\}.$$

· Azt mondjuk, hogy a $\pi^* \in \Pi(x_0,z_0)$ terv az (x_0,z_0) kiinduló helyzetből induló optimális megoldása az (SP) feladatnak, ha a fent szuprémum éppen a π^* útnál vétetik fel, azaz

$$w(x_0,z_0) = u(\pi^*,(x_0,z_0)).$$

8.5.8. definíció. Adott $\pi \in \Pi(x_0,z_0)$ terv mellett a t-edik lépés várható hozadéka:

$$e_0(\pi,(x_0,z_0)) = F(x_0,\pi_0,z_0), \text{ ha } t=0, \text{ és}$$
$$e_t(\pi,(x_0,z_0)) = \int_{Z^t} F(\pi_{t-1}(z^{t-1}),\pi_t(z^t),z_t)\mu^t(z_0,dz^t), \text{ ha } t \geq 1.$$

8.5.9.
Tudjuk, hogy $\mu^t : Z \times \mathcal{Z}^t \to [0,1]$ egy sztochasztikus mag a $(Z,\mathcal{Z})$ és $(Z,\mathcal{Z}^t)$ mérhető terek felett. Rögzített F nem negatív mérhető függvényre, $\pi \in \Pi$ tervre, és x_0 kiinduló állapotra az $e_t(\pi,(x_0,\cdot))$ függvény nem más, mint a

$$z^t \mapsto F(\pi_{t-1}(z^{t-1}),\pi_t(z^t),z_t)$$

függvénynek mint $Z^t \to \mathbb{R}$ függvénynek a Markov-transzformáltja a μ^t sztochasztikus mag szerint. Mivel nem negatív mérhető függvény Markov-transzformáltja is mérhető, ezért $e_t(\pi,(x_0,\cdot)) : Z \to \mathbb{R}$ nem negatív mérhető függvény.

Látható, hogy ezekkel a jelölésekkel a $\pi \in \Pi(x_0, z_0)$ terv értékére

$$u(\pi, (x_0, z_0)) = \sum_{t=0}^{\infty} \beta^t e_t(\pi, (x_0, z_0)),$$

és így $u(\pi, (x_0, \cdot)) : Z \to \mathbb{R}$ is nem negatív mérhető függvény.

Ilyen módon a $w(x_0, \cdot) : Z \to \mathbb{R}$ függvény mérhetőségét nem tudjuk biztosítani, hisz ez a függvény szuprémuma az $u(\pi, (x_0, \cdot))$ alakú függvényeknek, de az utak számossága nyilvánvalóan több, mint megszámlálható.

8.5.3. Bellman-egyenlet

8.5.10. definíció (Bellman-egyenlet megoldása)**.** Tekintsük az

$$((X, \mathcal{M}), (Z, \mathcal{Z}), \mu, F, \Gamma, \beta)$$

sztochasztikus programozást, és legyen $v : X \times Z \to \mathbb{R}$ mérhető nem negatív függvény. Azt mondjuk, hogy v *kielégíti a feladat Bellman-egyenletét*, ha minden $(x_0, z_0) \in X \times Z$ induló érték mellett

$$v(x_0, z_0) = \sup_{y \in \Gamma(x_0, z_0)} \left\{ F(x_0, y, z_0) + \beta \int_Z v(y, z) \mu(z_0, dz) \right\}.$$

8.5.11. definíció (optimal policy leképezés)**.** Tekintsük az

$$((X, \mathcal{M}), (Z, \mathcal{Z}), \mu, F, \Gamma, \beta)$$

sztochasztikus programozást, és legyen $v : X \times Z \to \mathbb{R}$ a feladat Bellman-egyenletének egy megoldása. Definiáljuk az ehhez a megoldáshoz tartozó optimal policy (op) leképezést a következő képpen:

$$G(x_0, z_0) = \left\{ y \in \Gamma(x_0, z_0) : v(x_0, z_0) = F(x_0, y, z_0) + \beta \int_Z v(y, z) \mu(z_0, dz) \right\}.$$

Azt mondjuk, hogy a $\pi \in \Pi$ megvalósítható tervet az optimal policy leképezés generálja, ha

$$\pi_t(z^t) \in G(\pi_{t-1}(z^{t-1}), z_t)$$

is teljesül.

8.5.12. állítás (optimal policy leképezés által generált út létezése)**.** *Tegyük fel, hogy a sztochasztikus programozás Bellman-egyenletének egy megoldásához tartozó G optimal policy leképezésnek létezik mérhető szelekciója, azaz létezik $h : X \times Z \to X$ mérhető függvény, amelyre $h(x, z) \in G(x, z)$ minden*

$(x, z) \in X \times Z$ esetén. Jelöljön g_t ezen mérhető szelekciók közül megszámlálhatóan sokat, tehát $t \in \mathbb{N}$. Ekkor tetszőleges $\pi_0 \in X$ állapot mellett

$$\pi_t(z^t) = g_t(\pi_{t-1}(z^{t-1}), z_t)$$

rekurzióval megadott függvényekre $\pi = (\pi_0, \pi_1, \cdots, \pi_t, \cdots)$ egy a G optimal policy leképezés által generált út.

Bizonyítás. Alkalmazzuk az út létezésére vonatkozó állítást a Γ transzformációs leképezés helyett a G optimal policy leképezésre. $\square$

8.5.13. állítás. *Tegyük fel, hogy a $v : X \times Z \to \mathbb{R}$ függvény az*

$$((X, \mathcal{M}), (Z, \mathcal{Z}), \mu, F, \Gamma, \beta)$$

programozás Bellman-egyenletének olyan megoldása, amely megoldáshoz tartozó optimal policy függvénynek van mérhető szelekciója. Legyen az $(x_0, z_0) \in \in X \times Z$ olyan pont, amelyből kiinduló minden $\pi \in \Pi(x_0, z_0)$ út esetén

$$\beta^t \int_{Z^t} v(\pi_{t-1}(z^{t-1}), z_t)\mu^t(z_0, dz^t) \to 0, \quad \textit{midőn} \;\; t \to \infty.$$

Ekkor minden az (x_0, z_0) pontból induló, az optimal policy leképezés által generált út egyben az (x_0, z_0)-ból induló optimális út is, továbbá a Bellman-egyenlet megoldásának értéke az (x_0, z_0) pontban megegyezik az SP feladat (x_0, z_0)-beli értékével, azaz a Bellman-egyenlet megoldása éppen az SP feladat értékfüggvénye.

Bizonyítás. Legyen π^* az op-leképezés által generált út, melyre $\pi^* \in \Pi(x_0, z_0)$. Megmutatjuk, hogy a tett feltételek mellett

$$w(x_0, z_0) = u(\pi^*, (x_0, z_0)) = v(x_0, z_0)$$

is fennáll. Ehhez legyen $\pi \in \Pi(x_0, z_0)$ egy tetszőlegesen választott megengedett út. Először azt mutatjuk meg, hogy

$$u(\pi^*, (x_0, z_0)) \geq u(\pi, (x_0, z_0)).$$

Ehhez induljunk ki a Bellman-egyenlet megoldásának definíciójából. Mivel π az (x_0, z_0)-ból induló megengedett út, ezért $\pi_0 \in \Gamma(x_0, z_0)$, tehát

$$v(x_0, z_0) = \sup_{y \in \Gamma(x_0, z_0)} \left\{ F(x_0, y, z_0) + \beta \int_Z v(y, z_1)\mu(z_0, dz_1) \right\} \geq$$
$$\geq F(x_0, \pi_0, z_0) + \beta \int_Z v(\pi_0, z_1)\mu(z_0, dz_1).$$

De persze minden $z_1 \in Z$ mellett v a (π_0, z_1) pontban is kielégíti a Bellman-egyenletet és $\pi_1(z_1) \in \Gamma(\pi_0, z_1)$, ami azt jelenti, hogy $x_1 = \pi_0$ jelöléssel:

$$v(x_1, z_1) = \sup_{y \in \Gamma(x_1, z_1)} \left\{ F(x_1, y, z_1) + \beta \int_Z v(y, z_2)\mu(z_1, dz_2) \right\} \geq$$
$$\geq F(x_1, \pi_1(z_1), z_1) + \beta \int_Z v(\pi_1(z_1), z_2)\mu(z_1, dz_2)$$

is fennáll. Teljesen analóg módon, kihasználva, hogy v az $(\pi_{t-1}(z^{t-1}), z_t)$ pontban is kielégíti a Bellman-egyenletet, és $\pi_t(z^t) \in \Gamma(\pi_{t-1}(z^{t-1}), z_t)$ azt kapjuk, hogy $x_t = \pi_{t-1}(z^{t-1})$ rövidítéssel:

$$v(x_t, z_t) = \sup_{y \in \Gamma(x_t, z_t)} \left\{ F(x_t, y, z_t) + \beta \int_Z v(y, z_{t+1})\mu(z_t, dz_{t+1}) \right\} \geq$$
$$\geq F(x_t, \pi_t(z^t), z_t) + \beta \int_Z v(\pi_t(z^t), z_{t+1})\mu(z_t, dz_{t+1}).$$

Az eddigi számolást visszahelyettesítve azt kapjuk, hogy

$$v(x_0, z_0) \geq F(x_0, \pi_0, z_0) +$$
$$+ \sum_{t=1}^{T} \beta^t \int_Z \int_Z \cdots \int_Z F(\pi_{t-1}(z^{t-1}), \pi_t(z^t), z_t)\mu(z_{t-1}, dz_t)\mu(z_{t-2}, dz_{t-1}) \cdots$$
$$\cdots \mu(z_0, dz_1) +$$
$$+ \beta^{T+1} \int_Z \int_Z \cdots \int_Z v(\pi_T(z^T), z_{T+1})\mu(z_t, dz_{t+1})\mu(z_{t-1}, dz_t) \cdots \mu(z_0, dz_1),$$

ami a Fubini-tétel szerint azzal ekvivalens, hogy

$$v(x_0, z_0) \geq F(x_0, \pi_0, z_0) + \sum_{t=1}^{T} \beta^t \int_{Z^t} F(\pi_{t-1}(z^{t-1}), \pi_t(z^t), z_t)\mu(z_0, dz^t) +$$
$$+ \beta^{T+1} \int_{Z^{T+1}} v(\pi_T(z^T), z_{T+1})\mu(z_0, dz^{T+1}) =$$
$$= u_T(\pi, (x_0, z_0)) + \beta^{T+1} \int_{Z^{T+1}} v(\pi_T(z^T), z_{T+1})\mu(z_0, dz^{T+1}) \to u(\pi, (x_0, z_0)).$$

Mivel $v(x_0, z_0) \geq u(\pi, (x_0, z_0))$ minden $\pi \in \Pi(x_0, z_0)$ útra igaz, ezért a szuprémum felső korlát volta miatt

$$v(x_0, z_0) \geq w(x_0, z_0).$$

De a fenti számolást újra gondolva abban a speciális esetben, ha π^* egy a v-hez tartozó optimal policy leképezés által generált út, azt kapjuk, hogy az összes egyenlőtlenség egyenlőséggel teljesül. Ez azt jelenti, hogy ha veszünk egy az előző állítás által garantált konkrét π^* optimal policy leképezés által generált utat, akkor

$$v(x_0, z_0) = u(\pi^*, (x_0, z_0))$$

is fennáll, amiből már

$$v(x_0, z_0) = u(\pi^*, (x_0, z_0)) = w(x_0, z_0)$$

valóban adódik. Ezt kellett belátni. □

8.5.14. következmény. *Az* $((X, \mathcal{M}), (Z, \mathcal{Z}), \mu, F, \Gamma, \beta)$ *programozás Bellman-egyenletének legfeljebb egy olyan megoldása lehet, amelynek van mérhető szelekciója.*

Az alábbiakban a fenti állítás megfordítását fogjuk látni, nevezetesen, hogy az SP feladat egy optimális útja az a Bellman-egyenlet megoldásához tatozó, optimal policy leképezés által generált út is. Ehhez némi előkészítésre van szükségünk.

8.5.15. lemma (indukciós lemma)**.**

$$\begin{aligned}
e_{t+1}(\pi, (x_0, z_0)) &= \int_Z e_t(c(\pi, z_1), (\pi_0, z_1))\mu(z_0, dz_1), \\
u(\pi, (x_0, z_0)) &= F(x_0, \pi_0, z_0) + \beta \int_Z u(c(\pi, z_1), (\pi_0, z_1))\mu(z_0, dz_1).
\end{aligned}$$

Bizonyítás. A sztochasztikus magok szorzatára vonatkozó Fubini-tétel szerint:

$$\begin{aligned}
&e_{t+1}(\pi, (x_0, z_0)) = \\
&= \int_{Z^{t+1}} F(\pi_t(z^t), \pi_{t+1}(z^{t+1}), z_{t+1})\mu^{t+1}(z_0, dz^{t+1}) = \\
&= \int_{Z^{t+1}} F(\pi_t(z_1, z_2^t), \pi_{t+1}(z_1, z_2^{t+1}), z_{t+1})\mu^{t+1}(z_0, dz^{t+1}) = \\
&= \int_Z \int_{Z^t} F(\pi_t(z_1, z_2^t), \pi_{t+1}(z_1, z_2^{t+1}), z_{t+1})\mu^t(z_1, dz_2^{t+1})\mu(z_0, dz_1) = \\
&= \int_Z \int_{Z^t} F(c_{t-1}^{\pi_0, z_1}(z_2^t), c_t^{\pi_0, z_1}(z_2^{t+1}), z_{t+1})\mu^t(z_1, dz_2^{t+1})\mu(z_0, dz_1) = \\
&= \int_Z e_t(c(\pi, z_1), (\pi_0, z_1))\mu(z_0, dz_1).
\end{aligned}$$

A második állításhoz ezt felhasználva:

$$\begin{aligned}
u(\pi,(x_0,z_0)) &= \sum_{t=0}^{\infty}\beta^t e_t(\pi,(x_0,z_0)) = \\
&= F(x_0,\pi_0,z_0)+\beta\sum_{t=1}^{\infty}\beta^{t-1}e_t(\pi,(x_0,z_0)) = \\
&= F(x_0,\pi_0,z_0)+\beta\sum_{t=0}^{\infty}\beta^{t}e_{t+1}(\pi,(x_0,z_0)) = \\
&= F(x_0,\pi_0,z_0)+\beta\sum_{t=0}^{\infty}\beta^{t}\int_Z e_t(c(\pi,z_1),(\pi_0,z_1))\mu(z_0,dz_1) = \\
&= F(x_0,\pi_0,z_0)+\beta\int_Z\sum_{t=0}^{\infty}\beta^{t}e_t(c(\pi,z_1),(\pi_0,z_1))\mu(z_0,dz_1) = \\
&= F(x_0,\pi_0,z_0)+\beta\int_Z u(c(\pi,z_1),(\pi_0,z_1))\mu(z_0,dz_1).
\end{aligned}$$

Ezt kellett belátni. □

8.5.16. állítás (optimális terv folytatása is optimális terv)**.** *Tegyük fel, hogy az op-leképezésnek létezik mérhető szelekciója és a diszkonttényezőre $\beta > 0$. Ekkor egy (x_0,z_0) pontból induló $\pi^* \in \Pi(x_0,z_0)$ optimális terv esetén $\mu(z_0,\cdot)$ majdnem minden $z_1 \in Z$ sokk bekövetkezte mellett a terv z_1 utáni $c(\pi,z_1)$ folytatása is optimális marad.*

Formálisabban, ha $w(x_0,z_0) = u(\pi^*,(x_0,z_0))$, akkor

$$w(\pi_0^*,z_1) = u(c(\pi^*,z_1),(\pi_0^*,z_1))$$

a $\mu(z_0,\cdot)$ mérték szerint, majdnem minden $z_1 \in Z$ mellett.

Bizonyítás. Legyen $g : X \times Z \to X$ az op-függvény egy mérhető szelekciója, és definiáljuk a

$$\pi^g = (\pi_0^*, \pi_1^g, \pi_2^g, \cdots, \pi_t^g, \cdots)$$

az (x_0,z_0) pontból megvalósítható tervet

$$\pi_t^g(z^t) = g(\pi_{t-1}^g(z^{t-1}), z_t)$$

definíció szerint. Láttuk, hogy így egy az op-függvény által generált $\pi^g \in$ $\in \Pi(x_0,z_0)$ tervet kapunk. Mivel π^* optimális út, ezért

$$u(\pi^*,(x_0,z_0)) \geq u(\pi^g,(x_0,z_0)).$$

Az indukciós lemma szerint tehát

$$F(x_0, \pi_0^*, z_0) + \beta \int_Z u(c(\pi^*, z_1), (\pi_0^*, z_1))\mu(z_0, dz_1) \geq$$
$$\geq F(x_0, \pi_0^*, z_0) + \beta \int_Z u(c(\pi^g, z_1), (\pi_0^*, z_1))\mu(z_0, dz_1),$$

amiből $\beta \neq 0$ felhasználásával az következik, hogy

$$\int_Z u(c(\pi^*, z_1), (\pi_0^*, z_1)) - u(c(\pi^g, z_1), (\pi_0^*, z_1))\mu(z_0, dz_1) \geq 0.$$

Másrészről a $c(\pi^g, z_1) \in \Pi(\pi_0^*, z_1)$ folytatás is az op-függvény által generált út, azt pedig már tudjuk, hogy az op-függvény által generált út optimális is, ami azt jelenti, hogy

$$u(c(\pi^*, z_1), (\pi_0^*, z_1)) - u(c(\pi^g, z_1), (\pi_0^*, z_1)) \leq 0.$$

Egy nem pozitív függvény integrálja persze csak úgy lehet nem negatív, ha a függvény majdnem mindenütt nulla (2.2.21), ami azt jelenti, hogy

$$w(\pi_0^*, z_1) = u(c(\pi^g, z_1), (\pi_0^*, z_1)) = u(c(\pi^*, z_1), (\pi_0^*, z_1)).$$

Ezt kellett belátni. □

8.5.17. állítás (SP értékfüggvényére is teljesül az indukciós lemma)**.** *Tegyük fel, hogy az op-leképezésnek létezik mérhető szelekciója. Ekkor egy (x_0, z_0) pontból induló $\pi^* \in \Pi(x_0, z_0)$ optimális terv esetén*

$$w(x_0, z_0) = F(x_0, \pi_0^*, z_0) + \beta \int_Z w(\pi_0^*, z_1)\mu(z_0, dz_1).$$

Bizonyítás. Az indukciós lemma és az előző állítás szerint:

$$\begin{aligned} w(x_0, z_0) &= u(\pi^*, (x_0, z_0)) = \\ &= F(x_0, \pi_0^*, z_0) + \beta \int_Z u(c(\pi^*, z_1), (\pi_0^*, z_1))\mu(z_0, dz_1) = \\ &= F(x_0, \pi_0^*, z_0) + \beta \int_Z w(\pi_0^*, z_1)\mu(z_0, dz_1). \end{aligned}$$ □

8.5.18. következmény. *Tegyük fel, hogy pozitív diszkonttényező mellett az op-függvénynek van mérhető szelekciója, valamint minden $\pi \in \Pi(x_0, z_0)$ mellett*

$$\int_{Z^t} v(\pi_{t-1}(z^{t-1}), z_t)\mu^t(z_0, dz^t) \to 0,$$

ahol $v : X \times Z \to X$ a Bellman-egyenlet megoldása. Láttuk, hogy ekkor a megoldás egyértelmű, tetszőleges az op-függvény által generált út egyben optimális út is, továbbá a Bellman-egyenlet megoldása egybeesik a SP feladat érték függvényével. Legyen most π^* a SP feladat egy optimális megoldása. Ekkor

$$w(x_0, z_0) = u(\pi^*, (x_0, z_0)) = v(x_0, z_0).$$

Ugyanezt megismételve az $x_1 = \pi_0^*$ állapotra, kapjuk, hogy

$$w(x_1, z_1) = u(c(\pi^*, z_1), (x_1, z_1)) = v(x_1, z_1) \quad m.m\ \mu(z_0, \cdot),$$

tehát ez előző tétel konklúziója a w értékfüggvény helyett a Bellman-egyenlet v megoldására is teljesül, ami azt jelenti, hogy

$$\pi_0^* \in G(x_0, z_0),$$

ahol G az op-policy leképezés. A fenti gondolatot iterálva kapjuk, hogy minden $t \geq 1$ mellett

$$\pi_t^*(z^t) \in G(\pi_{t-1}^*(z^{t-1}), z_t),$$

ami azt jelenti, hogy a π^* optimális pálya egyben az op-függvény által generált pálya is. A fenti feltételek fennállása esetén tehát az optimális pálya fogalma és az op-függvény által generált pálya fogalma egybeesik.

8.6. A sztochasztikus Bellman-egyenlet megoldhatósága

Az előzőekben azt láttuk, hogy ha egy függvény bizonyos feltételek mellett kielégíti a Bellman-egyenletet, akkor az egyben a szuprémum feladat megoldása is. Sőt, azt is láttuk, hogy ilyenkor vannak olyan tervek, amelyek mentén a feladat fel is veszi maximális értékét.

A most olyan feltételeket keresünk, amelyek teljesülése már a Bellman-egyenlet teljesülését vonják maguk után.

Először is, gondoljunk vissza a determinisztikus esetre. Ott a Bellman-egyenlet

$$v(x) = \sup_{y \in \Gamma(x)} \{F(x, y) + \beta v(y)\}$$

formában volt felírva, ahol $\Gamma(x)$ jelentette az x helyzetből elérhető állapotok nem üres, kompakt halmazát, $F(x, y)$ pedig az egylépéses hozadékfüggvény, ami azt írja le, hogy az x állapotból az y állapotba kerülés milyen hozadékot jelent. Az F-ről feltettük, hogy korlátos, folytonos függvény. Bevezettük a folytonos, korlátos függvényeken értelmezett

$$(T(f))(x) = \sup_{y \in \Gamma(x)} \{F(x, y) + \beta f(y)\}$$

operátort, és ennek fixpontját kerestük. Olyan feltételeket tettünk, amelyek segítségével a Banach-fixponttétel alkalmazhatóvá vált, így a fenti T operátor fixpontja nyilvánvalóan megoldása a Bellman-egyenletnek. Sőt azt is megmutattuk, hogy további feltevések mellett, garantálható a fenti egyértelműen létező fixpont szigorúan monoton növekedése, és szigorúan konkavitása is.

A szakasz üzenete nagyon egyszerű: Sztochasztikus esetre szó szerint átvihetők a fenti eredmények. Új feltétel, lényegében csak egyetlen egy van, ami a sztochasztikus sokkokra vonatkozik. Az összes többi feltétel a determinisztikus eset értelemszerű átírásával adódik.

Először is nézzünk rá a korábban bevezetett sztochasztikus függvény egyenletre. Adott F, Γ mellett olyan v függvényt keresünk, amelyre

$$v(x,z) = \sup_{y\in\Gamma(x,z)} \left\{ F(x,y,z) + \int_Z \beta v\,(y,z')\,\mu(z,dz') \right\},$$

ahol $\Gamma(x,z) \subseteq X \times Z$ az x állapotból a z sokk bekövetkezte esetén elérhető állapotok halmaza, és $F(x,y,z)$ az ebből a helyzetből az y állapotba lépés hozadéka. Persze azonnal szembetűnő, hogy a determinisztikus eset megoldásakor alkalmazott trükkre most is lehetőség nyílik. Csak azt kell tenni, hogy az adott f, $X \times Z$-n értelmezett folytonos, korlátos függvényhez rendeljük hozzá a

$$(T(f))(x,z) = \sup_{y\in\Gamma(x,z)} \left\{ F(x,y,z) + \int_Z \beta f\,(y,z')\,\mu(z,dz') \right\}$$

függvényt, és az így kapott operátornak keressük most is fixpontját. Persze a fenti integrálnak mint y függvényének folytonosságát valahogy biztosítanunk kell, hiszen most is a Berge-tétel az egyetlen reményünk arra, hogy a $T(f)$ függvény is folytonos legyen, ami pedig elengedhetetlen a Banach-fixponttétel alkalmazhatóságához. A rögtön bevezetendő sztochasztikus sokk-feltétel éppen ezt a célt szolgálja. Ha ezzel megbirkózunk, lényegében már készen is vagyunk, mert akkor a determinisztikus esethez hasonlóan, teljesen analóg módon kapjuk az elegendő feltételeket a Bellman-egyenlet megoldhatóságára.

8.6.1. Sokk feltétel

Emlékezzünk arra, hogy a sokkok struktúráját egy $(Z, \mathcal{Z})$ mérhető tér és egy $\mu : Z \times \mathcal{Z} \to [0,1]$ átmenetfüggvény határozza meg. Az alábbi feltétellel élünk a sokkokat leíró struktúrára:

8.6.1. definíció (Sokk feltétel)**.** A lenti két eset közül az egyik fennáll:
a) Z megszámlálható és $\mathcal{Z}$ a hatványhalmaz.
b) $Z \subseteq \mathbb{R}^l$ kompakt halmaz, $\mathcal{Z}$ a Z feletti Borel-halmazok halmaza, és a μ átmenetfüggvény Markov-transzformáltja teljesíti a Feller-feltételt.

Emlékezzünk arra, hogy mit jelent a Feller-feltétel! Tekintsük az $f : Z \to \mathbb{R}$ függvény mellett a

$$(Mf)(z) = \int_Z f(z')\mu(z, dz')$$

Markov-transzformáltat. Amennyiben folytonos és korlátos f mellett az Mf is folytonos (az persze nyilvánvaló, hogy korlátos), akkor a μ átmenetfüggvény teljesíti a Feller-feltételt.

Az a) eset fennállása esetén is szeretnénk folytonosságról beszélni, ezért hallgatólagosan feltesszük, hogy a Z megszámlálható halmazon a diszkrét metrika van megadva, ami persze csak annyit jelent, hogy minden függvény folytonos.

A következő jelölés bevezetése voltaképpen csak egy rövidítés a sztochasztikus Bellman-egyenlet felírásához.

8.6.2. definíció. Az $f : X \times Z \to \mathbb{R}_+$ függvény mellett legyen $N(f) : X \times \times Z \to \mathbb{R}_+$ a következőképpen definiálva.

$$(N(f))(y, z) = \int_Z f(y, z')\mu(z, dz').$$

Persze az éppen bevezetett N operátor kapcsolatban áll a Markov-transzformálttal: Ha f_y jelöli a függvény első változóban rögzített szeletét, azaz $f_y = f(y, \cdot)$, akkor

$$Nf(y, z) = Tf_y(z),$$

így a Bellman-egyenlet picit egyszerűbben írható:

$$v(x, z) = \sup_{y \in \Gamma(x,z)} \{F(x, y, z) + \beta Nf(y, z)\}.$$

Bevezetett N operátornak van három egyszerű tulajdonsága, amit érdemes szem előtt tartanunk:

8.6.3.
Az N operátor monotonitása nyilvánvaló, az absztrakt mérték szerinti integrál monotonitási tulajdonsága miatt. Ha ugyanis $f \leq g$, akkor

$$Nf(y, z) = \int_Z f(y, z')\ \mu(z, dz') \leq \int_Z g(y, z')\ \mu(z, dz') = Ng(y, z),$$

ami éppen azt jelenti, hogy $Nf \leq Ng$, azaz N egy monoton operátor.

Hasonlóan, az integrál linearitásából következik, hogy az első változójában konkáv f mellett Nf is éppen ilyen marad.

Végül azt vegyük észre, hogy a $\mu(z, \cdot)$ egy valószínűségi mérték, ezért $N(f + \alpha) = N(f) + \alpha$ is fennáll.

8.6.4. tétel. *Tegyük fel, hogy a* $(Z, \mathcal{Z})$ *mérhető térre és a* μ *átmenetfüggvényre teljesül a sokk-feltétel. Ekkor minden* $f : X \times Z \to \mathbb{R}$ *folytonos és korlátos függvény mellett az* Nf *függvény is folytonos és korlátos.*

Bizonyítás. Tegyük fel, hogy $(y_n, z_n) \to (y, z)$. A sokk-feltétel mindkét esetében a bizonyítás kulcsa a következő egyenlőtlenség teljesülése:

$$|Nf(y,z) - Nf(y_n, z_n)| \le |Nf(y,z) - Nf(y, z_n)| + |Nf(y, z_n) - Nf(y_n, z_n)|.$$

a) Megszámlálható Z esetében, $z_n \to z$ csak úgy lehet, hogy $z_n = z$ egy bizonyos N után, ezért a jobb oldali első tag zérus. A második tagot kell tehát csak becsülnünk, amikor $z = z_n$. No de:

$$|Nf(y,z) - Nf(y_n, z)| = \left| \int_Z f(y,z') - f(y_n, z')\mu(z, dz') \right| \le$$
$$\le \int_Z h_n(z')\mu(z, dz'),$$

ahol $h_n(z') = |f(y,z') - f(y_n,z')|$. Viszont $h_n \to 0$ pontonként az f folytonossága szerint, és persze $h_n \le 2\|f\|$, mivel f korlátos is, így persze figyelembe véve, hogy $\mu(z, Z) = 1$, adódik, hogy a h_n sorozatnak van véges integrálú majoránsa. Ergo a Lebesgue-konvergenciatétel szerint a jobb oldali második tag is tart zérushoz.

b) Amikor a Feller-feltétel jelenti a sokk-feltétel teljesülését, akkor az egyenlőtlenség jobb oldalának első tagja zérushoz tart, hiszen

$$Nf(y,z) - Nf(y, z_n) = Tf_y(z) - Tf_y(z_n),$$

és a Feller-feltétel éppen azt jelenti, hogy Tf_y folytonos. A második becsléséhez legyen $D = \{y_n : n \in \mathbb{N}\} \cup \{y\}$. Az $y_n \to y$ konvergencia miatt a D kompakt halmaz. Szorítsuk meg f-et a $D \times Z$ kompakt halmazra. Az f folytonossága miatt f egyenletesen is folytonos a $D \times Z$ kompakt halmazon. Ez azt jelenti, hogy minden $\varepsilon > 0$-hoz választható olyan N, hogy minden $n > N$ mellett $|f(y,z') - f(y_n,z')| < \varepsilon$ egyenlőtlenség teljesül a z' megválasztásától függetlenül. Ezért

$$|Nf(y,z_n) - Nf(y_n,z_n)| = \left| \int_Z f(y,z') - f(y_n,z')\mu(z_n, dz') \right| \le \varepsilon\mu(z_n, Z) = \varepsilon.$$

Azt kaptuk tehát, hogy az Nf függvény valóban folytonos a sokk-feltétel mindkét speciális esetében. Ezt kellett belátni. □

8.6.2. Banach fixpont tételének alkalmazásának feltételei

Definiáljuk azt az operátort, aminek fixpontja szolgáltathatja a sztochasztikus Bellman-egyenlet megoldását. Jelölje a továbbiakban $C(S)$ az $S = X \times Z$

halmazon értelmezett folytonos, korlátos függvények halmazát. Tudjuk, hogy ez Banach-tér, ha ellátjuk az $\|f\| = \sup\{|f(x,z)| : (x,z) \in S\}$ normával.

8.6.5. definíció. Adott $f \in C(S)$ mellett legyen

$$(Tf)(x,z) = \sup_{y \in \Gamma(x,z)} \{F(x,y,z) + \beta N f(y,z)\}.$$

Először is megvizsgáljuk a bevezetett T operátor tulajdonságait. Ahhoz, hogy a legegyszerűbb módon a Banach-fixponttétel segítségével biztosítani tudjuk a fixpont egyértelmű létezését, az kell, hogy T kontrakció legyen és a $C(S)$ teret önmagába képezze. Persze a determinisztikus esethez hasonlóan most is többre vágyunk: A Bellman-egyenlet szigorúan monoton növő megoldására, mi több, szigorúan konkáv megoldhatóságára.

Először nézzük a kontrakció problémáját. Most is a Blackwell-lemmát alkalmazzuk, ezért kell T monotonitása és a $T(f+\alpha) \le Tf + \alpha\beta$ tulajdonság teljesülése. Ez implikálja, hogy T kontrakció a β konstanssal, tehát ha olyan feltételeket találunk amelyek biztosítják, hogy $T(f) \in C(S)$ is legyen, akkor máris tudjuk, hogy a Bellman-egyenletnek létezik egyetlen folytonos, korlátos megoldása.

Ezután rátérünk majd annak vizsgálatára, hogy hogyan garantáljuk a fent megtalált egyértelmű megoldás szép tulajdonságait. A szigorú monotonitásra olyan feltételt kell találnunk, amely teljesülése esetén T minden monoton növekedő függvényhez szigorúan monoton növekedőt rendel, a megoldás konkavitását pedig úgy biztosíthatjuk, hogy olyan feltételeket találunk, amelyek garantálják, hogy T minden konkáv függvényhez szigorúan konkáv függvényt rendel.

Ugyanis, ha T mondjuk ilyen, akkor T operátort leszorítjuk a $C(S)$ azon alterére, amelyek a konkáv függvényeket tartalmazzák. Látható, hogy konkáv függvények pontonkénti határértéke, pláne egyenletes határértéke is konkáv marad, ami azt jelenti, hogy a szóban forgó altér $C(S)$-nek zárt részhalmaza, ergo ez az altér is Banach-tér. A Banach-fixponttételt itt alkalmazva kapunk egy fixpontot, ami persze megegyezik az egész $C(S)$-en talált fixponttal, hiszen a tétel ott is egyértelműséget garantált. Világos, hogy ez a fixpont konkáv, és T-t még egyszer alkalmazva rá, látjuk azt is hogy szigorúan konkáv. A szigorú monotonitást ezzel analóg módon garantáljuk.

A most megértett programhoz szükséges operátortulajdonságok következnek.

8.6.6. állítás. *Legyen $X \subseteq \mathbb{R}^k$ egy Borel-halmaz; tegyük fel a sztochasztikus sokkokat definiáló $(Z, \mathcal{Z})$ mérhető térről, és a $\mu : Z \times \mathcal{Z} \to [0,1]$ átmenetfüggvényről, hogy teljesítik a sokk-feltételt; az $F : \operatorname{graph}\Gamma \to \mathbb{R}_+$ hozamfüggvény legyen egy korlátos folytonos függvény; a $\Gamma : S \multimap X$ halmaz értékű folytonos függvény (tehát egyszerre alulról és felülről félig folytonos) rendeljen minden*

$s \in S$ pontpárhoz $\Gamma(s)$ nem üres kompakt halmazt; és legyen a diszkonttényezőre $\beta \in (0{,}1)$.

Ekkor T rendelkezik az alábbi tulajdonságokkal:

1. *monoton operátor, azaz ha $f, g \in C(S)$, amelyekre $f \leq g$, akkor $Tf \leq$ $\leq Tg$;*
2. *minden rögzített $\alpha \geq 0$ konstans mellett $T(f + \alpha) = Tf + \alpha\beta$;*
3. *minden $f \in C(S)$ esetén Tf is folytonos, korlátos függvény, továbbá az*

$$(x, z) \mapsto \{y \in \Gamma(x, z) : F(x, y, z) + \beta Mf(y, z) = Tf(x, z)\}$$

   egy nem üres kompakt halmazokat felvevő felülről félig folytonos halmazértékű hozzárendelés.

4. *Most tegyük fel még, hogy*

   (a) *a Γ leképezés monoton az első változójában, azaz*
   $x_1 \leq x_2 \implies \Gamma(x_1, z) \subseteq \Gamma(x_2, z)$;

   (b) *és az F hozadékfüggvény is szigorúan monoton az első változójában, azaz*
   $x_1 \neq x_2, x_1 \leq x_2$ mellett $F(x_1, y, z) < F(x_2, y, z)$ minden $z \in Z$-re és $y \in X$-re.

   Ekkor Tf az első változójában szigorúan monoton növő függvény, azaz

$$x_1 \leq x_2, x_1 \neq x_2 \implies Tf(x_1, z) < T(x_2, z)$$

   a z változótól függetlenül.

5. *Most azt tegyük fel, hogy*

   (a) *a Γ leképezés konkáv az első változójában, azaz*

$$\begin{aligned} x_1, x_2 \in X, \lambda \in (0{,}1) \implies \Gamma(\lambda x_1 + (1 - \lambda)x_2, z) &\supseteq \\ \supseteq \lambda\Gamma(x_1, z) + (1 - \lambda)\Gamma(x_2, z)&; \end{aligned}$$

   (b) *és az F hozadékfüggvény együttesen szigorúan konkáv az első két változójában, azaz $x_1, y_1, x_2, y_2 \in X, (x_1, y_1) \neq (x_2, y_2)\lambda \in (0{,}1)$ mellett*

$$\begin{aligned} &F(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2, z) > \\ &> \lambda F(x_1, y_1, z) + (1 - \lambda)F(x_2, y_2, z) \end{aligned}$$

   minden $z \in Z$ mellett.

Ekkor T az első változójában konkáv függvényhez az első változójában szigorúan konkávot rendel, azaz

$$x_1, x_2 \in X, \lambda \in (0{,}1) \implies f(\lambda x_1 + (1-\lambda)x_2, z) \geq \\ \geq \lambda f(x_1, z) + (1-\lambda) f(x_2, z)$$

implikáció teljesülése maga után vonja az

$$x_1, x_2 \in X; x_1 \neq x_2, \lambda \in (0{,}1) \implies \\ \implies Tf(\lambda x_1 + (1-\lambda)x_2, z) > \lambda Tf(x_1, z) + (1-\lambda) Tf(x_2, z)$$

implikációt z-től függetlenül.

Bizonyítás. Valamennyi állítás igazolása a determinisztikus eset párhuzamos állításának indoklásával analóg:

1) A monotonitás egyszerűen az N operátor monotonitásából látszik: Ha $f \leq g$, akkor

$$Tf(x,z) = \sup_{y \in \Gamma(x,z)} \{F(x,y,z) + \beta N f(y,z)\} \leq \\ \leq \sup_{y \in \Gamma(x,z)} \{F(x,y,z) + \beta N g(y,z)\} = Tg(x,z).$$

2) A kontrakciót biztosító feltétel is könnyű, mert tudjuk, hogy $N(f+\alpha) = N(f) + \alpha$. Emiatt

$$(T(f+\alpha))(x,z) = \\ = \sup_{y \in \Gamma(x,z)} \{F(x,y,z) + \beta (N(f+\alpha))(y,z)\} = \\ = \sup_{y \in \Gamma(x,z)} \{F(x,y,z) + \beta (N f(y,z) + \alpha)\} = \\ = \sup_{y \in \Gamma(x,z)} \{F(x,y,z) + \beta N f(y,z)\} + \beta\alpha = (T(f))(x,z) + \beta\alpha.$$

3) $T(f)$ folytonosságához tekintsük rögzített $(x,z) \in S$ mellett az

$$y \mapsto F(x,y,z) + \beta N f(y,z)$$

függvényt. A hozadékfüggvény folytonossága és a sokkokra vonatkozó feltevésünk szerint ez a függvény folytonos. Így a $\Gamma(x,z)$ kompakt halmazon felveszi a legnagyobb értékét, ergo a sup helyett max is írható. A 11.2.3. Berge-tétel szerint a tett feltevések elegendőek ahhoz, hogy az értékfüggvény folytonosan függjön a paraméterektől. Ez éppen azt jelenti, hogy Tf folytonos.

Azt is tudjuk még a Berge-tételből, hogy az argmax halmazértékű leképezés felülről félig folytonos. Tehát

$$(x,z) \mapsto \{y \in \Gamma(x,z) : F(x,y,z) + \beta Nf(y,z) = Tf(x,z)\}$$

tetszőleges $f \in C(S)$ mellett egy nem üres, kompakt értékű, felülről félig folytonos halmazértékű leképezés.

4) Tf szigorúan monotonitásához csak a Γ monotonitását és F szigorúan monotonitását kell használnunk:

$$\begin{aligned} Tf(x_1,z) = &\max_{y\in\Gamma(x_1,z)} \{F(x_1,y,z) + \beta Nf(y,z)\} \leq \\ \leq &\max_{y\in\Gamma(x_2,z)} \{F(x_1,y,z) + \beta Nf(y,z)\} < \\ < &\max_{y\in\Gamma(x_2,z)} \{F(x_2,y,z) + \beta Nf(y,z)\} = Tf(x_2,z). \end{aligned}$$

Itt persze felhasználtuk, hogy a fenti maximumok valóban felvétetnek és arra az y-ra alkalmaztuk az $F(\cdot,y,z)$ függvény szigorúan monotonitását, amelyre $F(x_1,y,z) + \beta Nf(y,z) = \max_{\Gamma(x_2,z)} \{F(x_1,\cdot,z) + \beta Nf(\cdot,z)\}$.

5) A Tf szigorúan konkavitásához:

$$\begin{aligned} Tf(x_1,z) &= F(x_1,y_1,z) + \beta Nf(y_1,z) \quad \text{valamely} \quad y_1 \in \Gamma(x_1,z), \\ Tf(x_2,z) &= F(x_2,y_2,z) + \beta Nf(y_2,z) \quad \text{valamely} \quad y_2 \in \Gamma(x_2,z) \end{aligned}$$

mellett. Így a Γ konkavitása miatt

$$\lambda y_1 + (1-\lambda)y_2 \in \lambda\Gamma(x_1,z) + (1-\lambda)\Gamma(x_2,z) \subseteq \Gamma(\lambda x_1 + (1-\lambda)x_2, z).$$

Tehát egy halmaz szuprémuma nem kisebb, mint bármelyik elem, ezért:

$$\begin{aligned} &Tf(\lambda x_1 + (1-\lambda)x_2, z) = \\ &= \max\{F(\lambda x_1+(1-\lambda)x_2, y, z)+\beta Nf(y,z) : y \in \Gamma(\lambda x_1 + (1-\lambda)x_2, z)\} \geq \\ &\geq F(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2, z) + \beta Nf(\lambda y_1 + (1-\lambda)y_2, z) \geq \\ &\geq F(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2, z) + \\ &\qquad\qquad + \lambda\beta Nf(y_1,z) + (1-\lambda)\beta Nf(y_2,z) > \\ &> \lambda(F(x_1,y_1,z) + \beta Nf(y_1,z)) + (1-\lambda)(F(x_2,y_2,z) + \beta Nf(y_2,z)) = \\ &= \lambda Tf(x_1,z) + (1-\lambda)Tf(x_2,z). \end{aligned}$$

A harmadik egyenlőtlenség Nf konkavitása, míg a szigorú egyenlőtlenség F szigorú konkavitása miatt teljesül.

Ezt kellett belátni. $\square$

8.6.3. A Bellman-egyenlet egzisztencia és unicitás tétele

8.6.7.
Most emlékezzünk vissza a Blackwell-lemmára: Ha T egy a korlátos függvények halmazán értelmezett monoton operátor, amelyre teljesül a

$$T(f+\alpha) \leq T(f) + \beta\alpha$$

egyenlőtlenség, akkor T egy kontrakció a β konstanssal.

Ugyanis: $f - g \leq \|f - g\|$ pontonként, ezért $f \leq g + \|f - g\|$ is fennáll pontonként, így a monotonitás és a feltétel miatt $T(f) \leq T(g) + \beta\|f - g\|$. Ebből először a $Tf - Tg \leq \beta\|f - g\|$ egyenlőtlenséget, majd f és g szerepét felcserélve a $|Tf - Tg| \leq \beta\|f - g\|$ egyenlőtlenséget kapjuk pontonként. Ekkor persze $\|Tf - Tg\| \leq \beta\|f - g\|$ is fennáll, ami éppen azt jelenti, hogy T valóban egy β konstansú kontrakció.

Mindent átgondoltunk a Bellman-egyenlet egzisztenciáját és unicitását garantáló feltételekről. Összefoglalásként pontosan megfogalmazzuk a fenti T operátor fixpontjával kapcsolatos tételeket:

8.6.8. tétel (A Bellman-egyenlet egzisztencia és unicitás tétele)**.** *Legyen $X \subseteq \subseteq \mathbb{R}^k$ egy Borel-halmaz; tegyük fel a sztochasztikus sokkokat definiáló $(Z, \mathcal{Z})$ mérhető térről, és a $\mu : Z \times \mathcal{Z} \to [0,1]$ átmenetfüggvényről, hogy teljesítik a sokk-feltételt; az $F : \operatorname{graph}\Gamma \to \mathbb{R}_+$ hozamfüggvény legyen egy korlátos, folytonos függvény; a $\Gamma : S \multimap X$ halmaz értékű folytonos leképezés rendeljen minden $s \in S$ pontpárhoz $\Gamma(s) \subseteq X$ nem üres kompakt halmazt; végül legyen a diszkonttényezőre $\beta \in (0,1)$.*

Ekkor a Bellman-egyenletnek létezik egyetlen folytonos, korlátos megoldása, azaz létezik egyetlen $v \in C(S)$ korlátos függvény, amelyre minden $(x, z) \in \in S$ esetén teljesül az

$$v(x,z) = \sup_{y \in \Gamma(x,z)} \left\{ F(x,y,z) + \beta \int_Z v(y,z')\mu(z,dz') \right\}$$

egyenlőség.

A fenti sup *helyett* max *is írható, majd bevezetve a*

$$G(x,z) = \left\{ y \in \Gamma(x,z) : v(x,z) = F(x,y,z) + \beta \int_Z v(y,z')\mu(z,dz') \right\}$$

optimal policy halmazértékű leképezést a kapott $G : S \multimap X$ hozzárendelés egy nem üres, kompakt értékű, és felülről félig folytonos függvény.

Bizonyítás. Tekintsük a fent definiált $T : C(S) \to C(S)$ operátort. Láttuk, hogy ez egy kontrakció a diszkonttényezővel mint konstanssal. Alkalmazható tehát a Banach-fixponttétel. Ez azt jelenti, hogy létezik egyetlen egy olyan $v \in$ $\in C(S)$, amire $Tv = v$, ami pont azzal ekvivalens, hogy v a Bellman-egyenlet egyetlen megoldása.

A felülről félig folytonosság a 11.2.3. Berge-tétellel adódik. Tetszőleges folytonos függvény mellett az argmax hozzárendelés, nem üres, kompakt értékű, hiszen azon pontok halmaza, ahol egy folytonos függvény felveszi a szuprémumát, egy zárt halmaz, és persze kompakt halmaz zárt részhalmaza kompakt. Teljesülnek tehát a Berge-tétel feltételei, ezért v felülről félig folytonos is. □

8.6.9. tétel (A Bellman-egyenlet egyetlen megoldása szigorúan monoton növő)**.** *Most tegyük fel a Bellman-egyenlet egzisztenciája és unicitásának garantáló tétel feltételei mellé, hogy a Γ függvény monoton növő az első változójában, és az F hozadékfüggvény szigorúan monoton növő az első változójában.*

Ekkor a Bellman-egyenlet egyetlen megoldása a fentiek mellett még szigorúan monoton növő is az első változójában.

Bizonyítás. Jelölje $C'(S)$ az S-en értelmezett folytonos, korlátos, monoton növő függvények alterét. Látható, hogy $C'(S)$ egy zárt részhalmaza a $C(S)$ Banach-térnek, ezért maga is Banach-tér. Az F-re és Γ-ra tett kiegészítő feltételek szerint $T : C'(S) \to C'(S)$ operátor. Itt is alkalmazhatjuk tehát a Banach-fixponttételt. Az emiatt létező v fixpont, nyilván a Bellman-egyenlet egyetlen megoldása, és $Tv = v$ miatt szigorúan monoton növekedő is. □

Most nézzük a szigorúan konkáv megoldás létezését garantáló állítást:

8.6.10. tétel (A Bellman-egyenlet egyetlen megoldása szigorúan konkáv)**.** *Tegyük fel a Bellman-egyenlet egzisztenciája és unicitásának garantáló tétel feltételei mellé, hogy a Γ függvény konkáv az első változójában, és az F hozadékfüggvény együttesen szigorúan konkáv az első két változójában.*

Ekkor a Bellman-egyenlet egyetlen megoldása a fentiek mellett még szigorúan konkáv is az első változójában.

Sőt ebben az esetben a $G(x, z)$ optimális policy leképezés minden értéke egyelemű halmaz, és a $G : S \to S$ függvény folytonos.

Bizonyítás. Jelölje $C'(S)$ az S-en értelmezett folytonos, korlátos, konkáv függvények alterét. Látható, hogy $C'(S)$ egy zárt részhalmaza a $C(S)$ Banach-térnek, ezért maga is Banach-tér. Az F-re és Γ-ra tett kiegészítő feltételek szerint $T : C'(S) \to C'(S)$ operátor. Itt is alkalmazhatjuk tehát a Banach-fixponttételt. Az emiatt létező v fixpont, nyilván a Bellman-egyenlet egyetlen megoldása, és $Tv = v$ miatt szigorúan konkáv is.

Mivel egy szigorúan konkáv és egy konkáv függvény összege szigorúan konkáv, ezért $G(x, z)$ egy szigorúan konkáv függvény maximumhelyeit tartalmazza. Persze szigorúan konkáv függvény két különböző pontban nem tudja felvenni a maximális értékét, így $G(x, z)$ egyelemű. No de egy halmazértékű függvény felülről félig folytonossága éppen folytonosságot jelent, ha a képhalmazok egyeleműek. □

8.6.4. Összegzés

Azt gondoltuk meg tehát, hogy a determinisztikus esetre vonatkozó egzisztencia és unicitás tétel szinte szó szerint átírható a sztochasztikus esetre is. Az egyetlen új feltétel a sztochasztikus sokkok struktúrájára vonatkozott, ami persze a determinisztikus esetben nincs is.

Külön érdemes felfigyelni a Z, tehát a sokkok halmaza megszámlálható feltételre, hiszen ekkor a sztochasztikus esetben sincs új feltétel. Szép látni, hogy abban az esetben mikor Z egyelemű, éppen a determinisztikus Bellman-egyenlet egzisztencia- és unicitási tételeit kapjuk speciális esetként. Elmondhatjuk tehát, hogy nem igazán van értelme sztochasztikus és determinisztikus esetről beszélni, hiszen a determinisztikus dinamikus programozás nem más, mint a sztochasztikus dinamikus programozás az egyelemű sokk-tér speciális esetében.

9. fejezet

Sztochasztikus mátrixok

A fejezetben a sztochasztikus sokkok struktúráját megadó átmenetfüggvényt vizsgáljuk abban a speciális esetben, mikor a sokkok $Z = \{z_1, \cdots, z_l\}$ halmaza véges.

Mivel egy véges halmazon megadott valószínűségi mérték nem más, mint egy $p = (p_1, \cdots, p_l)$ *valószínűségeloszlás* – $p_i \geq 0$ és $\sum_{k=1}^{l} p_k = 1$ –, ezért egy $Q : Z \times \mathcal{P}(Z) \to \mathbb{R}$ átmenetfüggvény megadása azonos l darab valószínűségeloszlás megadásával. Az i-edik eloszlás, a $Q(z_i, \cdot)$ mérték, ezért az átmenetfüggvény intuíciója szerint a $Q(z_i, \{z_j\})$ annak valószínűsége, hogy a rendszert a z_i sokk bekövetkezte után periódusban a z_j sokk éri.

9.0.11.
Ha $f : Z \to \mathbb{R}$ véges halmazon értelmezett függvény és μ egy a véges halmaz hatványhalmazán értelmezett mérték, akkor

$$\int_Z f \, d\mu = \sum_{k=1}^{l} \int_{\{z_k\}} f \, d\mu = \sum_{k=1}^{l} f(z_k) \, \mu(\{z_k\}) .$$

Speciálisan a Q^n átmenetfüggvény-szorzatra felírva az 8.4.2. pont alatti egyenlőséget

$$Q^{n+1}(z_i, \{z_j\}) = \int_Z Q^n(z, \{z_j\}) \, Q(z_i, dz) = \sum_{k=1}^{l} Q^n(z_k, \{z_j\}) \, Q(z_i, \{z_k\}) . \tag{9.1}$$

Vezessük be most a

$$\Pi_{i,j} = Q(z_i, \{z_j\})$$

jelölést. Világos, hogy Π egy $l \times l$-es *sztochasztikus mátrix*, azaz olyan nem negatív mátrix, melynek minden sorösszege éppen 1. A fenti (9.1) $n = 1$

speciális esetében, a Q^2 szorzathoz rendelt mátrix éppen a $\Pi^2 = \Pi \cdot \Pi$ szorzatmátrix, hiszen a jobb oldal

$$\sum_{k=1}^{l} \Pi_{i,k}\Pi_{k,j},$$

ami a mátrix szorzás definíciója szerint éppen a Π^2 mátrix i-edik sorának j-edik eleme. Ugyanígy kapjuk $n = 2$ alkalmazásával, hogy a Q^3 sztochasztikus szorzathoz rendelt mátrix éppen Π^3, és teljesen hasonlóan, hogy a Q^n szorzathoz rendelt mátrix éppen a Π^n mátrixhatvány.

A Chapman–Kolmogorov-azonosság (8.4.14) tehát a

$$\Pi^{n+m} = \Pi^n \cdot \Pi^m$$

egyenlőséget jelenti, ami itt egyszerűen a mátrixszorzás asszociativitásának következménye.

Meggondoltuk a következő állítást.

9.0.12. állítás. *A sztochasztikus mátrixok és a véges halmaz feletti sztochasztikus magok kölcsönösen egyértelműen meghatározzák egymást. E kölcsönösen egyértelmű megfeleltetésben a sztochasztikus mag n-edik hatványához rendelt mátrix azonos az ezen maghoz rendelt mátrix n-edik hatványával.*

9.0.13.

A Π^n sztochasztikus mátrix i-edik sorának j-edik eleme tehát annak a valószínűsége, hogy a rendszer a z_i sokk-állapotból n periódus leteltének hatására a z_j sokk-helyzetbe kerül.

A továbbiak előtt nézzünk egy konkrét, ám egyszerű példát. Tegyük fel, hogy $Z = \{z_1, z_2\}$, tehát csak két fajta sokkot ismerünk. Mindkét sokkra igaz, hogy az ismétlődés valószínűsége $3/4$, továbbá az egyik sokk után a másik sokk bekövetkeztének valószínűsége mindkét sokk mellett $1/4$. Megkérdezhető, hogy ha kezdetben megfigyeltük mondjuk a $\{z_1\}$ sokk bekövetkeztét, akkor sok-sok sokk után mi a valószínűsége a z_1 vagy a z_2 sokk bekövetkeztének. Hosszú távon nézve melyiknek nagyobb a gyakorisága?

E sztochasztikus magot leíró mátrix a

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}.$$

A P^n első sora tartalmazza azt a valószínűségeloszlást, amit z_1 sokkból indulva n lépés után valószínűsítünk, és a második sor az a valószínűségeloszlás, amit z_2 sokkból indulva kapunk. A P^n mátrixot kell tehát nagy n értékekre közelítenünk.

Jelölje $a_1 = \frac{3}{4}$ és $b_1 = \frac{1}{4}$, továbbá a_n és b_n a P^n mátrix első sorának első és második eleme. Ekkor a mátrix szorzás definíciója szerint,

$$\begin{pmatrix} a_{n+1} & b_{n+1} \\ b_{n+1} & a_{n+1} \end{pmatrix} = \begin{pmatrix} a_n & b_n \\ b_n & a_n \end{pmatrix} \cdot \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}.$$

Áttérve az elemek definíciójára azt kapjuk, hogy

$$a_{n+1} = \frac{1}{4}\left(3a_n + b_n\right),$$
$$b_{n+1} = \frac{1}{4}\left(3b_n + a_n\right).$$

Oldjuk meg a rekurziót: Világos, hogy a_n monoton fogyó és b_n monoton növő sorozat, melyre $a_n > \frac{1}{2} > b_n$ teljesül minden $n \in \mathbb{N}$ mellett. Monoton, korlátos sorozatokként mindkét sorozat konvergens. Jelölje $\alpha = \lim_{n\to\infty} a_n$ és $\beta = \lim_{n\to\infty} b_n$. A rekurziót alkalmazva azt kapjuk, hogy

$$\alpha = \frac{1}{4}\left(3\alpha + \beta\right),$$
$$\beta = \frac{1}{4}\left(3\beta + \alpha\right).$$

Figyelembe véve, hogy P^n egy sztochasztikus mátrix, azaz $a_n+b_n = 1$ teljesül minden $n \in \mathbb{N}$ mellett, így kapjuk, hogy $\alpha + \beta = 1$ is fennáll. Ez viszont csak úgy lehetséges, ha $\alpha = \beta = \frac{1}{2}$.

Azt kaptuk tehát, hogy a P mátrix hatványai pontonként konvergálnak, és

$$P^n \to \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

Ez azt jelenti, hogy bármelyik kezdeti sokkból indulunk is ki, nagyszámú sokk bekövetkezte után körülbelül azonos valószínűséggel számíthatunk mindkét sokk előfordulására.

Még vegyük észre azt is, hogy a határmátrix minden p sorára a $pP = p$ egyenlőség teljesül, azaz minden sor egyben invariáns eloszlás is. Látni fogjuk, hogy pontosan mi a jelenség általános oka.

9.0.14.

A dinamikus programozási feladat értelmezése szerint a Z alaphalmaz a sokkok tere, tehát Z elemeire mint a feladatban rejlő bizonytalanság okára tekintünk, ahogyan ezt eddig is tettük. E bizonytalanság leírására szolgál a Z-én értelmezett sztochasztikus mag. A szóhasználat egyszerűsítése érdekében a továbbiakban nem fogjuk hangsúlyozni, hogy Z elemei voltaképpen sokkok, hanem Z elemeit a rendszer sokk-állapotainak, vagy még rövidebben állapotainak mondjuk. Ha adott egy $p = (p_1, p_2, \dots, p_l)$, kezdeti valószínűségeloszlás, az azt jelenti, hogy rendre $p_1, p_2, \dots, p_l$ annak valószínűsége, hogy rendszerünk a $z_1, z_2, \dots, z_l$ helyzetben van, azaz a rendszert $z_1, z_2, \dots, z_l$ sokkok érik. Tegyük fel, hogy a Z l-elemű halmazon adott egy a sokkokra jellemező sztochasztikus magot leíró $\Pi \in \mathbb{R}^{l\times l}$-es sztochasztikus mátrix. Ekkor $p\Pi$ valószínűségeloszlás interpretációja szerint a $p\Pi$ vektor k-adik koordinátája annak valószínűsége, hogy a rendszer egy periódus elteltével után a z_k sokknak megfelelő állapotba érkezik. Hasonlóan $p\Pi\,(\Pi) = p\Pi^2$ a két egymás utáni periódus elteltével kapott valószínűségeloszlás, és $p\Pi^n$ az n darab időszak utáni eloszlás.

9.1. Normálalak

Tegyük fel, hogy adott egy sztochasztikus magot definiáló Π mátrix, a $Z = \{z_1, \dots, z_l\}$ halmaz felett. Világos, hogy a fenti felsorolásban az elemek sorrendje is fontos, hiszen a mátrix reprezentáció megadásában a sorrend is szerepel. Ha viszont ebben a mátrixban felcseréljük a i-edik oszlopot a j-edik oszloppal, majd az i-edik sort a j-edik sorral, akkor ugyanennek a sztochasztikus magnak a mátrix reprezentációját kapjuk, azzal az egyetlen különbséggel, hogy a Z halmaz elemeinek sorrendjét megváltoztattuk: a $z_1, \dots, z_i, \dots, z_j, \dots z_l$ sorrendről áttértünk a $z_1, \dots, z_j, \dots, z_i, \dots z_l$ sorrendre. Ugyanis attól, hogy a „zivatar" nevű jelenséget z_i-vel vagy z_j-vel jelölöm, vagy a „köd" nevű jelenséget a z_j vagy z_i szimbólumok melyikével illetem, még magáról a jelenségről és annak előfordulásairól nem mondtam semmit. Mivel minden permutáció transzpozíciók szorzata, ezért a sztochasztikus magot csak formálisan változtatja meg egy rögzített π permutációnak a sorokra és az oszlopokra gyakorolt hatása. Ez pusztán annyit jelent, hogy az elemek eredeti sorrendjéről áttérünk a $z_{\pi_1}, \dots, z_{\pi_l}$ sorrendre. A sor–oszlop teréken átesett mátrix ugyanannak a jelenségnek a reprezentációja, de már az alaphalmaz π-vel permutált sorrendjére felírva.

A szakaszban azt fogjuk vizsgálni, hogy egy sztochasztikus mátrix hogyan írható fel a legáttekinthetőbb formában, megengedve az oszlopok és sorok átrendezését. Először is egy rövidítést vezetünk be a valószínűségeloszlások halmazára.

9.1.1. definíció. Jelölje

$$\Delta_l = \left\{ p = (p_1, \dots, p_l) \in \mathbb{R}^l : p_k \geq 0, \sum_{k=1}^{l} p_k = 1 \right\}$$

A következő állítás voltaképpen magától értetődő, és csak a terminológiát szeretném vele rögzíteni.

9.1.2. állítás. *A $p \in \mathbb{R}^l_+$ nem negatív vektorra $p \in \Delta_l$ akkor és csak akkor, ha $\langle p; u\rangle = 1$, ahol $u = (1, \dots, 1) \in \mathbb{R}^l$ az úgynevezett* összegző vektor. *A* *valószínűségeloszlások $\Delta_l \subseteq \mathbb{R}^l$ halmaza egy konvex, kompakt halmaz. Ha $p \in$ $\in \Delta_l$ egy valószínűségeloszlás és P egy sztochasztikus mátrix, akkor $pP \in \Delta_l$ is fennáll. A $P \in \mathbb{R}_+^{l \times l}$ nem negatív elemű mátrix pontosan akkor sztochasztikus mátrix, ha $Pu = u$. Ha P és Q sztochasztikus mátrixok, akkor azok $P \cdot Q$ szorzata is az.*

Bizonyítás. Tán csak az utolsó állítást érdemes indokolni: Ha u az összegző vektor, akkor $(PQ)u = P(Qu) = Pu = u$ szerint a szorzatmátrix is egy sztochasztikus mátrix. □

9.1.3. definíció (invariáns eloszlás)**.** Legyen $p \in \Delta_l$ egy valószínűségeloszlás és P egy sztochasztikus mátrix. Azt mondjuk, hogy p *invariáns eloszlása* P-nek, ha a

$$pP = p$$

egyenlőség teljesül.

Az invariáns eloszlás interpretációja könnyen érthető: Ha egy periódus elteltével nem változik az egyes sokk-állapotok bekövetkeztének valószínűsége, akkor az ilyen eloszlást a sokkokat leíró sztochasztikus mátrixra nézve invariánsnak mondjuk.

9.1.4.
A *Brouwer-fixpont-tétel* szerint egy X véges dimenziós normált tér $K \subseteq X$ kompakt, konvex halmazán, ha adott egy $f : K \to K$ folytonos leképezés, akkor annak van $f(x) = x$ fixpontja. Mivel $\Delta_l \subseteq \mathbb{R}^l$ egy kompakt, konvex halmaz, és a

$$p \mapsto pP$$

függvény lineárisan, ergo folytonosan, képezi a Δ_l halmazt saját magára, ezért a fenti fixponttétel alkalmazható. Ez azt jelenti, hogy minden sztochasztikus mátrixnak van legalább egy invariáns eloszlása. Később látni fogjuk, hogy ez az eredmény, a viszonylag mély eszköznek számító, Brouwer-fixpont-tétel nélkül is adódik.

9.1.5. definíció. Adott P sztochasztikus mátrix mellett azt mondjuk, hogy

z_j **következménye** z_i**-nek,** ha létezik $N \in \mathbb{N}$ egész szám, melyre $P_{i,j}^N > 0$ teljesül;

z_i **tranziens,** ha van olyan különböző következménye, melynek ő maga nem következménye. Formálisan ha létezik $j \neq i$, és $N \in \mathbb{N}$, hogy $P_{i,j}^N > 0$, ám $P_{j,i}^n = 0$ minden $n \in \mathbb{N}$ mellett;

z_i **visszatérő,** ha nem tranziens, azaz ha minden nála különböző következményének maga is következménye.

Világos például, hogy az iménti egyszerű példában mindkét sokk-állapot tranziens.

9.1.6.
A következményének lenni, egy tranzitív reláció. Hiszen ha z_s következménye z_t-nek és z_r következménye z_s-nek, akkor létezik $N, M \in \mathbb{N}$, amelyekre

$$P_{t,s}^N > 0 \qquad P_{s,r}^M > 0.$$

No de a Chapman–Kolmogorov-azonosság szerint

$$P_{t,r}^{N+M} = \sum_{k=1}^{l} P_{t,k}^N P_{k,r}^M \geq P_{t,s}^N P_{s,r}^M > 0,$$

ami éppen azt jelenti, hogy z_r következménye z_s-nek.

E reláció tranzitivitása játssza a kulcsszerepet a normálalakra vonatkozó tételben. Az első következmény:

9.1.7. állítás. *Minden tranziens elemnek van visszatérő következménye, ezért minden sztochasztikus mátrixnak van legalább egy visszatérő eleme.*

Bizonyítás. Ha nincs tranziens elem, akkor nyilvánvaló az állítás.

Legyen $s_1 \in Z$ tranziens elem. Ez azt jelenti, hogy létezik $s_2 \neq s_1$, $s_2 \in Z$ következménye, melyre s_1 nem következménye s_2-nek. Az

$$(s_1, s_2)$$

tehát különböző elemek úgy sorba téve, hogy pontosan a nagyobb indexű következménye a kisebb indexűnek.

Ha s_2 visszatérő, akkor készen vagyunk. Ha s_2 is tranziens, akkor létezik $s_3 \neq s_2$, $s_3 \in Z$ következménye s_2-nek, de s_2 nem következménye s_3-nak. Világos, hogy $s_3 \neq s_1$, hiszen s_1 nem következménye s_2-nek, de s_3 igen. Vegyük még azt is észre, hogy s_1 sem következménye s_3-nak, hiszen ellenkező esetben a következményének lenni reláció tranzitivitása biztosítaná, hogy s_2 következménye s_3-nek. Az

$$(s_1, s_2, s_3)$$

tehát különböző elemek úgy sorba téve, hogy éppen a nagyobb indexű a következménye a kisebb indexűnek.

Ha s_3 visszatérő, akkor készen is vagyunk. Ha s_3 is tranziens, akkor van olyan $s_4 \neq s_3$, $s_4 \in Z$ következménye s_3-nak, amelyre s_3 nem következménye s_4-nek. Először is látható, hogy $s_4 \notin \{s_1, s_2\}$ is fennáll, hiszen s_1 és s_2 nem következménye s_3-nak, ám s_4 igen. Másodszor pedig az $\{s_1, s_2\}$ halmaz egyike sem következménye s_4-nek, hiszen ellenkező esetben a tranzitivitás szerint s_3-nak lenne nála kisebb indexű következménye. Az

$$(s_1, s_2, s_3, s_4)$$

tehát különböző elemek úgy sorba téve, hogy éppen a nagyobb indexű a következménye a kisebb indexűnek.

Az eljárás folytatásával tehát vagy találunk egy visszatérő elemet, vagy ha $s_1, \dots, s_{l-1}$ mindegyike tranziens, akkor utoljára alkalmazva a fenti eljárást, végül sorba tesszük a Z halmaz valamennyi elemét

$$(s_1, \dots, s_{l-1}, s_l)$$

módon úgy, hogy pontosan a nagyobb indexű a következménye a kisebb indexűnek. Persze így a sorban az utolsó elem olyan, hogy nincs nálánál különböző következménye sem, ergo visszatérő. □

9.1.8. definíció (következmény-zárt halmaz)**.** Legyen adott a Q sztochasztikus mag a Z halmazon. Azt mondjuk, hogy az $E \subseteq Z$ halmaz *következmény-zárt*, ha az alábbi két ekvivalens feltevés egyike fennáll.

1. Minden $z \in E$ mellett $Q(z, E) = 1$.
2. Minden E-beli pont minden következménye is E-beli.

Legyen ugyanis $P = [\pi_{i,j}]$ a maghoz tartozó sztochasztikus mátrix. Ha $z_j \notin E$ és $z_i \in E$, akkor

$$\pi^2_{i,j} = \sum_{k=1}^{l} \pi_{i,k}\pi_{k,j}.$$

Ebben az összegben ha k olyan, hogy $z_k \in E$, akkor $Q(z_k, E) = 1$ miatt $\pi_{k,j} = 0$. Ha viszont egy k-ra $z_k \notin E$, akkor $Q(z_i, E) = 1$ szerint $\pi_{i,k} = 0$. Így látjuk, hogy $\pi^2_{i,j} = 0$. A fenti gondolat ismétlésével kapjuk, hogy minden $N \in \mathbb{N}$ esetén $\pi^N_{i,j} = 0$, ami éppen azt jelenti, hogy z_i-nek nincs az E halmazon kívüli következménye.

Megfordítva, ha $z_i \in E$ és z_i minden következménye E-beli, akkor speciálisan minden egy lépéses következmény is az E halmazba esik, azaz $\pi_{i,j} = 0$ minden $z_j \notin E$ mellett. Így

$$Q(z_i, E) = \sum_{\substack{k=1 \\ z_k \in E}}^{l} \pi_{i,k} = \sum_{k=1}^{l} \pi_{i,k} = 1.$$

Ezt kellett belátni.

9.1.9.
Ha Z egy véges halmaz és $P = [\pi_{i,j}]$ a sztochasztikus mag reprezentálta sztochasztikus mátrix, úgy a $\{z_{i_1}, z_{i_2}, \dots, z_{i_r}\}$ részhalmaz pontosan akkor következmény-zárt, ha minden $i \in \{i_1, \dots, i_r\}$ esetén az i-edik sornak az $j \notin \{i_1, \dots, i_r\}$ sorszámú oszlopaira $\pi_{i,j} = 0$.

Ez úgy is fogalmazható, hogy a P mátrixból megtartva csak az $i_1, \ldots, i_r$ indexekkel rendelkező sorokat és oszlopokat, a kapott mátrix is egy $r \times r$ méretű sztochasztikus mátrix.

Most példát mutatunk következmény-zárt halmazra és olyanra is, amely nem ilyen.

9.1.10. állítás. *A visszatérő elemek halmaza egy következmény-zárt halmaz. A tranziens elemek halmazának minden nem üres részhalmaza egy nem következmény-zárt halmaz.*

Bizonyítás. Legyen ugyanis u egy visszatérő elem, és v annak következménye. Persze feltehető, hogy $v \neq u$. Ha w a v-nak v-tól különböző következménye, akkor látni fogjuk, hogy v is w következménye. Ez nyilvánvaló $w = u$ esetben. Egyébként pedig a tranzitivitás szerint w következménye u-nak is, és u visszatérő volta igazolja, hogy u is következménye w-nek. Újra használva a tranzitivitást, kapjuk, hogy v következménye w-nek.

Ha a tranziens elemek halmazának F egy nem üres, következmény-zárt részhalmaza, akkor pusztán az ezen indexek feletti részmátrix is egy sztochasztikus mátrix. E részmátrixnak van tehát visszatérő eleme, ami nem lehetséges, hiszen ez az elem visszatérő lenne az eredeti mátrix szerint is, márpedig F éppen a tranziens elemeket tartalmazza. $\square$

9.1.11. definíció (ergodikus halmaz)**.** Legyen Q egy sztochasztikus mag a véges Z halmaz felett. A Z halmaz részhalmazai közül a nem üres, minimális, következmény-zárt halmazokat *ergodikus* halmaznak mondjuk.

A minimalitáson azt értjük, hogy az $E \subseteq Z$ ergodikus halmaznak nincs nem üres valódi következmény-zárt részhalmaza, ergo ha $G \subseteq E$, $G \neq \varnothing$ egy következmény-zárt halmaz, az csak úgy lehetséges, ha $G = E$.

9.1.12. állítás. *Legyen P sztochasztikus mátrix a Z halmaz felett. Jelölje E a visszatérő és F a tranziens elemek halmazát. Ekkor*

1. *Z előáll $Z = F \cup E$ diszjunkt egyesítés alakjában, ahol $E \neq \varnothing$.*
2. *Ha a tranziens elemek F halmaza nem üres, akkor F nem következmény-zárt halmaz.*
3. *Az visszatérő elemek E halmaza előáll mint véges sok, diszjunkt, nem üres, ergodikus halmaz egyesítése.*

Bizonyítás. Az első állítást már igazoltuk az előbbiekben, hiszen minden elem vagy tranziens vagy visszatérő, és minden sztochasztikus mátrixnak van legalább egy visszatérő eleme.

A második állítást még erősebb formában is láttuk: A tranziens elemek tetszőleges nem üres részhalmaza nem alkot következmény-zárt halmazt.

Vezessük be a visszatérő elemek E halmazán a következő relációt: Az $u, v \in$ $\in E$ mellett

$$u \simeq v \iff u = v, \text{ vagy } u \text{ következménye } v\text{-nek}$$

Persze e reláció reflexív. Ha $u \simeq v$ és $v \neq u$, akkor u következménye v-nek. De v is visszatérő, így minden nála különböző következményének, ergo u-nak is, maga v is következménye, azaz $v \simeq u$ fennáll, megmutatva, hogy $\simeq$ reláció szimmetrikus is. Mivel a következményének lenni egy tranzitív reláció, ezért $\simeq$ is tranzitív. Összefoglalva: a $\simeq$ reláció egy ekvivalenciareláció az E halmaz felett. Az alaphalmaz végessége miatt E előáll

$$E = \cup_{k=1}^{M} E_k$$

legfeljebb l darab ekvivalenciaosztály egyesítéseként.

Már csak azt kell igazolni, hogy minden egyes E_k ekvivalenciaosztály ergodikus halmaz. Az E_k következmény-zárt halmaz, hiszen ha z_j következménye $z_i \in E_k \subseteq E$-nek, akkor a visszatérő elemek következmény-zártsága szerint $z_j \in E$ is fennáll. Definíció szerint tehát $z_j \simeq z_i$. No de e reláció szimmetrikus is, ergo z_j és z_i azonos ekvivalenciaosztályba esnek, azaz $z_j \in E_k$.

Legyen most $G \subseteq E_k$ egy nem üres, következmény-zárt részhalmaz. Ha $z \in G$ és $w \in E_k, z \neq w$, akkor z és w egyazon ekvivalenciaosztály elemei, így $w \simeq z$. Ez definíció szerint azt jelenti, hogy w következménye z-nek. No de G egy következmény-zárt halmaz, tehát $w \in G$ is fennáll. Megmutattuk tehát, hogy $E_k \subseteq G$ is teljesül, ami éppen azt jelenti, hogy E_k egy nem üres, minimális, következmény-zárt halmaz. □

9.1.13. következmény. *Minden sztochasztikus mátrixhoz van olyan permutáció, amelyet a sorokra és az oszlopokra alkalmazva az alábbi módon partícionált mátrixhoz jutunk.*

$$\begin{array}{c|ccccc}
 & F & E_1 & E_2 & \dots & E_M \\ \hline
F & R_{0,0} & R_{0,1} & R_{0,2} & \dots & R_{0,M} \\
E_1 & & R_{1,1} & & \dots & \\
E_2 & & & R_{2,2} & \dots & \\
\vdots & \vdots & & & \ddots & \vdots \\
E_M & & & & \dots & R_{M,M}
\end{array}$$

9.1.14. állítás. *Egy sztochasztikus mátrixnak pontosan akkor van egyetlen egy ergodikus halmaza, ha létezik $s_1 \in Z$, amely tetszőleges $s \neq s_1, s \in Z$ elem következménye.*

Bizonyítás. Tegyük fel, hogy egyetlen ergodikus részhalmaz van. A visszatérő elemeken a fent bevezetett $\simeq$ ekvivalenciarelációnak ekkor egyetlen egy ekvivalenciaosztálya van, azaz bármelyik visszatérő elem következménye bármelyik másik visszatérőnek. Tudjuk (9.1.7), hogy ha van tranziens elem, akkor annak van visszatérő következménye, így a tranzitivitás szerint bármelyik visszatérő elem rendelkezik a kívánt tulajdonsággal.

Ha s_1 következménye minden más elemnek, akkor az tranziens nem lehet, hiszen tranziens elem nem következménye az egyik elemnek. Az s_1 tehát olyan visszatérő elem, amely minden visszatérő elemnek következménye is. Ez azt jelenti, hogy a $\succeq$ relációnak egyetlen ekvivalenciaosztálya van, ezért csak egyetlen ergodikus részhalmaz van. □

9.2. Invariáns eloszlások

9.2.1. állítás. *Legyen* Π *egy sztochasztikus mátrix, és képezzük az*

$$A_n = \frac{1}{n}\sum_{t=0}^{n-1}\Pi^t$$

közepeket. Ekkor ez a sorozat a sztochasztikus mátrixok körében konvergens, azaz létezik Q *sztochasztikus mátrix, amelyre*

$$A_n \to Q$$

Itt a konvergenciát a mátrix elemeire kell érteni. A Q *határmátrixra teljesül a*

$$\Pi Q = Q = Q\Pi$$

azonosság.

Bizonyítás. Világos, hogy A_n minden $n \in \mathbb{N}$ természetes szám mellett sztochasztikus mátrixok konvex lineáris kombinációjaként maga is sztochasztikus mátrix. A mátrix elemei tehát nem negatívak és legfeljebb 1 az értékük. Egy korlátos valós sorozat pontosan akkor konvergens, ha tetszőleges részsorozatának azonos a határértéke. Ezt a tuljadnoságát fogjuk A_n mátrixsorozatnak is igazolni.

Legyen tehát A_{n_k} egy olyan részsorozat, amelyre $A_{n_k} \to Q$. Világos, hogy Q is sztochasztikus mátrix. Ekkor:

$$\Pi A_{n_k} = \frac{1}{n_k}\sum_{t=0}^{n_k-1}\Pi^{k+1} = \frac{1}{n_k}\sum_{t=1}^{n_k}\Pi^k = A_{n_k} + \frac{1}{n_k}\left(\Pi^{n_k} - I\right).$$

Itt az utolsó tag a zérus mátrixhoz tart, hiszen egy korlátos sorozatnak és egy végtelenhez tartó sorozatnak a hányadosa a mátrixsorozat minden rögzített eleme. Így $\Pi A_{n_k} \to Q$. A szorzás folytonossága szerint persze $\Pi A_{n_k} \to \Pi Q$, így $Q = \Pi Q$. Innen nyilvánvaló indukció mutatja, hogy $Q = \Pi^t Q$ minden $t \in \mathbb{N}$. Mindkét oldalon képezve az n elemű $\frac{1}{n}$ számokkal súlyozott konvex kombinációt már azt kapjuk, hogy $Q = A_n Q$ is fennáll. Teljesen hasonlóan, jobbról szorzásra áttérve, azt kapjuk, hogy

$$QA_n = Q = A_n Q$$

teljesül minden $n \in \mathbb{N}$ mellett.

Legyen most A_{n_j} egy másik konvergens részsorozat, amelyre $A_{n_j} \to R$. Az egész eddigi gondolatot ismételhetjük Q helyett R-re, így azt kapjuk, hogy

$$RA_n = R = A_n R.$$

No de $QA_{n_j} \to QR$ és $A_{n_j} Q \to RQ$ így felhasználva, hogy a bal oldali mátrixszorzatok értéke a konstans Q

$$QR = Q = RQ.$$

No persze R és Q szerepe szimmetrikus, emiatt felcserélhetjük egyiket a másikkal. Így az

$$RQ = R = QR$$

egyenlőség is fennáll, amiből már $Q = R$ is következik. Ezt kellett belátni. □

9.2.2. állítás. *Legyen Π egy sztochasztikus mátrix és Q az $A_n = \frac{1}{n}\sum_{t=0}^{n-1}\Pi^t$ közepek határértéke. Ekkor a Q sztochasztikus mátrix sorai invariáns eloszlásai Π-nek, sőt Π invariáns eloszlásainak halmaza azonos a Q mátrix sorainak konvex burkával.*

Tetszőleges $p_0 \in \Delta_l$ valószínűségeloszlás esetén a $p_{t+1} = p_t \Pi$ iteráció számtani közepeiből álló sorozat konvergál a Π egy invariáns eloszlásához. Pontosabban:

$$\frac{1}{n}\sum_{t=0}^{n-1} p_t \to p_0 Q.$$

Bizonyítás. Legyen $p_0 \in \Delta_l$ tetszőlegesen rögzítve. Világos, hogy

$$p_1 = p_0\Pi, p_2 = p_1\Pi = p_0\Pi^2, \ldots, p_t = p_0\Pi^t.$$

Az $A_n \to Q$ miatt $p_0 A_n \to p_0 Q$, és $p_0 A_n = \frac{1}{n}\sum_{t=0}^{n-1} p_0 \Pi^t$, ami épp az iterált sorozat számtani közepeinek konvergenciáját jelenti.

A $Q\Pi = \Pi$ azonosság szerint

$$(p_0 Q)\,\Pi = p_0\,(Q\Pi) = p_0 Q.$$

Ez éppen azt jelenti, hogy $p_0 Q$ egy invariáns eloszlása Π-nek. Mivel ez minden $p_0 \in \Delta_l$ mellett igaz, ezért Q sorainak tetszőleges konvex kombinációja a Π egy invariáns eloszlását adja.

Legyen most $p\Pi = p, p \in \Delta_l$, azaz egy invariáns eloszlás. Ekkor $p_0 = = p$ választással az összes iteráltra $p_t = p$, és persze $pA_n = p$ is fennáll. Alkalmazva az eddig már igazoltakat kapjuk, hogy

$$p = pA_n = pQ$$

Ez éppen azt jelenti, hogy a p invariáns eloszlás a Q sorainak konvex kombinációja, tehát a Π-re invariáns eloszlások halmaza éppen Q sorainak konvex burka. $\square$

9.2.3.
Az imént azt mutattuk meg, hogy

$$\{p_0 Q : p_0 \in \Delta_l\} = \{p \in \Delta_l : p\Pi = p\}.$$

Ezek következménye, hogy a Q mátrix egyszerűen mutatja azt az esetet, mikor Π-nek csak egyetlen egy invariáns eloszlása van.

9.2.4.
A Π-nek pontosan akkor van egyetlen egy invariáns eloszlása, ha Q minden sora azonos.

9.3. Invariáns eloszlások unicitása

Természetesen merül fel, hogy amennyiben a Banach-fixponttételt alkalmazzuk az

$$p \mapsto p\Pi, p \in \Delta_l$$

folytonos függvényre, akkor az invariáns eloszlások unicitására is elegendő feltételt kapunk. Látnunk kell, hogy ezzel az ötlettel csak elegendő feltételt kaphatunk az invariáns eloszlás egyértelműségére. Legyen ugyanis a sztochasztikus magot megadó mátrix a

$$\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Ekkor tetszőleges $\alpha \in [0,1]$ mellett $(\alpha, 1-\alpha)\,\Pi = (1-\alpha, \alpha)$, tehát egyetlen fixpont van az $\left(\frac{1}{2}, \frac{1}{2}\right)$ eloszlás. Az is világos, hogy $\|f(p) - f(q)\| = \|p - q\|$ minden $p, q \in \Delta_2$ mellett, ami azt jelenti, hogy az invariáns eloszlás unicitása anélkül is fennállhat, hogy a $p \mapsto p\Pi$ leképezés kontraktív módon képezné Δ_l-et Δ_l-be.

Most olyan erős feltételt keresünk, ami már kontrakciót is biztosítja.

9.3.1. lemma. *Legyen* Π *egy sztochasztikus mátrix. Definiálja*

$$\varepsilon_j = \min\{\pi_{i,j} : i = 1, \ldots, l\} \quad \textit{és} \quad \varepsilon = \sum_{j=1}^{l} \varepsilon_j.$$

Ekkor az $f : \Delta_l \to \Delta_l, f(p) = p\Pi$ *leképezésre az*

$$\|f(p) - f(q)\| \leq (1-\varepsilon)\|p-q\|$$

egyenlőtlenség minden $p, q \in \Delta_l$ *mellett teljesül.*

Bizonyítás. A következő számolásban azt használjuk, hogy $p, q \in \Delta_l$ és hogy a Π egy sztochasztikus mátrix.

$$\begin{aligned}
&\|f(p) - f(q)\| = \|(p-q)\Pi\| = \\
&= \sum_{j=1}^{l}\left|\sum_{i=1}^{l}(p_i - q_i)\pi_{i,j}\right| = \sum_{j=1}^{l}\left|\sum_{i=1}^{l}(p_i-q_i)(\pi_{i,j}-\varepsilon_j) + \sum_{i=1}^{l}(p_i-q_i)\varepsilon_j\right| = \\
&= \sum_{j=1}^{l}\left|\sum_{i=1}^{l}(p_i-q_i)(\pi_{i,j}-\varepsilon_j) + \varepsilon_j\sum_{i=1}^{l}(p_i-q_i)\right| = \\
&= \sum_{j=1}^{l}\left|\sum_{i=1}^{l}(p_i-q_i)(\pi_{i,j}-\varepsilon_j) + \varepsilon_j 0\right| \leq \\
&\leq \sum_{i=1}^{l}\sum_{j=1}^{l}|p_i-q_i|\,(\pi_{i,j}-\varepsilon_j)\,| = \\
&= \sum_{i=1}^{l}|p_i-q_i|\sum_{j=1}^{l}(\pi_{i,j}-\varepsilon_j) = \sum_{i=1}^{l}|p_i-q_i|(1-\varepsilon) = (1-\varepsilon)\|p-q\|.
\end{aligned}$$

Ezt kellett belátni. □

Világos, hogy az előző tételben $\varepsilon > 0$ pontosan akkor teljesül, ha van olyan $z_j \in Z$ sokk-állapot, amelybe egy lépés hatására minden z_i állapotból pozitív valószínűséggel kerül a rendszer.

9.3.2.

Ha a Π sztochasztikus mátrixra teljesül, hogy van olyan oszlopa, amelyben minden elem zérustól különböző, akkor a folyamatnak csak egy invariáns eloszlása van és tetszőleges $\pi_0 \in \Delta_l$ mellett a $p_{k+1} = p_k\Pi$ rekurzió konvergál az egyetlen invariáns eloszláshoz.

Ebben az esetben ugyanis $\varepsilon = \sum_{i=1}^{l}\varepsilon_i > 0$, ezért az $f : \Delta_l \to \Delta_l$, $f(p) = p\Pi$ leképezés kontrakció a $(\Delta_l, \|\cdot\|)$ Banach-téren. A Banach-fixponttétel szerint a leképezésnek van egyetlen egy fixpontja, ami persze definíció szerint a Π egyetlen egy invariáns eloszlása, és a $p_{k+1} = f(p_k)$ iteráció a fixponthoz konvergál tetszőleges $p_0 \in \Delta_l$ kiindulási eloszlás mellett.

Most ezt a gondolatot általánosítjuk.

9.3.3. állítás. *Legyen* Π *egy sztochasztikus mátrix. Jelölje rögzített* $N \in \mathbb{N}$ *természetes szám mellett*

$$\varepsilon_j^{(N)} = \min\left\{\pi_{i,j}^{(N)} : i = 1, \dots, l\right\} \qquad \textit{és} \qquad \varepsilon^{(N)} = \sum_{j=1}^{l} \varepsilon_j^{(N)},$$

ahol $\pi_{i,j}^{(N)}$ *jelöli a* Π^N *mátrix* i*-edik sorának* j*-edik elemét.*

Ha létezik $N \in \mathbb{N}$*, melyre* $\varepsilon^N > 0$*, akkor a* Π*-nek csak egyetlen invariáns eloszlása van és a* $p_{k+1} = p_k\Pi$ *rekurzióval megadott sorozat tetszőleges* $p_0 \in \in \Delta_l$ *vektorból indulva az invariáns eloszláshoz konvergál.*

Bizonyítás. Ha $q \in \Delta_l$ invariáns eloszlása Π-nek, akkor az invariáns eloszlása Π^N-nek is, hiszen

$$q\Pi^N = (q\Pi)\,\Pi^{N-1} = q\Pi^{N-1} = (q\Pi)\,\Pi^{N-2} = q\Pi^{N-2} = \dots = q\Pi = q.$$

Alkalmazzuk az eddig belátottakat a Π^N szorzatra. Tudjuk hát, hogy Π^N-nek pontosan egy invariáns eloszlása van, ezért Π-nek sem lehet több. Persze minden sztochasztikus mátrixnak van invariáns eloszlása, ezért Π-nek is pontosan egy inveriáns eloszlása van.

Most azt mutatjuk meg, hogy a tételben definiált rekurzív sorozat tetszőlegesen választott $p_0 \in \Delta_l$ mellett konvergál ehhez a $p^\star \in \Delta_l$ egyértelműen létező invariáns eloszláshoz. Tekintsük a

$$k \mapsto p_0\Pi^{kN}$$

sorozatot. Az előzőek szerint $\lim_{k\to\infty} p_0\Pi^{kN} = p^\star$.

Hasonlóan, nézzük most a

$$k \mapsto p_0\Pi^{kN+1}$$

sorozatot. Mivel $p_0\Pi^{kN+1} = (p_0\Pi)\,\Pi^{kN}$, ezért e sorozat tekinthető úgy, mint a $p_1 = p_0\Pi$ pontból induló rekurziója a Π^N szorzatnak, ezért $\lim_{k\to\infty} p_0\Pi^{kN+1} = = p^\star$.

Teljesen azonos módon a

$$k \mapsto p_0\Pi^{kN+2}$$

sorozatról is látjuk, hogy $\lim_{k\to\infty} p_0\Pi^{kN+2} = p^\star$.

Ezt ismételhetjük, egészen a

$$k \mapsto p_0\Pi^{kN+N-1}$$

sorozatig és persze $\lim_{k\to\infty} p_0\Pi^{kN+N-1} = p^\star$.

Azt kaptuk tehát, hogy a tételben megadott rekurzív sorozat N darab a $p^\star$ ponthoz konvergáló sorozat összefésülésével keletkezik. Világos, hogy az ugyanahhoz a ponthoz konvergáló sorozatok összefésültje is a szóbanforgó ponthoz konvergál. $\square$

9.3.4. állítás. *Ha egy sztochasztikus mátrixnak csak egyetlen ergodikus halmaza van, akkor annak pontosan egy invariáns eloszlása van.*

Bizonyítás. Jelölje $R = \frac{1}{2}(I + \Pi)$. Mivel I és Π sztochasztikus mátrixok, ezért R mint ezek konvex kombinációja is egy sztochasztikus mátrix. Persze ebből az is következik, hogy R valamennyi hatványa is sztochasztikus mátrix. Most vegyük észre, hogy egy $p \in \Delta_l$ eloszlás pontosan akkor invariánsa Π és R egyikének ha az invariáns eloszlása a másiknak. Ugyanis

$$pR = p \iff \frac{1}{2}p + \frac{1}{2}p\Pi = p \iff \frac{1}{2}p\Pi = \frac{1}{2}p \iff p\Pi = p.$$

Ha Π-nek egyetlen ergodikus halmaza van, akkor létezik j, hogy z_j következménye minden nála különböző z_i-nek. Ez azt jelenti, hogy létezik j, hogy minden $i \neq j$-hez létezik $n_i \in \mathbb{N}$, hogy $\pi_{i,j}^{n_i} > 0$. A választott z_j visszatérő, hiszen az ergodikus halmaz egyik eleme. Mivel minden visszatérő elem saját magának is következménye, ezért $i = j$-hez is van a $\pi_{i,i}^{n_i} > 0$ elvárásnak megfelelő n_i egész. Jelölje $N = \max\{n_1, n_2, \ldots, n_l\}$. Persze a Binomiális tétel a mátrixgyűrűben is érvényes, ezért

$$R^N = \frac{1}{2^N} \sum_{i=0}^{N} \binom{N}{i} \Pi^i.$$

Látjuk tehát, hogy R^N mátrix j-edik oszlopában minden elem pozitív. Alkalmazva az előző tételt, azt kapjuk, hogy R-nek csak egyetlen invariáns eloszlása van, emiatt persze Π-nek is csak egy invariáns eloszlása van. $\square$

III. rész

Függelék

10. fejezet

A Hausdorff- és a Banach–Tarski-paradoxonról

Delians: „How can we be rid of the plague?”
Delphic Oracle: „Construct a cubic altar of the double the size of the existing one.”
Banach and Tarski: „Can we use the Axiom of Choice?”
Stan Wagon (1985) [Wagon (1985)]

Egy narancs véges sok részre bontható olyan módon, hogy a részekből két narancs is összerakható; vagy egy biliárd golyó és Stephan Banach életnagyságú szobra feldarabolható páronként egybevágó részekre [Stromberg (1979)]. Ezek a szokásos népszerűsítő formái Hausdorff, Banach és Tarski idevágó tételeinek:

Adott két korlátos halmaz $\mathbb{R}^3$-ban, amelyeknek a belseje nem üres. Ekkor e két halmaz felbontható véges sok diszjunkt részhalmazra úgy, hogy az egyes részhalmazok egymásba vihetők a tér egybevágósági transzformációinak segítségével.

A fenti állítás egyrészt a 20. századi matematika egyik legnagyobb meglepetése és jól mutatja a háromdimenziós térbeli végtelen halmazok képzeletbeli jellegét. A tétel kizárja a nem triviális, végesen additív forgatás- és eltolásinvariáns mértékek létezését az $\mathbb{R}^n$ $(n \geq 3)$ tér egész hatványhalmazán; és rámutat az olyan mértékkiterjesztési eljárások szükségszerűségére, mint például a Lebesgue-mérték szokásos Caratheodory-féle konstrukciója.

Másrészt emlékezzünk arra, hogy mennyire természetes és könnyű gondolat az, hogy a végtelen halmaz fogalmát karakterizálhatjuk azzal a tulajdonsággal, hogy az valamilyen értelemben azonos, ekvipotens, egy valódi részhalmazával. Itt észre kell vennünk, hogy a végtelen fogalmának megértésére való törekvés az emberi tudásvágy egyik legrégebbi megjelenési formája, és valószínűleg pszichénk egészen mélyen kódolt része. A valódi részhalmazzal való azonosság a Szentháromság fogalmában is előkerül. Ott is Isten végtelenségének megnyilvánulását láthatjuk benne. Ez utóbbi végtelenségfogalomnak bizonnyal semmi köze a halmazelméleti végtelenség fogalmához, de újabb érv amellett, hogy a Banach–Tarski-paradoxonra ne mint egy matematikai szörnyszülöttre gondoljunk, hanem a végtelenség egy következményére.

10.1. Hausdorff-paradoxon

Először Vitali programját gondoljuk át, persze sokkal absztraktabb környezetben, majd bevezetjük a kongruens, az ekvidekompozábilis és a paradox halmazok és csoportok fogalmát. Példát is adunk paradox halmazra, amiből már könnyen fog következni, hogy a háromdimenziós egységgömb S^2 felülete, egy megszámlálható halmaztól eltekintve megduplázható. A továbbiakban S^n jelöli az $n+1$-dimenziós egységgömb felületét.

10.1.1. Particionálás egy csoport hatásaként

10.1.1. definíció (a G csoport hat az X halmazon)**.** Legyen G egy csoport és X egy rögzített halmaz. Azt mondjuk, hogy *G az X-en hat*, ha minden $g \in G$ -hez létezik egyértelműen egy $\widehat{g} : X \to X$ bijekció, melyre

$$\widehat{gh} = \widehat{g} \circ \widehat{h} \quad \text{és} \quad \widehat{1} = \mathrm{id}. \tag{10.1}$$

Azt mondjuk, hogy G *fixpontmentesen hat* az X-en, ha minden $g \in G$, $g \neq 1$ esetén $\widehat{g}$ egy fixpontmentes bijekciója X-nek.

10.1.2.
Amennyiben G az X-en hat, úgy a

$$\widehat{g^{-1}} = (\widehat{g})^{-1}$$

azonosság is teljesül, hiszen $g \cdot g^{-1} = 1$ miatt $\widehat{g} \circ \widehat{g^{-1}} = \mathrm{id}$ is fennáll.

10.1.3.
Legyen G egy az X halmazon ható csoport, és $M \subseteq X$ egy rögzített részhalmaz. Vezessük be a következő $\mathcal{P}(G) \to \mathcal{P}(X)$ operációt:

$$\begin{aligned} B \subseteq G \text{ esetén } B^{\star} &\doteq \{x \in X : \exists m \in M,\ \exists g \in B, x = \hat{g}(m)\} = \\ &= \cup \{\hat{g}(M) : g \in B\}. \end{aligned}$$

Jelölje tetszőleges $B \subseteq G$ és tetszőleges $g \in B$ mellett $gB \doteq \{gh : h \in B\} \subseteq$ $\subseteq X$ halmazt. Ekkor

$$(gB)^{\star} = \hat{g}\,(B^{\star})\,,$$

ugyanis

$$\begin{aligned}\hat{g}\,(B^{\star}) &= \cup\left\{\hat{g}\left(\hat{h}\,(M)\right) : h \in B\right\} = \\ &= \cup\left\{\widehat{gh}\,(M) : h \in B\right\} = \left\{\widehat{f}\,(M) : f \in gB\right\} = (gB)^{\star}\,.\end{aligned}$$

10.1.4.
Érdemes a következő példákat a későbbi tételek folyamán is szem előtt tartani.

1. Legyen G egy csoport és $X \doteq G$. Adott $g \in G$ mellett a $\widehat{g} : G \to G$ bijekció legyen a g-vel képzett balszorzás, azaz

   $$\widehat{g}\,(h) \doteq gh.$$

   Könnyen látható, hogy ez valóban bijekció és (10.1) is fennáll, tehát G valóban egy a G-n ható csoport. Sőt: G fixpontmentesen hat G-n.

2. Legyen $n \geq 2$ természetes szám mellett $G = SO\,(n+1)$, azaz az S^n forgatásainak csoportja és $X = S^n$. Adott $g \in G$ esetén $\widehat{g} = g$. Látható, hogy $SO\,(n+1)$ hat az S^n-en, de nem fixpontmentesen. Ugyanis $n \geq 2$ miatt, a forgatás tengelye és az $n+1$-dimenziós egységgömb két közös pontja egyben a forgatás fixpontja is.

3. Legyen $G = SO\,(2)$, azaz az $\mathbb{R}^2$ sík S^1 egységkörének forgatáscsoportja és $X = S^1$. Adott $g \in G$ esetén $\widehat{g} = g$. Az $SO\,(2)$ csoport fixpontmentesen hat az S^1-en.

4. Legyen G az $\mathbb{R}^2$ sík S^1 egységkörének racionális szögekkel való forgatáscsoportja, és $X = S^1$. Adott $g \in G$ esetén $\widehat{g} \doteq g$. A G csoport persze fixpontmentesen hat az S^1-en.

5. Legyen $X = [0{,}2\pi)$, és G a $[0{,}2\pi)$-beli racionális számok additív csoportja a modulló 2π-vel számított összeadás művelettel ellátva. Adott $g \in G$ esetén $\widehat{g} : [0{,}2\pi) \to [0{,}2\pi)$ a

   $$\widehat{g}\,(r) \doteq g + r \qquad \mathrm{mod}\ \ (2\pi)$$

   módon megadott függvény. Világos, hogy ez csak átfogalmazása az előző példának, és G fixpontmentesen hat $[0{,}2\pi)$-en.

6. Legyen $X = \mathbb{R}$ és G a racionális számok additív csoportja. Adott $g \in G$ esetén $\widehat{g} : \mathbb{R} \to \mathbb{R}$ a

   $$\widehat{g}\,(r) \doteq g + r$$

   módon definiált függvény. Világos, hogy G fixpontmentesen hat X-en.

10.1.5. állítás. *Legyen G egy az X halmaz fölött ható csoport. Jelölje $G(x) \doteq \{\widehat{g}(x) : g \in G\}$ az x pont pályáját.*

(1) *Ekkor a pályákból álló halmazrendszer az X alaphalmaz egy partícióját alkotja, azaz*

$$\{G(x) : x \in X\} \text{ partíciója } X\text{-nek.} \tag{10.2}$$

(2) *Tegyük most fel, hogy G fixpontmentesen hat az X halmazon. Vegyünk ki a fenti partíció minden halmazából pontosan egy-egy elemet és jelölje M az így kapott reprezentáns halmazt. Ekkor a $\{\widehat{g}(M) : g \in G\}$ halmazrendszer partíciója X-nek, amely ekvipotens G-vel, azaz*

$$\hat{g}(M) \neq \varnothing; \quad g_1 \neq g_2 \Rightarrow \hat{g}_1(M) \cap \hat{g}_2(M) = \varnothing; \quad X = \cup_{g\in G}\hat{g}(M). \tag{10.3}$$

(3) *Rögzítsük az iménti M reprezentáns halmazt, és tekintsünk a G egy $G = \cup_{i=}^{n} B_i$ partícióját. Ekkor $X = \cup_{i=1}^{n} B_i^{\star}$ az X partícióját adja ahol $\star$ a 10.1.3. szakaszban bevezetett transzformáció.*

Bizonyítás. (1) Először is $x \in G(x)$ minden $x \in X$ mellett fennáll, hiszen az $1 \in G$ multiplikatív egységelemre $\widehat{1} = \mathrm{id}$. Másodszor, ha $u \in G(x) \cap G(y)$ azaz valamely $r, s \in G$ mellett $\widehat{r}(x) = u = \widehat{s}(y)$ és $z \in G(x)$, akkor $z \in G(y)$ is fennáll. Ugyanis ha $z = \widehat{t}(x)$, valamely $t \in G$ esetén, akkor

$$z = \widehat{t}(x) = \widehat{t}\left(\widehat{r}^{-1}(u)\right) = \widehat{t}\left(\widehat{r^{-1}}(u)\right) = \widehat{t}\left(\widehat{r^{-1}}(\widehat{s}(y))\right) = (\widehat{tr^{-1}s})(y) \in G(y).$$

Megmutattuk tehát, hogy $G(x) \subseteq G(y)$. Hasonlóan kapjuk, hogy $G(y) \subseteq$ $\subseteq G(x)$, tehát

$$G(x) \cap G(y) \neq \varnothing \Rightarrow G(x) = G(y),$$

ami azt jelenti, hogy (10.2) valóban egy partícióját alkotja X-nek.

(2) Most tegyük fel, hogy $\{\widehat{g} : g \in G, g \neq 1\}$ fixpontmentes bijekciók. Persze minden $x \in X$ mellett $x \in G(x)$ és arra az egyetlen $m \in M$-re, amelyre $m \in G(x)$, az $x \in G(x) = G(m)$ is teljesül, ami épp azt jelenti, hogy létezik $g \in G$, melyre $x = \widehat{g}(m)$.

Tegyük fel most, hogy valamely $g, h \in G$ mellett $\widehat{g}(M) \cap \widehat{h}(M) \neq \varnothing$, azaz létezik $m, m' \in M$, melyekre $\widehat{g}(m) = \widehat{h}(m')$. Ekkor persze

$$G(m) \cap G(m') \neq \varnothing,$$

ezért $G(m) = G(m')$. De mivel a (10.2) partíció minden részhalmazából csak egyetlen elemet vettünk ki az M halmaz képzéséhez, ezért $m = m'$ is teljesül. Így $\widehat{g}(m) = \widehat{h}(m)$, amiből $m = \left(\widehat{h}\right)^{-1}(\widehat{g}(m)) = (\widehat{h^{-1}g})(m)$ következik. Tehát a $h^{-1}g \in G$ egy olyan elem, melyhez rendelt bijekciónak van fixpontja. A feltételünk szerint ez csak úgy lehetséges, ha $h^{-1}g = 1$, azaz $h = g$.

(3) Az állítás szinte csak átfogalmazása (2)-nek. $\square$

A bizonyítandó állítás szempontjából kitérő, de érdemes felfigyelni arra, hogy a fenti 10.1.5. állítás szoros kapcsolatban van a nem mérhető halmaz szokásos konstrukciójával.

10.1.6.
Tekintsük először a 4. vagy 5. példát. Az ott konstruált M halmaz biztosan nem Lebesgue-mérhető, hiszen az előző állítás szerint a $[0{,}2\pi)$ intervallum előáll

$$[0{,}2\pi) = \bigcup_{g\in\mathbb{Q}} (g + M)$$

diszjunkt egyesítés alakban. Mivel a Lebesgue-mérték eltolás-invariáns, ezért M mérhetőségéből a $g+M$ eltolt halmazok mérhetősége is következne, majd a $2\pi = \lim_{n\to\infty} n\lambda(M)$, ami nyilván ellentmondás, hiszen ez utóbbi szám csak 0 vagy $+\infty$ lehet.

A 6. példa alkalmazása is $\mathbb{R}$ egy nem Lebesgue-mérhető halmazához vezet. A 10.1.5. állításbeli M halmaz konstrukcióját csak egy kicsit kell megváltoztatnunk. Egyrészt minden $x \in \mathbb{R}$ mellett van olyan $y \in \mathbb{R}$, $0 \le y \le 1$, hogy $x - y \in \mathbb{Q}$. Ez azt jelenti, hogy $G(x) \cap [0{,}1] \ne \varnothing$ minden $x \in \mathbb{R}$ esetén, tehát a reprezentánsokat választhatjuk a $[0{,}1]$ intervallumból, így az $M \subseteq [0{,}1]$ tartalmazás feltehető. Másrészt

$$[0{,}1] \subseteq \bigcup_{g\in\mathbb{Q}|g|\le 2} (g + M) \doteq H,$$

hiszen $x \in [0{,}1]$-hez létezik $g \in \mathbb{Q}$ melyre $x \in g + M$, azaz $x - g = m \in [0{,}1]$. Ezért $|g| = |x - m| \le 2$. Azt kaptuk tehát, hogy a H halmazra $[0{,}1] \subseteq H \subseteq$ $\subseteq [-2{,}2] + [0{,}1] = [-2{,}3]$. Amennyiben tehát M Lebesgue-mérhető lenne, akkor H is Lebesgue-mérhető lenne, és H Lebesgue-mértékére $1 \le \lambda(H) \le$ ≤ 5 állna fenn. De H definíciója szerint csak $\lambda(H) = 0$, vagy $\lambda(H) = \infty$ lehetséges, ami ugyanazt az ellentmondást adja, mint az előző példa.

10.1.7.
A fenti példából az is világos, hogy nincs $\mathcal{P}\left(S^1\right)$-en μ megszámlálhatóan additív forgatás-invariáns mérték, amelyre $0 < \mu\left(S^1\right) < \infty$. Ebből azonnal következik, hogy nincs olyan valós, eltolás-invariáns μ mérték az egész $\mathcal{P}(\mathbb{R})$-en, amelyre $0 < \mu([0{,}1]) < 1$, ezért az $A \mapsto A \times [0{,}1]^{n-1}$ leképezés segítségével rögtön láthatjuk, hogy nincs eltolás-invariáns μ mérték a $\mathcal{P}(\mathbb{R}^n)$ halmazon értelmezve, amelyre $\mu([0{,}1]^n) = 1$.

A 10.1.5. állításra úgy is nézhetünk, mint egy általános eszköz nem mérhető halmazok konstrukciójához.

10.1.2. Ekvidekompozábilis halmazok

10.1.8. definíció (kongruens halmazok, $A \cong_G B$)**.** Legyen G egy az X halmazon ható csoport, valamint $A, B \subseteq X$. Azt mondjuk, hogy az *A halmaz G-kongruens a B halmazzal* ($A \cong_G B$), ha létezik $g \in G$, amelyre $\widehat{g}\,(A) = B$.

10.1.9. definíció (ekvidekompozábilis halmazok, $A \sim_G B$)**.** Legyen G egy olyan csoport, amely az X halmazon hat, továbbá $A, B \subseteq X$. Azt mondjuk, hogy az *A és a B halmazok G-ekvidekompozábilisek* ($A \sim_G B$), ha létezik A-nak és B-nek olyan véges partíciója, amelynek elemei páronként G-kongruensek egymással.

Magyarul: létezik $A = \cup_{i=1}^{n} A_i$ és $B = \cup_{i=1}^{n} B_i$ diszjunkt felbontás, amelyekre $A_i \cong_G B_i$ minden $i = 1, \ldots, n$ mellett.

Elképzelhető, hogy egy halmaz ekvidekompozábilis egy valódi részhalmazával:

10.1.10. állítás. *Legyen G egy az X halmazon ható csoport, valamint $P \subseteq \subseteq A \subseteq X$, és tegyük fel, hogy létezik $g \in G$ és létezik $Q \subseteq X$, amelyekre*

$$P \subseteq Q \subseteq A \quad \textit{valamint} \quad \hat{g}\,(Q) = Q \setminus P.$$

Ekkor

$$A \sim_G (A \setminus P)\,.$$

Bizonyítás. Világos, hogy az $A = Q \cup (A \setminus Q)$ valamint az $A \setminus P = (Q \setminus P) \cup \cup (A \setminus Q)$ diszjunkt egyesítések, valamint $\hat{g}\,(Q) = Q \setminus P$ és $\hat{1}\,(A \setminus Q) = = A \setminus Q$. □

A fenti állítás feltételeit kielégítő konkrét transzformációkat találunk példaként a 10.2.1. lemmában és a 10.1.13. állításban.

10.1.11. definíció (paradox halmaz)**.** Legyen G egy csoport, amely az X halmazon hat és $A \subseteq X$. Azt mondjuk, hogy az *A halmaz G-paradox*, ha léteznek $B, C \subseteq A$; $B \cap C = \varnothing$ diszjunkt részhalmazok, amelyek külön-külön G-ekvidekompozábilisek A-val, azaz $B \sim_G A$ és $C \sim_G A$ is fennáll.

A 10.2.5. állításban látjuk majd, hogy a fenti definícióban $B \cup C = X$ is feltehető. Ehhez a Banach–Schröder–Bernstein-tétel kell, de amíg nincs szükségünk rá, addig nem érdemes kihasználni.

10.1.12. definíció (paradox csoport)**.** Egy csoportot *paradoxnak* mondunk, ha úgy tekintve mint egy saját maga felett ható csoport (lásd a 10.1.4. szakasz 1. példát), a G halmaz G-paradox.

Az alábbiakban példát adunk paradox csoportra. Ha arra gondolunk, hogy az egész számok azért ekvipotensek a páros számokkal mert mindegyiket meg kell szoroznunk 2-vel, talán nem is olyan meglepő, hogy milyen könnyen találkozhatunk egy ártatlan paradox csoporttal.

10.1.13. állítás. *Minden két elem generálta szabad csoport paradox.*

Bizonyítás. Szabad csoport elemeinek redukált szóként való előállítása egyértelmű, értelmes tehát a következő halmaz megadása. Jelölje $W(\alpha)$ a G csoport α-val kezdődő redukált szavaiból álló részhalmazát, amennyiben $\alpha = \sigma^{\pm 1}$ vagy $\alpha = \tau^{\pm 1}$, ahol σ és τ a két generáló eleme a G szabad csoportnak. Világos, hogy G előáll az alábbi diszjunkt egyesítésként:

$$G = W(\sigma) \cup W\left(\sigma^{-1}\right) \cup W(\tau) \cup W\left(\tau^{-1}\right) \cup \{\mathrm{id}\}.$$

Jelölje $P \doteq W(\tau) \cup W\left(\tau^{-1}\right) \cup \{\mathrm{id}\}$ és $Q \doteq P \cup W(\sigma)$. Látható, hogy $\sigma(Q) = W(\sigma) = Q \setminus P$, ami a 10.1.10. állítás szerint azt jelenti, hogy $G \sim W(\sigma) \cup W\left(\sigma^{-1}\right)$. A fenti gondolatban σ helyett τ-t írva, azt kapjuk, hogy $G \sim W(\tau) \cup W\left(\tau^{-1}\right)$, ami e két utóbbi halmaz diszjunktsága miatt igazolja az állítást. □

10.1.14. állítás. *Tekintsünk egy $A \subseteq X$ halmazt, amely G-paradox és legyen $A' \subseteq X$; $A \sim_G A'$. Ekkor A' is G-paradox halmaz.*

Bizonyítás. Tegyük fel, hogy $B, C \subseteq A$ diszjunkt halmazok, amelyekre $B \sim_G A$ és $C \sim_G A$ is fennáll. Létezik $A = \cup_{k=1}^{n} A_k$ partíció és léteznek $g_k \in G$ elemek amelyekkel $A' = \cup_{k=1}^{n} \hat{g}_k(A_k)$ az A' egy partícióját alkotja. Legyen $B' \doteq \cup_{k=1}^{n} g_k(A_k \cap B)$ és hasonlóan $C' \doteq \cup_{k=1}^{n} g_k(A_k \cap C)$. Világos, hogy $B' \cap C' = \varnothing$ és $B \sim_G B'$ valamint $C \sim_G C'$. A $\sim_G$ reláció tranzitivitása miatt tehát $B' \sim_G A'$ és $C' \sim_G A'$. Ezt kellett belátni. □

10.1.3. Hausdorff-paradoxon

Először is azt kell látnunk, hogy amennyiben van egy paradox csoportunk akkor ennek segítségével könnyen kaphatunk más paradox halmazokat is. Sőt, az alábbi állítás megfordítása is igaz (lásd [Su (1990)] és [Wagon (1985)]), de itt most érdektelen.

10.1.15. állítás. *Ha G az X halmazon fixpontmentesen ható paradox csoport, akkor X is G-paradox.*

Bizonyítás. Legyenek $B, C \subseteq G$ diszjunkt részhalmazai G-nek, amelyek különkülön ekvidekompozábilisek G-vel. Léteznek tehát

$$B = \cup_{i=1}^{n} B_i \quad \text{és} \quad C = \cup_{j=1}^{m} C_j$$

partíciók, valamint $g_i, h_j \in G$ elemek $i = 1, \ldots, n$, és $j = 1, \ldots, m$, amelyekkel

$$\cup_{i=1}^{n} g_i B_i = G \quad \text{és} \quad \cup_{j=1}^{m} h_j C_j = G$$

is egy-egy partícióját alkotja G-nek. A 10.1.5. állításból már tudjuk, hogy ha az M halmaz a $\{G(x) : x \in X\}$ partíció tetszőleges reprezentánsaiból áll, akkor $\{\widehat{g}(M) : g \in G\}$ halmazrendszer is partíciója X-nek, ezért a G egy particionálása az X egy particionálását indukálja a 10.1.3. szakaszban bevezetett $\star : \mathcal{P}(G) \to \mathcal{P}(X)$ transzformáció segítségével. Tehát $B^\star = \cup B_i^\star$ particionálását adja $B^\star$-nak és $G^\star = \cup (g_i B_i)^\star$ particionálását adja $G^\star$-nak. De

$$B^\star = \cup_{i=1}^n B_i^\star \quad \sim_G \quad \cup_{i=1}^n \hat{g}_i (B_i^\star) = \cup_{i=1}^n (g_i B_i)^\star = (\cup g_i B_i)^\star = G^\star = X.$$

Hasonlóan $C^\star \sim_G X$ is fennáll. □

A probléma tehát abban áll, hogy milyen módon kapunk $SO(3)$ forgatás csoport két elem generálta szabad részcsoportját?

10.1.16. állítás (Wierczkowski, 1958)**.** *Az S^2 gömbfelület $SO(3)$ forgatás csoportjának van két elem generálta szabad részcsoportja.*

Bizonyítás. Azt kell megmutatnunk, hogy van $SO(3)$-nak két eleme – ϕ és ρ –, amelyek által generált részcsoportban a redukált szavak előállítása egyértelmű. Ez azzal ekvivalens, hogy a $\phi^{\pm 1}$ és a $\rho^{\pm 1}$ jelekből az identitás transzformáció legalább egy hosszú szóként nem rakható ki.

Legyen ϕ a z-tengely körüli arccos $\frac{1}{3}$ fokkal való forgatás és ρ az x-tengely körüli ugyanekkora szögű forgatás. Ekkor ϕ-nek és ϕ^{-1}-nek valamint ρ-nak és ρ^{-1}-nek a mátrix reprezentációja:

$$\phi^{\pm 1} = \begin{pmatrix} \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0 \\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \rho^{\pm 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} \\ 0 & \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}.$$

Tekintsünk tehát a ϕ és ρ generálta részcsoportban egy w redukált szót, amely egyenlő az identitás transzformációval és legalább egy hosszú. Feltehető, hogy w utolsó betűje ϕ, hiszen ezt elérhetjük a $\phi^{-1} w \phi$ konjugálással. *Megmutatjuk, hogy amennyiben w egy a $\phi^{\pm 1}$ és a $\rho^{\pm 1}$ elemekből alkotott redukált szó, melynek utolsó betűje ϕ, akkor*

$$w[1{,}0{,}0] = \left[a, b\sqrt{2}, c\right]/3^k$$

alakú, ahol $a, b, c \in \mathbb{Z}$ egészek, továbbá 3 nem osztója b-nek. Ebből bőven következik, hogy w nem lehet az identitás transzformáció. (Látjuk majd, hogy a gyengébb állítást nehezebb lenne indukcióval bizonyítani.)

A w hossza szerinti indukció következik. Amennyiben w egy hosszú, azaz csak ϕ-ből áll, akkor

$$w[1{,}0{,}0] = \phi[1{,}0{,}0] = \left[1{,}2\sqrt{2}{,}0\right]/3,$$

ami valóban megfelel az állításnak

$$a_1 \doteq 1; \quad b_1 \doteq 2; \quad c_1 \doteq 0$$

jelöléssel.

Tegyük fel most, hogy $k > 1$ és minden k-nál rövidebb redukált szó teljesíti állításunkat, valamint w éppen k hosszú. Ekkor

$$w = \phi^{\pm 1} w' \quad \text{vagy} \quad w = \rho^{\pm 1} w',$$

ahol

$$w'\,[1{,}0{,}0] = \left[a_{k-1}, b_{k-1}\sqrt{2}, c_{k-1}\right]/3^{k-1}$$

alakban írható fel, persze $a_{k-1}, b_{k-1}, c_{k-1} \in \mathbb{Z}$ és b_{k-1} nem osztható 3-mal. Amennyiben $w = \phi^{\pm 1} w'$ alakú, akkor az alábbi egyszerű számolgatás szerint:

$$\begin{aligned}3^{k-1} w\,[1{,}0{,}0] &= 3^{k-1}\phi^{\pm 1} w'\,[1{,}0{,}0] = \phi^{\pm 1}\left[a_{k-1}, b_{k-1}\sqrt{2}, c_{k-1}\right] = \\ &= a_{k-1}\left[\frac{1}{3}, \pm\frac{2\sqrt{2}}{3}, 0\right] + b_{k-1}\sqrt{2}\left[\mp\frac{2\sqrt{2}}{3}, \frac{1}{3}, 0\right] + c_{k-1}\,[0{,}0{,}1] = \\ &= \left[a_{k-1} \mp 4b_{k-1}, \pm 2\sqrt{2}a_{k-1} + \sqrt{2}b_{k-1}, 3c_{k-1}\right]/3 = \\ &= \left[a_k, b_k\sqrt{2}, c_k\right]/3,\end{aligned}$$

ahol tehát

$$a_k \doteq a_{k-1} \mp 4b_{k-1}; \quad b_k \doteq b_{k-1} \pm 2a_{k-1}; \quad c_k = 3c_{k-1}. \tag{$\dagger$}$$

Hasonlóan, ha $w = \rho^{\pm} w'$, akkor

$$\begin{aligned}3^{k-1} w\,[1{,}0{,}0] &= 3^{k-1}\rho^{\pm 1} w'\,[1{,}0{,}0] = \rho^{\pm 1}\left[a_{k-1}, b_{k-1}\sqrt{2}, c_{k-1}\right] = \\ &= a_{k-1}\,[1{,}0{,}0] + b_{k-1}\sqrt{2}\left[0, \frac{1}{3}, \pm\frac{2\sqrt{2}}{3}\right] + c_{k-1}\left[0, \mp\frac{2\sqrt{2}}{3}, \frac{1}{3}\right] = \\ &= \left[3a_{k-1}, \sqrt{2}b_{k-1} \mp 2\sqrt{2}c_{k-1}, \pm 4b_{k-1} + c_{k-1}\right]/3 = \\ &= \left[a_k, b_k\sqrt{2}, c_k\right]/3,\end{aligned}$$

ahol most

$$a_k \doteq 3a_{k-1}; \quad b_k \doteq b_{k-1} \mp 2c_{k-1}; \quad c_k \doteq c_{k-1} \pm 4b_{k-1}. \tag{$\ddagger$}$$

Az ($\dagger$) és ($\ddagger$) miatt $a_k, b_k, c_k \in \mathbb{Z}$ valóban egész számok. Azt kell még megmutatnunk, hogy 3 nem osztja a b_k egész számot. Ha $k = 2$, akkor $b_2 = 2 \pm 2$ vagy $b_2 = 2 \mp 0$, tehát b_2 csak 0,2 vagy 4 lehet. Nézzük most a $k > 2$

esetet! Négy esetet fogunk szétválasztani.
1. Ha $w_k = \phi^{\pm 1}\rho^{\pm 1}w_{k-2}$ alakú, akkor (†) miatt $b_k = b_{k-1} \pm 2a_{k-1}$ és (‡) szerint $b_k = b_{k-1} \pm 2 \cdot 3a_{k-2}$. Így az indukciós feltevés szerint 3 nem osztója b_{k-1}-nek ezért nem osztója b_k-nak sem.
2. Ha $w_k = \rho^{\pm 1}\phi^{\pm 1}w_{k-2}$ alakban van felírva, akkor (‡) miatt $b_k = b_{k-1} \mp \mp 2c_{k-1}$ és (†) figyelembevételével $b_k = b_{k-1} \mp 2 \cdot 3c_{k-2}$, ezért 3 ebben az esetben sem osztója b_k-nak.
3. Ha $w_k = \phi^{\pm 1}\phi^{\pm 1}w_{k-2}$ alakú, akkor (†)-et kell kétszer alkalmaznunk, így

$$\begin{aligned} b_k &= b_{k-1} \pm 2a_{k-1} = b_{k-1} \pm 2\left(a_{k-2} \mp 4b_{k-2}\right) = b_{k-1} \pm 2a_{k-2} - 8b_{k-2} = \\ &= b_{k-1} + \left(b_{k-2} \pm 2a_{k-2}\right) - 9b_{k-2} = 2b_{k-1} - 9b_{k-2}. \end{aligned}$$

Itt az utolsó lépésben (†)-et alkalmaztuk harmadszorra is k helyett a $k-1$ számra. Mivel az indukciós feltevés szerint 3 nem osztója b_{k-1}-nek, ezért 3 nem osztója b_k-nak sem.
4. Ha $w_k = \rho^{\pm 1}\rho^{\pm 1}w_{k-2}$ alakban van felírva, akkor most (‡)-et kell kétszer alkalmaznunk, és úgy átalakítani, hogy a (‡) újbóli alkalmazásával az előzőhöz hasonló eset álljon elő.

$$\begin{aligned} b_k &= b_{k-1} \mp 2c_{k-1} = b_{k-1} \mp 2\left(c_{k-2} \pm 4b_{k-2}\right) = b_{k-1} \mp 2c_{k-2} - 8b_{k-2} = \\ &= b_{k-1} + \left(b_{k-2} \mp 2c_{k-2}\right) - 9b_{k-2} = 2b_{k-1} - 9b_{k-2}. \end{aligned}$$

Beláttuk tehát, hogy 3 nem osztója b_k-nak egyik esetben sem. □

Érdekes látni, hogy bizonyításunk egyedül itt lenne hibás, ha azt $\mathbb{R}^2$-re próbálnánk alkalmazni, hiszen a sík egységkörének forgatásai nyilvánvaló módon kommutatív csoportot alkotnak, márpedig az triviális, hogy egy kommutatív csoportnak csak az egy-egy elem generálta részcsoportjai lehetnek szabad csoportok.

10.1.17. állítás (Hausdorff, 1914)**.** *A $\mathbb{R}^3$ gömbfelszínének létezik olyan megszámlálható részhalmaza, amelynek komplementere $SO\,(3)$-paradox.*

Bizonyítás. Legyen G az $SO\,(3)$ két elem által generált szabad részcsoportja, amely 10.1.13. állítás miatt paradox. Világos, hogy a négy jelből kirakható szavak halmaza legfeljebb megszámlálható számosságú és minden forgatás, amely az identitástól különbözik éppen két pontot hagy érintetlenül. Jelölje tehát D az alábbi megszámlálható számosságú halmazt:

$$D \doteq \left\{x \in S^2 : \exists g \in G \; g\,(x) = x \text{ és } g \neq \mathrm{id}\right\}$$

Ha $g \in G$ esetén létezik $x \in S^2 \setminus D$, melyre $g\,(x) = x$, akkor az csak úgy lehetséges, ha $g = \mathrm{id}$ is teljesül. Ha $x \in S^2$ esetén $g\,(x) \in D$, akkor létezik $h \in G, h \neq \mathrm{id}$, amelyre $h\,(g\,(x)) = g\,(x)$, ezért $x = g^{-1}hg\,(x)$, ahol $g^{-1}hg \neq$ $= \mathrm{id}$. Eszerint $x \in D \Rightarrow g\,(x) \in D$ is fennáll minden $g \in G$ esetén, emiatt

$g\left(S^2 \setminus D\right) = S^2 \setminus D$ is teljesül. A G paradox csoport tehát $S^2 \setminus D$-n hat, sőt fixpontmentesen hat. Alkalmazható ezért a 10.1.15. állítás, amely szerint az $S^2 \setminus D$ halmaz G-paradox, tehát $SO\,(3)$-paradox is. □

10.2. Banach–Tarski-paradoxon

Hausdorff a fenti állítás segítségével megmutatta, hogy nincs az S^2 gömbfelület egész hatványhalmazán értelmezett μ forgatás-invariáns végesen additív halmazfüggvény, melyre $\mu\left(S^2\right) = 1$. A bizonyítása szerint egy ilyen μ minden megszámlálható halmazt 0-ba visz, így az állítás nyilvánvaló következménye a fenti Hausdorff-paradoxonnak.

Banach egyrészt megmutatta, hogy a Hausdorff-paradoxonból a megszámlálható halmaz kihagyható, ezért rögtön látszik, hogy ilyen mérték valóban nincsen. Másrészt észrevette, hogy könnyen áttérhetünk B^3 részhalmazainak vizsgálatára, sőt $\mathbb{R}^3$ két olyan részhalmaza, amely egyrészt nem túl nagy (belefér egy gömbbe), másrészt nem túl kicsi (beleírható egy gömb) szintén egymásba darabolható.

10.2.1. Gyenge alak

A Hausdorff-paradoxon megszámlálható halmazát az alábbi konkrét transzformációkkal fogjuk ignorálni:

10.2.1. lemma. · *Legyen $P \subseteq S^1$ egy megszámlálható halmaz és $\tau_\alpha : \mathbb{R}^2 \to \mathbb{R}^2; \tau_\alpha \in SO\,(2)$ az α szögű forgatás. Ekkor megszámlálhatóan sok α kivételével minden $\alpha \in [0,2\pi)$ esetén*

$$\left(\cup_{n=1}^{\infty} \tau_\alpha^n\,(P)\right) \cap P = \varnothing.$$

· *Legyen $P \subseteq S^2$ egy megszámlálható halmaz. Ekkor létezik Q megszámlálható halmaz és $\omega : \mathbb{R}^3 \to \mathbb{R}^3; \omega \in SO\,(3)$ forgatás, hogy*

$$P \subseteq Q \subseteq S^2 \quad \textit{és} \quad \omega\,(Q) = Q \setminus P.$$

· *Létezik $0 \in N \subseteq B^3$ megszámlálható halmaz, és létezik $r : \mathbb{R}^3 \to \mathbb{R}^3; r \in G_3$ egybevágósági transzformáció, amelyre*

$$r\,(N) = N \setminus \{0\}\,.$$

Alkalmazva tehát az a 10.1.10. állítást azt kapjuk, hogy amennyiben P megszámlálható, akkor

$$S^2 \sim_{SO(3)} \left(S^2 \setminus P\right) \quad \textit{valamint} \quad B^3 \sim_{G_3} \left(B^3 \setminus \{0\}\right).$$

Bizonyítás. (1) Valamely $\alpha \in [0{,}2\pi)$ mellett legyen $z_\alpha \doteq e^{i\alpha}$. Ha erre az α-ra és valamely $n \in \mathbb{N}$-re $\tau_\alpha^n(P) \cap P \neq \varnothing$, az azt jelenti, hogy létezik $x, y \in P$, amelyre $z_\alpha^n x = y$, azaz $z_\alpha = \sqrt[n]{\frac{y}{x}}$. Mivel egy komplex számnak csak n darab n-edik gyöke van, ez azt jelenti, hogy adott $(x, y, n) \in P \times P \times \mathbb{N}$-hez, csak véges sok α létezik, amelyre $\tau_\alpha^n(P) \cap P \neq \varnothing$. Megismételve ezt $P \times P \times \mathbb{N}$ minden elemére, azt kapjuk, hogy csak megszámlálhatóan sok olyan $\alpha \in [0{,}2\pi)$ szög van, amelyre $(\cup_{n=1}^{\infty} \tau_\alpha^n(P)) \cap P \neq \varnothing$.

(2) Legyen most $P \subseteq S^2 \subseteq \mathbb{R}^3$ egy megszámlálható halmaz. Válasszunk egy olyan L egyenest, amely S^2-et P-től különböző pontokban metszi. Mivel P csak megszámlálható számosságú, ilyen egyenes biztosan van. Legyen $\omega_\alpha : \mathbb{R}^3 \to \mathbb{R}^3$ ezen L egyenes mint tengely körüli α szögű forgatás. Világos, hogy csak megszámlálható olyan sík van, amely L-re merőleges és tartalmaz P-beli pontot. Erre a legfeljebb megszámlálhatóan sok síkmetszetre alkalmazva a lemma első felében már igazolt állítást, azt kapjuk, hogy megszámlálhatóan sok α szög kivételével tetszőleges α-ra

$$(\cup_{n=1}^{\infty} \omega_\alpha^n(P)) \cap P = \varnothing.$$

Egy ilyen α mellett tekintsük a $Q \doteq P \cup (\cup_{n=1}^{\infty} \omega_\alpha^n(P))$ diszjunkt egyesítést. Világos, hogy $P \subseteq Q \subseteq S^2$; Q megszámlálható számosságú; és $\omega(Q) = \cup_{n=1}^{\infty} \omega_\alpha^n(P) = Q \setminus P$.

(3) Jelölje $u \doteq (1{,}0{,}0) \in S^2$. A lemma második állítását alkalmazva legyen $u \in Q \subseteq S^2$ megszámlálható halmaz, valamint ρ az a forgatás, amelyre

$$\rho(Q) = Q \setminus \{u\}.$$

Legyen $N \doteq \frac{1}{2}Q - \frac{1}{2}u$. Világos, hogy $N \subseteq B$, amelyre $0 \in N$. Definiáljuk az alábbi r egybevágósági transzformációt:

$$r(x) = \rho\left(x + \frac{1}{2}u\right) - \frac{1}{2}u.$$

Ha $x \in N$, akkor $x = \frac{1}{2}q - \frac{1}{2}u$ valamely $q \in Q$-val, így

$$r(x) = \rho\left(\frac{1}{2}q - \frac{1}{2}u + \frac{1}{2}u\right) - \frac{1}{2}u = \frac{1}{2}\rho(q) - \frac{1}{2}u,$$

amiből

$$r(N) = \frac{1}{2}\rho(Q) - \frac{1}{2}u = \frac{1}{2}(Q \setminus \{u\}) - \frac{1}{2}u = N \setminus \{0\}. \qquad \square$$

10.2.2. állítás (Banach–Tarski-paradoxon gyenge alakja)**.** *Az S^2 gömbfelület $SO(3)$-paradox, valamint a B^3 gömb G_3-paradox.*

Bizonyítás. A Hausdorff-paradoxon szerint $S^2 \setminus P$ egy $SO(3)$-paradox halmaza a gömbfelszínnek. Az előző lemma miatt viszont $\left(S^2 \setminus P\right) \sim_{SO(3)} S^2$, ezért a 10.1.14. állítást alkalmazva azt kapjuk, hogy S^2 valóban $SO(3)$-paradox halmaz.

A gömbfelület egy részhalmazához hozzárendelve ugyanezen gömbfelület-részhez tartozó gömbcikket, könnyen látható, hogy $B^3 \setminus \{0\}$ pedig $SO(3)$-paradox részhalmaza, ezért G_3-paradox részhalmaza is a B^3 gömbnek. Az előző lemma miatt viszont $B^3 \setminus \{0\} \sim_{G_3} B^3$, ezért a fentihez hasonlóan a 10.1.14. állítást alkalmazva kapjuk, hogy B^3 valóban egy G_3-paradox halmaz. □

10.2.2. Erős alak

Az erős alak igazolásának egyetlen trükkje, hogy a számosságok elméletéből jól ismert Schröder–Bernstein-ekvivalenciatételt kell élesítenünk bijekciók helyett távolságtartó bijekciókra. Ehhez először bevezetünk egy relációt, amelynek szimmetrikus része az ekvidekompozábilitás relációja lesz.

10.2.3. definíció (ekvidekompozábilis beágyazás)**.** Tegyük fel, hogy G egy az X halmazon ható csoport, $A, B \subseteq X$. Azt mondjuk, hogy A ekvidekompozábilis része B-nek, ha létezik $B' \subseteq B$, amelyre $A \sim_{G_3} B'$. Ezt $A \preceq_G B$ módon jelöljük.

10.2.4. állítás (Banach–Schröder–Bernstein)**.** *Legyen G egy az X halmazon ható csoport és $A, B \subseteq X$. Ekkor $B \preceq_G A$ és $A \preceq_G B$ esetén $A \sim_G B$ is fennáll.*

Bizonyítás. Legyen $A = \cup_{i=1}^n A_i$ diszjunkt felbontás és $g_i \in G$, amelyekre $\cup_{i=1}^n \hat{g}_i(A_i) = B' \subseteq B$ egy partíciója B'-nek. Hasonlóan, $B = \cup_{j=1}^m B_j$ partíció és $f_j \in G$, amelyekre $\cup_{j=1}^m \hat{f}_j(B_j) = A' \subseteq A$ is partíciója A'-nek. Legyenek a $g : A \to B', g(x) \doteq \hat{g}_i(x), x \in A_i$ és az $f : B \to A', f(x) \doteq \hat{f}_j(x), x \in B_j$ módon definiált bijekciók. Tekintsük a

$$\Phi : \mathcal{P}(A) \to \mathcal{P}(A); \qquad \Phi(D) \doteq A \setminus f(B \setminus g(D))$$

halmaz–halmaz leképezést. Világos, hogy $D_1 \subseteq D_2$ esetén $\Phi(D_1) \subseteq \Phi(D_2)$ teljesül. Megmutatjuk, hogy létezik Φ-nek fixpontja. Ehhez tekintsük a

$$\mathcal{H} \doteq \{C \subseteq A : C \subseteq \Phi(C)\}$$

halmazrendszert. Mivel $\varnothing \in \mathcal{H}$, ezért $\mathcal{H}$ nem üres, így legyen $D \doteq \cup\mathcal{H}$ az összes $\mathcal{H}$-beli halmaz egyesítése. Persze $C \in \mathcal{H}$ mellett $C \subseteq D$, ezért $C \subseteq \Phi(C) \subseteq \Phi(D)$, amiből $D \subseteq \Phi(D)$ következik. Így tehát $\Phi(D) \subseteq \subseteq \Phi(\Phi(D))$, ezért $\Phi(D) \in \mathcal{H}$, ergo $\Phi(D) \subseteq D$ is teljesül. Találtunk tehát

$D \subseteq X$ halmazt, amelyre

$$A \smallsetminus D = f\left(B \smallsetminus g\left(D\right)\right) \tag{†}$$

teljesül. Legyen $i = 1, \dots, n$ és $j = 1, \dots, m$ esetén

$$E_i \doteq A_i \cap D; \quad h_i \doteq g_i; \quad E_{n+j} \doteq \hat{f}_j\left(B_j \smallsetminus g\left(D\right)\right); \quad h_{n+j} \doteq f_j^{-1}.$$

Világos, hogy $D = \cup_{i=1}^{n} E_i$ partíciója D-nek; $B \smallsetminus g\left(D\right) = \cup_{j=}^{m}\left(B_j \smallsetminus g\left(D\right)\right)$ partíciója $B \smallsetminus g\left(D\right)$-nek, ezért (†) miatt

$$A \smallsetminus D = \cup_{j=1}^{m} \hat{f}_j\left(B_j \smallsetminus g\left(D\right)\right) = \cup_{j=1}^{m} E_{n+j}$$

partíciója $A \smallsetminus D$-nek. Együtt tehát

$$A = \cup_{k=1}^{n+m} E_k$$

partíciója A-nak.

Mivel $\cup_{i=1}^{n} \hat{g}_i\left(A_i\right) = B' \subseteq B$ partíciója B'-nek, ezért

$$\cup_{k=1}^{n} \hat{h}_k\left(E_k\right) = \cup_{i=1}^{n} \hat{g}_i\left(A_i \cap D\right) = g\left(D\right) \subseteq B'$$

partíciója $g\left(D\right)$-nek. Világos a $\hat{h}_{n+j}\left(E_{n+j}\right) = B_j \smallsetminus g\left(D\right)$ azonosság, emiatt pedig $\cup_{k=n+1}^{n+m} \hat{h}_k\left(E_k\right)$ partíciója $B \smallsetminus g\left(D\right)$-nek. Azt kaptuk tehát, hogy

$$\cup_{k=1}^{n+m} \hat{h}_k\left(E_k\right)$$

partíciója B-nek. Ezt kellett belátni. □

Első következményként egyszerűsíthetjük a paradox halmaz definícióját.

10.2.5. állítás (paradox halmaz felbontásáról)**.** *Legyen G egy az X halmazon ható csoport. Legyen $C \subseteq A$ olyan halmaz, amelyre $C \sim_G A$. Ekkor tetszőleges $C \subseteq Q \subseteq A$ esetén $Q \sim_G A$ is fennáll.*

Az $A \subseteq X$ halmaz pontosan akkor G-paradox, ha A előáll $A = B \cup C$ diszjunkt egyesítés alakban, ahol B és C is G-ekvidekompozábilisek A-val.

Bizonyítás. Világos, hogy a tartalmazások miatt $C \preceq_G Q \preceq_G A$. No de $A \preceq$ $\preceq_G C$ is fennáll, ezért a Banach–Schröder–Bernstein-ekvivalenciatétel miatt $Q \sim_G A$ is teljesül. A második állítás már az első következménye. □

10.2.6. állítás (Diszjunkt duplázás)**.** *Legyen $A \subseteq \mathbb{R}^3$ egy zárt gömb és A_1 ennek olyan eltoltja, hogy $A \cap A_1 = \varnothing$. Ekkor $A \sim_{G_3} A \cup A_1$.*

Bizonyítás. A gyenge alak és az előző állítás szerint az A gömbnek van olyan $A = B \cup C$ diszjunkt előállítása, ahol B és a C részhalmaz is G_3-ekvidekompozábilis A-val. Ekkor persze C is G_3-ekvidekompozábilis A_1-gyel. Mivel A_1 és A diszjunktak, ezért $A = B \cup C \sim_{G_3} A \cup A_1$. □

10.2.7. állítás (n-szerezés)**.** *Legyen $A \subseteq \mathbb{R}^3$ egy zárt gömb, és $A_1, \dots, A_n$ ennek véges számú eltoltja. Ekkor*

$$A \sim_{G_3} \cup_{j=1}^{n} A_j$$

Bizonyítás. Persze n szerinti teljes indukció, és az $n = 1$ eset csak azt állítja, hogy $A \sim_{G_3} A$. Tegyük fel, hogy, hogy $A \sim_{G_3} \cup_{j=1}^{n-1} A_j$ és lássuk be n-re. Világos, hogy $A_n \setminus \cup_{j=1}^{n-1} A_j$ az A_n részhalmazaként ekvidekompozábilis A egy részhalmazával, tehát $A_n \setminus \cup_{j=1}^{n-1} A_j \preceq_{G_3} B$, ahol B az A-nak olyan eltolja, melyre $A \cap B = \varnothing$. Ezért

$$\cup_{j=1}^{n} A_j = \left(\cup_{j=1}^{n-1} A_j\right) \cup \left(A_n \setminus \cup_{j=1}^{n-1} A_j\right) \preceq_{G_3} A \cup B \sim_{G_3} A.$$

No de $A \preceq \cup_{j=1}^{n} A_j$ triviális, ezért az ekvivalenciatétel miatt $A \sim_{G_3} \cup_{j=1}^{n} A_j$ is valóban teljesül. □

10.2.8. állítás (Banach–Tarski-paradoxon erős alakja)**.** *Legyen Q és T az $\mathbb{R}^3$ egy-egy korlátos részhalmaza, amelyeknek belseje sem üres. Ekkor $Q \sim_{G_3} T$, azaz a Q és a T halmazok G_3-ekvidekompozábilisek egymással.*

Bizonyítás. Legyen A egy olyan zárt gömb, melyre $A \subseteq Q$ és melynek bizonyos B eltoltjára $B \subseteq T$. Mivel Q teljesen korlátos, ezért vannak A-nak olyan $A_1, \dots, A_n$ eltoltjai, amelyekre $Q \subseteq \cup_{j=1}^{n} A_j$. Világos, hogy a tartalmazások és az előző tétel szerint

$$A \preceq_{G_3} Q \preceq_{G_3} \cup_{j=1}^{n} A_j \sim_{G_3} A,$$

így újra az ekvivalenciatételt használva azt kapjuk, hogy $A \sim_{G_3} Q$. Ugyanezt kapjuk A-nak a bizonyítás elején definiált B eltoltjára és T-re, azaz $B \sim_{G_3} T$. Összevetve: $Q \sim_{G_3} A \sim_{G_3} B \sim_{G_3} T$ valóban teljesül. Ezt kellett belátni. □

„Csak öt kenyerünk van és két halunk”
Mt. 14.17

11. fejezet

Berge maximumtétele

A Berge-tétel azt állítja, hogy egy jól viselkedő feltételes szélsőérték feladatnak a megoldása a feltételt leíró paraméterek folytonos függvénye, és az optimumot létrehozó pontok halmaza folytonosan függ a paraméterektől.

E gondolat kicsit pontosabb megfogalmazása a következő. Adott x paraméter állapot esetén a feltételi halmazt $\Gamma(x)$-szel jelöljük. Az optimalizálás e feltételi halmaz felett folyik. A célfüggvény is függhet a feladat paramétereitől, ami azt jelenti, hogy adott x mellett keressük a

$$\sup_{y \in \Gamma(x)} f(x, y)$$

optimumot. Persze más x-re ez várhatóan más lesz. Az optimum értékét $m(x)$-szel jelölve a feladat megoldásának értékét kapjuk a paraméterek függvényében. Az optimális pontok halmazát $\mu(x)$-szel jelöljük. Olyan feltételeket keresünk, amelyek biztosítják a paraméterekhez e maximalizáló elemek $\mu(x)$ halmazát rendelő hozzárendelés és az m értékfüggvény valamilyen értelemben vett folytonosságát.

Az eredeti forrás [Berge (1959)], ennek ma elérhető verziója [Berge (1997)]. Egy modern és nagyon szép összefoglaló található [Aliprantis–Border (2006), chap. 17.5]-ben. E fejezet is elsősorban ez utóbbin alapul.

Először is tisztázzuk a halmazértékű függvények legfontosabb folytonossági koncepcióit és azok kapcsolatát. A második szakaszban fokozatosan igazoljuk a maximumtételt a különböző folytonossági feltevések mellett. A fejezetben a könyvtől eltérően kicsit erősebb topológiai ismereteket tételezünk fel. Feltesszük, hogy az olvasó ismeri az általánosított sorozat fogalmát[1], és azok ele-

[1] Az általánosított sorozat használata elsősorban ott problémás és kétséges, mikor a részsorozatok fogalma is felmerül. A probléma, hogy néhány egymástól lényegesen különböző részsorozat-fogalom is természetesen merül fel. Itt mi részsorozatot nem fogunk használni.

mi topológiai vonatkozásait, például, hogy az általánosított sorozatok konvergenciájával a folytonosság és a zártság is jellemezhető. Ha az olvasót zavarja a fejezet általános fogalmazása, akkor nem sokat veszít azzal, ha az összes állítást metrikus térre, vagy akár véges dimenziós euklideszi terek esetében gondolja meg, ekkor persze általánosított sorozaton a szokásos sorozatot kell értenie. Mivel ezen egyszerűsítő feltételek a gondolatok indoklását érdemben nem változtatják, inkább az általános alakban fogalmazzuk meg a tételeket.

Legelső lépésként definiáljuk a négy szóba jövő folytonossági koncepciót:

11.0.9. definíció. Az X és Y topologikus terek közti $\Gamma : X \multimap Y$ halmazértékű hozzárendelést

felülről félig folytonosnak nevezünk egy x_0 pontban, ha minden olyan V nyílt halmazhoz, melyre $\Gamma(x_0) \subseteq V$ létezik olyan U nyílt környezete az x_0 pontnak, hogy minden $x \in U$ esetén $\Gamma(x) \subseteq U$ is fennáll;

alulról félig folytonosnak nevezünk egy x_0 pontban, ha minden olyan V nyílt halmazhoz, melyre $\Gamma(x_0) \cap V \neq \varnothing$ létezik olyan U nyílt környezete az x_0 pontnak, hogy minden $x \in U$ esetén $\Gamma(x) \cap U \neq \varnothing$ is fennáll;

folytonosnak mondunk az x_0 pontban, ha az alulról félig folytonos és egyben felülről félig folytonos leképezés is x_0-ban;

zárt leképezésnek nevezünk, ha a

$$\operatorname{graph}\Gamma = \{(x, y) \in X \times Y : y \in \Gamma(x)\}$$

gráfja az $X \times Y$ szorzattér zárt részhalmaza.

A szokásoknak megfelelően a leképezés felülről, alulról félig folytonosnak vagy folytonosnak nevezünk, ha az az értelmezési tartománya valamennyi pontjában felülről, alulról félig folytonos vagy folytonos.

Zárt leképezésnek lenni egy globális koncepció, nincs ennek megfelelő lokális fogalom.

11.0.10.

Egy halmaz pontosan akkor zárt, ha minden benne lévő konvergens általánosított sorozat határértéke is a halmazhoz tartozik. Speciálisan ez azt jelenti, hogy a Γ pontosan akkor zárt leképezés, ha

$$\forall\alpha\,(x_\alpha, y_\alpha) \in \operatorname{graph}\Gamma, \lim_\alpha (x_\alpha, y_\alpha) = (x, y) \implies (x, y) \in \operatorname{graph}\Gamma$$

implikáció fennáll. Mivel az (x_α, y_α) pár konvergenciája külön az x_α és külön az y_α általánosított sorozatok konvergenciáját jelenti, ezért a zárt leképezés egy ekvivalens átfogalmazása a következő: Minden olyan (x_α, y_α) párra, amelyre $y_\alpha \in \Gamma(x_\alpha)$, $\lim_\alpha x_\alpha = x$, $\lim_\alpha y_\alpha = y$ fennállnak, az $y \in \Gamma(x)$ tartalmazás is teljesül.

11.1. A zártgráf-tétel erősítése

A felülről félig folytonos és a zárt leképezések kapcsolatáról a legegyszerűbb állítás az úgynevezett zártgráf-tétel.

11.1.1. állítás (zártgráf-tétel)**.** *Legyen az X topologikus tér és az Y egy kompakt, T_2 topologikus tér. A $\varphi : X \multimap Y$ halmazértékű leképezés, amelyre a $\varphi(x)$ képhalmazok zártak. Ekkor φ zárt leképezés mivolta és felülről félig folytonossága egymással ekvivalens feltevések.*

A tételt ebben a formájában csak a szakasz végén igazoljuk. Ennek oka, hogy a zártgráf-tétel ebben az alakjában gyengébb annál, minthogy céljainknak megfelelne. Mind a szükségességet mind az elégségességet tekintve a fentinél kicsit erősebb alakra van szükségünk. A zártgráf-tételt nem fogjuk használni, de a szakasz végén megmutatjuk, hogy a következő két állítás valóban erősítései a zártgráf-tétel szükségességi és elegendőségi állításainak.

11.1.2. állítás. *Legyen a $\varphi : X \multimap Y$ halmazértékű leképezés, és tegyük fel, hogy az Y topologikus tér T_2 tulajdonságú, azaz bármely két különböző pontnak vannak diszjunkt környezetei.*

Ha φ egy felülről félig folytonos leképezés, melynek értékei kompakt halmazok, akkor φ zárt leképezés is.

Bizonyítás. Legyen $(x_\alpha, y_\alpha) \to (x, y)$ konvergens általánosított sorozat, melyre minden α mellett $(x_\alpha, y_\alpha) \in \operatorname{graph} \varphi$. Meg kell mutatnunk, hogy $(x, y) \in$ $\in \operatorname{graph} \varphi$, azaz $y \in \varphi(x)$ is teljesül. Tegyük fel indirekt, hogy $y \notin \varphi(x)$. Mivel $\varphi(x)$ kompakt és az Y tér T_2, ezért létezik V és W nyílt halmaz, amelyekre

$$y \in V, \varphi(x) \subseteq W, V \cap W = \varnothing.^2$$

Kihasználva φ felülről félig folytonosságát, találunk olyan $U \in \tau(x)$ környezetet, melyre $\varphi(z) \subseteq W, \forall z \in U$ fennáll. Mivel $x_\alpha \to x$ ezért elég nagy α mellett $x_\alpha \in U$, tehát $y_\alpha \in \varphi(x_\alpha) \subseteq W$. Más oldalról $y_\alpha \to y \in V$ is fennáll, ezért elég nagy α mellett $y_\alpha \in V \cap W = \varnothing$, ami ellentmondás. □

11.1.3. állítás. *Legyenek $\varphi, \psi : X \multimap Y$ halmazértékű leképezések, $x_0 \in X$ egy rögzített pont. Tegyük fel, hogy a $\psi(x_0)$ halmaz kompakt, φ zárt leképezés, és ψ felülről félig folytonos x_0-ban. Ekkor a $\varphi \cap \psi$ leképezés is felülről félig folytonos az x_0 pontban.*

Bizonyítás. Legyen a W nyílt halmazra $\varphi(x_0) \cap \psi(x_0) \subseteq W$. Keresünk $N \in$ $\in \tau(x_0)$ nyílt környezetet, amelyre minden $x \in N$ mellett $\varphi(x) \cap \psi(x) \subseteq W$.

[2] $\forall z \in \varphi(x)$-re a T_2 tulajdonság szerint $\exists V_z, W_z$ diszjunkt nyílt halmaz, amelyekre $z \in W_z, y \in V_z$. A $\varphi(x)$ kompaktsága szerint legyen $\{W_{z_1}, \ldots, W_{z_N}\}$ véges lefedés. Ekkor $W = \cup_{k=1}^{N} W_{z_k}$ és $V = \cap_{k=1}^{N} V_{z_k}$.

Ha $\psi(x_0) \subseteq W$, akkor alkalmazva ψ felülről folytonosságát, létezik $N \in \in \tau(x_0)$, hogy minden $x \in N$ mellett

$$\varphi(x) \cap \psi(x) \subseteq \psi(x) \subseteq W.$$

Egyébként, tekintsük a $K = \psi(x_0) \setminus W$ nem üres, kompakt halmazt. Világos, hogy $K \cap \varphi(x_0) = \varphi(x_0) \cap \psi(x_0) \cap W^c = \varnothing$. Így tehát minden $y \in \in K$ mellett $(x_0, y) \notin \operatorname{graph}\varphi$. Felhasználva, hogy φ grafikonja zárt halmaz, minden $y \in K$, hoz létezik $U_y \in \tau(x_0)$ környezet és $V_y \in \tau(y)$ környezet, hogy

$$(U_y \times V_y) \cap \operatorname{graph}\varphi = \varnothing.$$

Persze az $\{V_y : y \in K\}$ rendszer fedése a K kompakt halmaznak, ezért létezik véges $K \subseteq \cup_{k=1}^{N} V_{y_k}$ lefedés. Jelölje $V = \cup_{k=1}^{N} V_{y_k}$ nyílt halmazt. Persze

$$\psi(x_0) \subseteq K \cup W \subseteq V \cup W.$$

A ψ függvény felülről félig folytonossága szerint létezik $U \in \tau(x_0)$ nyílt környezet, hogy minden $z \in U$ mellett

$$\psi(z) \subseteq V \cup W.$$

Ha $z \in \cap_{k=1}^{N} U_{y_k}$, akkor $(\cap_{k=1}^{N} U_{y_k} \times V) \cap \operatorname{graph}\varphi = \varnothing$ szerint

$$\varphi(z) \cap V = \varnothing$$

így a két utóbb kiemelt sort összevetve azt kapjuk, hogy minden $z \in U \cap \cap \left(\cap_{k=1}^{N} U_{y_k}\right)$ esetén $\varphi(z) \cap \psi(z) \subseteq W$. $\square$

Csak a fenti két állításra lesz szükségünk, de a teljesség kedvéért nézzük hogyan következik ezekből a zártgráf-tétel:

a zártgráf-tétel igazolása. A szükségességet és az elegendőséget külön-külön indokoljuk:

1. Tegyük fel először, hogy φ felülről félig folytonos. Mivel $\varphi(x)$ képhalmaz zárt, és Y T_2 tulajdonságú, ezért $\varphi(x)$ a kompakt Y halmaz zárt részhalmazaként maga is kompakt. Alkalmazható tehát a 11.1.2. állítás, ergo φ valóban egy zárt leképezés.
2. Most azt tegyük fel, hogy φ zárt. Legyen $\psi : X \multimap Y$ az a konstans függvény, amelyre minden $x \in X$ mellett $\psi(x) = Y$. Világos, hogy φ-re és ψ-re fennállnak a 11.1.3. állítás feltételei, hiszen a konstans függvény triviálisan felülről félig folytonos. Így a $\varphi = \varphi \cap \psi$ függvény is felülről félig folytonos. $\square$

A metszet leképezés folytonossága az optimalizáció elméletben kulcsfontosságú kérdés. Ennek oka, hogy a feltételi halmaznak valamilyen paramétertől való függését írjuk le halmazértékű leképezések megadásával. Nincs természetesebb annál, mint a feltételi halmazt valamely korlátozó feltételrendszer együttesének teljesülésével megadni, ergo halmazértékű leképezések metszeteként. A bizonyított 11.1.3. állítás egy jól használható elegendő feltételt ad a metszet leképezés felülről félig folytonosságára. Az alulról félig folytonosság ennél sokkal nehezebb kérdés. A 11.1.3. állításhoz hasonló környezetben a metszet leképezés alulról félig folytonossága nem is teljesül, ennél finomabb feltételekre van szükség. Az érdeklődő olvasó egy bevezetést kap a metszet leképezés alulról félig folytonosságának problémakörébe [Dancs et al. (2011)]-ben.

11.2. A maximumtétel

Az értékfüggvény folytonossága pusztán a következő két állítás következménye. Ez a Berge-tétel egyszerűbb része. A maximalizáló elemek alkotta halmaz paraméterektől függésének felülről félig folytonosságához kell minden, amit eddig átgondoltunk.

11.2.1. állítás (alulról félig folytonos eset)**.** *Legyenek (X,τ) és (Y,τ) topologikus terek, $\Gamma : X \multimap Y$ halmazértékű függvény, amelyre minden $x \in X$ mellett $\Gamma(x) \neq \varnothing$ részhalmaza Y-nak, továbbá $f : \operatorname{graph}\Gamma \to \mathbb{R}$ függvény. Ha f alulról félig folytonos függvény és Γ alulról félig folytonos halmazértékű leképezés, akkor az*

$$m(x) = \sup\{f(x,y) : y \in \Gamma(x)\}$$

értékfüggvény is alulról félig folytonos.

Bizonyítás. Legyen $x_0 \in X$ rögzített, melyre $f(x_0) > \alpha$. Ekkor létezik $y \in \in \Gamma(x_0)$, melyre $f(x_0, y) > \alpha$. Az f alulról félig folytonosságát használva kapunk $U \in \tau(x_0)$ és $V \in \tau(y)$ környezetet, hogy

$$(U \times V) \cap \operatorname{graph}(\Gamma) \subseteq \{(u,v) \in \operatorname{graph}\Gamma : f(u,v) > \alpha\}. \tag{11.1}$$

Világos, hogy $\Gamma(x_0) \cap V \neq \varnothing$. A Γ alulról félig folytonosságát használva kapunk olyan $R \in \tau(x_0)$ nyílt környezetet, hogy minden $x \in R$ mellett $\Gamma(x) \cap \cap V \neq \varnothing$. Így ha $x \in R \cap U$ tetszőleges eleme f értelmezési tartományának, akkor létezik $y \in V$, amelyre $(x,y) \in \operatorname{graph}\Gamma$. Így a fent kiemelt (11.1) szerint $m(x) \geq f(x,y) > \alpha$ teljesül az x_0 pont $R \cap U$ nyílt környezetén. Ezt kellett belátni. □

11.2.2. állítás (felülről félig folytonos eset)**.** *Legyenek (X,τ) és (Y,τ) topologikus terek, $\Gamma : X \multimap Y$ halmazértékű függvény, amelyre minden $x \in X$*

mellett $\Gamma(x) \neq \varnothing$ *és kompakt részhalmaza* Y*-nak, továbbá* $f : \operatorname{graph}\Gamma \to \mathbb{R}$ *függvény. Ha* f *felülről félig folytonos függvény és* Γ *felülről félig folytonos halmazértékű leképezés, akkor az*

$$m(x) = \sup\{f(x,y) : y \in \Gamma(x)\}$$

értékfüggvény is felülről félig folytonos.

Bizonyítás. Legyen $x_0 \in X$ pont, melyre $m(x_0) < \alpha$. Ez ezt jelenti, hogy minden $y \in \Gamma(x_0)$ mellett $f(x_0, y) < \alpha$. Mivel f felülről félig folytonos a $\operatorname{graph}\Gamma$ halmazon, ezért minden $y \in \Gamma(x_0)$-hoz létezik $U_y \in \tau(x_0)$ és $V_y \in \in \tau(y)$ az X, illetve az Y térben nyílt halmaz, amelyre

$$(U_y \times V_y) \cap \operatorname{graph}\Gamma \subseteq \{(u,v) \in \operatorname{graph}\Gamma : f(u,v) < \alpha\}. \tag{11.2}$$

Világos, hogy $\Gamma(x_0) \subseteq \cup_{y\in\Gamma(x_0)} V_y$. A $\Gamma(x_0) \subseteq Y$ kompaktsága szerint, létezik véges sok $y_1, \ldots, y_N \in \Gamma(x_0)$, melyekre a $\Gamma(x_0) \subseteq \cup_{k=1}^N V_{y_k}$ véges lefedés is fennáll. Válasszunk most a Γ felülről félig folytonossága szerint olyan $U \in \in \tau(x_0)$ nyílt környezetet, amelyre minden $x \in U$ esetén $\Gamma(x) \subseteq \cup_{k=1}^N V_{y_k}$. Amennyiben $x \in U \cap \left(\cap_{k=1}^N U_{y_k}\right)$ és $(x,y) \in \operatorname{graph}\Gamma$, akkor létezik k index, melyre $y \in V_{y_k}$, így persze erre a k-ra

$$(x,y) \in U_{y_k} \times V_{y_k},$$

ezért $f(x,y) < \alpha$ is fennáll (11.2) szerint. Mivel a felülről félig folytonos $f(x,\cdot)$ függvény felveszi maximumát a $\Gamma(x)$ kompakt halmazon, ezért minden $x \in U \cap \left(\cap_{k=1}^N U_{y_i}\right)$ mellett $m(x) < \alpha$ teljesül. Ezt kellett belátni. □

Minden a rendelkezésünkre áll a közgazdasági irodalomban sokszor használatos maximumtétel összefoglalásához.

11.2.3. tétel (folytonos eset, Berge)**.** *Legyenek* (X,τ) *és* (Y,τ) *topologikus terek,* $\Gamma : X \multimap Y$ *halmazértékű függvény, amelyre minden* $x \in X$ *mellett* $\Gamma(x) \neq \varnothing$ *és kompakt részhalmaza* Y*-nak, továbbá* $f : \operatorname{graph}\Gamma \to \mathbb{R}$ *függvény. Ha* f *folytonos függvény és* Γ *folytonos halmazértékű leképezés, akkor az*

$$m(x) = \sup\{f(x,y) : y \in \Gamma(x)\}$$

$m : X \to \mathbb{R}$ *értékfüggvény is folytonos.*

Ha feltesszük még,

1. *hogy az* Y *topologikus tér* T_2 *tulajdonságú, vagy azt,*
2. *hogy az* f *célfüggvény* $\operatorname{graph}\Gamma$ *értelmezési tartományáról a teljes* $X\times Y$*-ra folytonosan kiterjeszthető,*

akkor a

$$\mu(x) = \{y \in \Gamma(x) : f(x, y) = m(x)\}^{3}$$

$\mu : X \multimap Y$ halmazértékű leképezés felülről félig folytonos.

Bizonyítás. A már igazolt alulról félig folytonos (11.2.1) és felülről félig folytonos (11.2.2) eset együtt az m értékfüggvény folytonosságát igazolja.

Az Y topologikus tér T_2 tulajdonsága esetében megmutatjuk, hogy μ egy zárt leképezés. Legyen tehát $(x_\alpha, y_\alpha) \to (x, y)$ egy konvergens általánosított sorozat, melyre minden α mellett $y_\alpha \in \mu(x_\alpha)$. Először is vegyük észre, lásd a 11.1.2. állítást, hogy Γ a felülről félig folytonosság miatt egy zárt leképezés, ezért $y \in \Gamma(x)$ fennáll. Másodszor, az f függvény $(x, y) \in \operatorname{graph}\Gamma$ pontbeli, és az m értékfüggvény x-beli folytonossága szerint

$$f(x, y) = \lim_\alpha f(x_\alpha, y_\alpha) = \lim_\alpha m(x_\alpha) = m(x). \tag{11.3}$$

Azt kaptuk tehát, hogy y kielégíti μ definícióját, ergo $(x, y) \in \operatorname{graph}\mu$, tehát μ valóban egy zárt leképezés. Nyilván $\mu(x) \subseteq \Gamma(x)$, Γ kompakt értékű, felülről félig folytonos leképezés és még azt is megmutattuk, hogy μ zárt. Láttuk a 11.1.3. állításban, hogy ekkor a

$$\mu(x) = \mu(x) \cap \Gamma(x)$$

metszet leképezés is egy felülről félig folytonos függvény.

Most tegyük fel, hogy f folytonosan kiterjeszthető a teljes szorzatra. Jelölje a kiterjesztett függvényt $F : X \times Y \to \mathbb{R}$. Legyen

$$\varphi(x) = \{y \in Y : F(x, y) = m(x)\}$$

Világos, hogy $\varphi : X \multimap Y$ egy zárt leképezés, hiszen ha $(x_\alpha, y_\alpha) \to (x, y)$, amelyre $y_\alpha \in \varphi(x_\alpha)$ akkor minden további feltevés nélkül $F(x, y)$ értelmes, így a kiemelt (11.3) egyenlőséget megismételhetjük f helyett F-fel. Kapjuk tehát, hogy $y \in \varphi(x)$, azaz φ valóban zárt leképezés. Persze minden $x \in X$ mellett

$$\mu(x) = \varphi(x) \cap \Gamma(x),$$

ezért most is alkalmazva a metszet felülről félig folytonosságát biztosító 11.1.3. állítást, már a μ felülről félig folytonossága is adódik. □

[3] A maximalizáló μ leképezést néha a rövidebb $\mu(x) = \operatorname{argmax}\{f(x, y) : y \in \Gamma(x)\}$ módon jelölik.

Irodalomjegyzék

[Aliprantis–Border (2006)] Aliprantis, Charalambos D. and Border, Kim C., *Infinite dimensional analysis.* Springer, Berlin, third edition, 2006. ISBN 978-3-540-32696-0; 3-540-32696-0. A hitchhiker's guide.

[Berge (1959)] Berge, Claude, *Espaces topologiques: Fonctions multivoques.* Collection Universitaire de Mathématiques, Vol. III. Dunod, Paris, 1959.

[Berge (1997)] Berge, Claude, *Topological spaces*, Including a treatment of multi-valued functions, vector spaces and convexity. Translated from the French original by E. M. Patterson, Reprint of the 1963 translation. Dover Publications Inc., Mineola, NY, 1997.

[Browder (1996)] Browder, Andrew, *Mathematical analysis.* An introduction. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1996. ISBN = 0-387-94614-4. DOI: 10.1007/978-1-4612-0715-3. URL `http://dx.doi.org/10.1007/978-1-4612-0715-3`.

[Dancs (1992)] Dancs István, *Bevezetés a matematikai analízisbe I, II.* AULA, Budapest, 1992.

[Dancs–Puskás (2001)] Dancs István és Puskás Csaba, *Vektorterek.* AULA, Budapest, 2001. ISBN 963-9345-53-9.

[Dancs et al. (2011)] Dancs, I., Medvegyev, P. and Magyarkuti, Gy., Normability via the convergence of closed and convex sets, *J. Optim. Theory Appl.* **150** (3) (2011), 675–682. ISSN 0022-3239. DOI: 10.1007/s10957-011-9835-1. URL `http://dx.doi.org/10.1007/s10957-011-9835-1`.

[Folland (1999)] Folland, Gerald B., *Real analysis.* Pure and Applied Mathematics (New York). Modern techniques and their applications. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York, second edition. 1999. ISBN 0-471-31716-0.

[Irigoyen et al. (2002)] Irigoyen, C., Rossi-Hansberg, E. and Wright, M. L. J., *Solutions manual for Recursive methods in economic dynamics*, Harvard University Press, 2002. ISBN 9780674008885. URL `http://books.google.hu/books?id=u0e7AAAAIAAJ`.

[Pál–Sali (2009)] Pál Jenő és Sali András, *Mértékelmélet konzultáció*. Konzultáció jegyzet. 2009. URL `https://sites.google.com/site/mertekkonz09/feladatok`,

[Rudin (1987)] Rudin, Walter, *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987. ISBN 0-07-054234-1.

[Schilling (2005)] Schilling, René L., *Measures, integrals and martingales*. Cambridge University Press, New York, 2005. ISBN 978-0-521-61525-9; 0-521-61525-9. DOI 10.1017/CBO9780511810886. URL `http://dx.doi.org/10.1017/CBO9780511810886`.

[Stokey–Lucas (1989)] Stokey, Nancy L. and Lucas Jr., Robert E., *Recursive methods in economic dynamics*. With the collaboration of Edward C. Prescott. Harvard University Press, Cambridge, MA, 1989. ISBN 0-674-75096-9.

[Stromberg (1979)] Stromberg, Karl, The Banach–Tarski Paradox. *Math. Monthly* **86** (3) (1979), 151–161.

[Su (1990)] Su, Francis Edward, *The Banach–Tarski Paradox*. Ph.D. Minor Thesis. Harvard University, 1990. URL `http://www.math.hmc.edu/~su/papers.dir/banachtarski.pdf`.

[Wagon (1985)] Wagon, Stan, *The Banach–Tarski Paradox*. Encycl. of Math. and its Appl. vol. 24, Cambridge University Press, Cambridge, 1985.

Tárgymutató