Números p-ádicos

Instituto de Matemática da Universidade Federal de Alagoas
Outono 2019

Herivelto Martins Borges Filho

e

Enno Nagel

Sumário

Introdução

Seja p um número primo. Os números p-ádicos, assim como os números reais, são limites de sequências dos números racionais; com a diferença que tais limites são tomados com base na norma p-ádica, ao invés da norma usual em $\mathbb{R}$.

Introduzidos há cerca de cem anos na Aritmética, vêm à tona recentemente mais e mais aplicações analíticas, por exemplo, na Geometria Diferencial ([GS92]) ou nos Sistemas Dinâmicos ([RL03]).

Definição Provisória

Mais precisamente, a norma *p-ádica* de um número inteiro a é definida por

$$|a|_p := p^{-v}, \text{ onde } p^v \text{ é a maior potência de } p \text{ que divide } a;$$

isto é, a norma p-ádica $|\cdot|_p$ mede quantas vezes p divide a. Em contraste com a norma usual, observe que a norma p-ádica de p^n *diminui* quando o expoente n *cresce*. Vejamos alguns exemplos:

- Para $p = 2$ temos $|12|_2 = |2^2 3|_2 = 2^{-2}$, e
- para $p = 3$ temos $|12|_3 = |3^1 \cdot 2^2|_3 = 3^{-1}$.

Em analogia à expansão decimal dos números reais, por exemplo,

$$\sqrt{2} = 1,414\ldots = 1 + 4 \cdot 10^{-1} + 1 \cdot 10^{-2} + 4 \cdot 10^{-3} \ldots$$

todo número p-ádico a tem uma expansão p-ádica,

$$a = a_{-\mathrm{N}} p^{-\mathrm{N}} + \cdots + a_0 + a_1 p^1 + \cdots, \qquad \text{com } a_i \in \{0, \ldots, p-1\}$$

que converge em relação à norma p-ádica. Ao compararmos, trocamos

- as potências de 10^{-1} pelas de p, e
- os algarismos decimais $0, \ldots, 9$ pelos algarismos $0, \ldots, p-1$, e
- a norma da convergência pela norma p-ádica.

Por exemplo, para $p = 2$, obtemos uma expansão binária infinita

$$10011\ldots = 2^0 + 2^3 + 2^4 + \cdots .$$

A contra-intuição que quanto maior a potência de p, menor a sua norma p-ádica justifica-se na resolução de equações polinomiais módulo p^n: Por exemplo, para $p = 5$, são zeros p-adicamente mais e mais próximos, isto é, módulo p, p^2, p^3, p^4 ... de $X^2 + 1 \equiv 0$ os números inteiros

$$4, 4 + 1 \cdot 5, 4 + 1 \cdot 5 + 3 \cdot 5^2, 4 + 1 \cdot 5 + 3 \cdot 5^2 + 1 \cdot 5^3 + \cdots$$

que podemos encontrar iterativamente:

1. Para $X^2 \equiv -1 \bmod 5$, existem as duas soluções $x \equiv \pm 2 \bmod 5$.

2. Para $X^2 \equiv -1 \bmod 5^2$, uma solução x satisfaz em particular $X^2 \equiv -1 \bmod 5$, logo, $x \equiv \pm 2 \bmod 5$. Optemos pela solução $x \equiv 2 \bmod 5$. Logo, buscamos $x = 2 + x_1 \cdot 5$ tal que

   $$x^2 + 1 \equiv 5 + 20x_1 \equiv 5^2,$$

   logo 5^2 divide $5 + 20x_1$, se, e somente se, 5 divide $1 + 4x_1$, logo $x_1 \equiv 1 \bmod 5$. Obtemos $x = 2 + 1 \cdot 5$ como solução de $X^2 \equiv -1 \bmod 5^2$.

3. Para $X^2 \equiv -1 \bmod 5^n$ para $n > 2$, procedamos por indução: Suponha que tenhamos obtido uma solução $\tilde{x}$ para $\tilde{x}^2 \equiv -1 \bmod 5^n$. A solução x módulo 5^{n+1} terá a forma $x = \tilde{x} + x_n 5^n$. Como $\tilde{x}^2 \equiv -1 \bmod 5^n$, isto é, $\tilde{x}^2 + 1 = 5^n a$ para algum inteiro a, obtemos

   $$0 \equiv x^2 + 1 \equiv \tilde{x}^2 + 1 + 2\tilde{x}x_n 5^n \equiv 5^n(a + 2\tilde{x}x_n) \bmod 5^{n+1}$$

   Logo $a + 2\tilde{x}x_n \equiv 0 \bmod 5$. Como $\tilde{x}^2 \equiv 4 \bmod 5$, o primo 5 não divide $\tilde{x}$, e esta equação tem uma única solução.

(Em geral, é o Lema de Hensel que garante que a resolubilidade desta equação módulo p implique a p-ádica.) A solução p-ádica

$$4 + 1 \cdot 5 + 3 \cdot 5^2 + 1 \cdot 5^3 + \cdots$$

assim obtida converge p-adicamente à solução $X^2 + 1 = 0$.

Observamos que para x em $\mathbb{Z}$, vale $x = 0$ se, e somente se, $x \equiv 0 \bmod n$ para qualquer inteiro n, ou, equivalentemente, pelo Teorema Chinês dos Restos se, e somente se,

$$x \equiv 0 \bmod p^n \quad \text{para qualquer primo } p \text{ e expoente } n.$$

Logo, para P em $\mathbb{Z}[X]$, quanto maior o expoente n para o qual $P(x) \equiv 0 \bmod p^n$, mais próximo está x de um zero inteiro.

Estas sequências assim obtidas formam um anel, o anel dos inteiros p-ádicos $\mathbb{Z}_p$. Acabamos de calcular uma solução p-ádica de $X^2+1 = 0$ para $p = 5$. Qual é a expansão p-ádica de -1, a solução de $X+1 = 0$? Esta solução já era conhecida por Euler, mas ainda não tinha um sentido matematicamente rigoroso: Temos

$$-1 = (p-1) + (p-1)\cdot p + (p-1)\cdot p^2 + \cdots \tag{0.1}$$

De fato, a série geométrica

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots$$

para $x = p$ resulta em

$$\frac{1}{1-p} = 1 + p + p^2 + \cdots$$

e multiplicação por $p-1$ leva-nos a (0.1). Enquanto no mundo real o significado da série de (0.1) permanece obscuro, no mundo p-ádico converge perfeitamente.

Conteúdo

Após a introdução dos números p-ádicos, compararemos as noções fundamentais da Análise real e p-ádica, as suas analogias e diferenças. Uma delas é a condição da diferenciabilidade, que para funções sobre os números p-ádicos se revela mais restritiva. Tal restrição dificulta a verificação da diferenciabilidade para ordens superiores; esta dificuldade será amenizada por duas outras descrições da diferenciabilidade: o polinômio de Taylor e bases ortonormais. Notadamente, pela base ortonormal dos polinômios de Mahler, obteremos uma caracterização geométrica para as integrais sobre os números p-ádicos.

Para uma leitura mais satisfatória, este livro pressupõe, além dos conhecimentos matemáticos do ensino médio, uma familiaridade

- com a noção de um *anel* recordada em Apêndice A (a grosso modo, um conjunto com 0 e 1 sobre o qual operam $+$, $-$ e $\cdot$ tais que a lei distributiva, associativa e comutativa seja satisfeita; por exemplo, o anel dos números inteiros), e

- com a *aritmética modular* recordada em Apêndice B (a grosso modo, a divisão com resto aplicada às operações $+$ e $\cdot$ sobre um conjunto $\{0, 1, \ldots, n-1\}$ para n em $\mathbb{N}$).

1. História

Seja p um número primo. Os números p-ádicos tornaram-se populares quando Helmut Hasse (1898 – 1979, matemático alemão) mostrou na sua tese em 1921 que, para certas equações polinomiais inteiras (isto é, cujos coeficientes são números inteiros), a existência de soluções *locais* garante a existência de soluções inteiras: Ou seja, solução sobre $\mathbb{R}$ (isto é, com entradas reais) e módulo p^n para todo primo p e n em $\mathbb{N}$ garante a solução sobre $\mathbb{Z}$. Elaboremos:

1.1. Soluções de Equações Polinomiais

Dado um polinômio P em $\mathbb{Z}[X_1, \ldots, X_d]$, queremos encontrar as soluções em $\mathbb{Z}^d$ para a equação $P(x_1, \ldots, x_d) = 0$.

Por exemplo, $X^2 + Y^2 = Z^2$ tem a solução inteira $3^2 + 4^2 = 5^2$ ou $5^2 + 12^2 = 13^2$; pelo Teorema de Pitágoras, mostra que existe um triângulo retângulo cujos lados todos têm comprimentos inteiros.

Em geral, esta é porém uma tarefa dificílima como o revela o exemplo a seguir: A equação $X^n + Y^n = Z^n$ não tem solução cujas entradas são inteiros positivos para $n > 2$. É o (Último) Teorema de Fermat-Wiles; conjeturado pelo matemático francês Pierre de Fermat em 1637, demonstrado pelo matemático inglês Sir Andrew Wiles em 1995 por métodos sobretudo p-ádicos.

A primeira etapa na busca de soluções inteiras é encontrar soluções módulo p^n para todo primo p e expoente n em $\mathbb{N}$. Ao fixarmos p, esta existência de soluções para todos os expoentes n é resumida pela existência de uma solução nos *números p-ádicos inteiros* $\mathbb{Z}_p$ (cuja definição é iminente, caro leitor):

Teorema 1.1. *Seja* P *um polinômio em* $\mathbb{Z}[X_1, \ldots, X_d]$. *A equação* $P(x) \equiv 0 \bmod p^n$ *tem solução para todo* $n \in \mathbb{N}$ *se, e somente se,* $P(x) = 0$ *tem solução em* $\mathbb{Z}_p$.

Destacamos que os números p-ádicos inteiros não são apenas um registro para as congruências módulo p, p^2, ... A verdade é que $\mathbb{Z}_p$ é bem diferente dos anéis finitos $\mathbb{Z}/p\mathbb{Z}$, $\mathbb{Z}/p^2\mathbb{Z}$, ...; por exemplo, $\mathbb{Z}_p$ é não enumerável e tem característica 0.

1.2. Local e Global

A condição de existir solução módulo p^n para todo primo p e expoente n, é necessária, mas, em geral, não é suficiente. De fato, como etapa final, resta a responder: Quando é que

- a resolubilidade *local*, isto é,
  - sobre todos os $\mathbb{Z}_p$ para p primo, e
  - sobre os números reais $\mathbb{R}$,

é suficiente para

- a resolubilidade *global*, isto é sobre $\mathbb{Z}$?

Esta filosofia de obter resultados *globais* (sobre $\mathbb{Z}$) a partir de resultados *locais* (sobre os $\mathbb{Z}_p$ e $\mathbb{R}$) é recorrente na Teoria dos Números moderna; chama-se a *compatibilidade local-global*. Contudo, esta implicação raramente vale sem hipóteses adicionais. Por exemplo, os resultados *locais*, p-ádicos, obtidos por Sir Andrew Wiles e os seus colaboradores não bastaram em si para resolver o Último Teorema de Fermat; grande parte do trabalho consiste em deduzir o resultado global dos resultados locais.

Felizmente, no caso de *formas quadráticas*, isto é, polinômios da forma

$$\sum_{i,j=1,\dots,d} a_{i,j} X_i X_j$$

(por exemplo, aquele do Teorema de Pitágoras $X^2 + Y^2 - Z^2$), a resolubilidade local é de fato suficiente:

Teorema (Hasse-Minkowski). *Uma forma quadrática com coeficientes em $\mathbb{Z}$ tem solução não-trivial sobre $\mathbb{Z}$ se, e somente se, tem solução não-trivial sobre $\mathbb{Z}_p$ para todo p e sobre $\mathbb{R}$.*

Demonstração: A demonstração para um número de incógnitas $d < 4$ será dada em Apêndice E. □

Claro, se temos uma solução sobre $\mathbb{Z}$, em particular sobre todos os $\mathbb{Z}_p$ e sobre $\mathbb{R}$. A graça é que a existência de a priori *diferentes* soluções sobre os $\mathbb{Z}_p$ e sobre $\mathbb{R}$ implica a existência de uma solução *comum* sobre $\mathbb{Z}$ por trás delas. A facilidade de encontrar soluções sobre os $\mathbb{Z}_p$ e $\mathbb{R}$ permite uma classificação das formas quadráticas sobre eles que leva pelo Teorema de Hasse-Minkowski ao resultado seguinte:

Proposição. *Para uma forma quadrática com coeficientes em $\mathbb{Z}$ em > 4 incógnitas ter uma solução não-trivial sobre $\mathbb{Z}$, é suficiente ter uma solução sobre $\mathbb{R}$.*

Graças ao Teorema do Valor Intermediário, achar uma solução sobre $\mathbb{R}$ é bem mais simples do que sobre $\mathbb{Z}$.

2. Números p-ádicos

Seja p um número primo. Descreveremos os *números p-ádicos inteiros*, denotados por $\mathbb{Z}_p$, por três vias diferentes:

2.1. Explicitamente via a expansão p-ádica

A primeira descrição é análoga à expansão decimal de um número real no intervalo $[0,10]$ da forma

$$a_0 + a_1 10^{-1} + a_2 10^{-2} + \cdots, \quad \text{com os algarismos } a_0, a_1, a_2, \ldots \text{ em } \{0,1,\ldots,9\}.$$

Em contraste, para a expansão p-ádica,

- ao invés da base 10^{-1}, a base é p, e
- ao invés dos algarismos $\{0,\ldots,9\}$, os algarismos são $\{0,\ldots,p-1\}$.

Por exemplo, para $p = 2$, temos a expansão *binária* usual. Em geral:

Definição 2.1 (expansão p-ádica infinita). Seja

$$\mathbb{Z}_p := \left\{ a_0 + a_1 p^1 + a_2 p^2 + \cdots \quad \text{com } a_0, a_1, a_2, \ldots \in \{0,\ldots,p-1\} \right\}.$$

Neste ponto, esta expansão p-ádica é apenas definida como uma série infinita formal, isto é, como sequência de expansões p-ádicas finitas (seus truncamentos finitos):

$$a_0 + a_1 p^1 + a_2 p^2 + \cdots := (a_0, a_0 + a_1 p, a_0 + a_1 p + a_2 p^2, \ldots).$$

Por agora não podemos interpretá-la como série convergente; esta questão será discutida nas subseções seguintes. Assim, formalmente,

$$\mathbb{Z}_p = \left\{ (a_0, a_0 + a_1 p, a_0 + a_1 p + a_2 p^2, \ldots) \in \prod_{n=1,2,\ldots} \{0,1,\ldots,p^n - 1\} \right\} \tag{2.1}$$

com a_0, a_1, a_2, $\ldots \in \{0,\ldots,p-1\}$.

Como a notação sugere, $\mathbb{Z}$ está contido em $\mathbb{Z}_p$: enquanto os inteiros não-negativos são as expansões finitas, isto é, para as quais existe um N tal que $a_{N+1} = a_{N+2} = \cdots = 0$, as expansões dos inteiros negativos serão discutidas na próxima subseção (por Equação (2.2)), são as para que existem N tal que $a_{N+1} = a_{N+2} = \cdots = p-1$).

Nota. Percebemos duas diferenças com a expansão decimal dos números reais:

- Todo número real tem um único sinal ±, mas isto não vale para os números p-ádicos. Em particular, $\mathbb{Z}_p$ não é ordenado.
- A expansão p-ádica é única. Ao passo que, por exemplo, $0,\bar{9} = 1$ em $\mathbb{R}$.

Isto tudo define $\mathbb{Z}_p$ apenas como *conjunto*, mas não ainda como *anel*. Falta definir a adição e a multiplicação nos truncamentos finitos. Para isto, ao invés de defini-las em $\mathbb{Z}$, basta ver cada truncamento finito $a_0 + a_1 p + \cdots + a_n p^n$ como elemento em $\mathbb{Z}/p^n\mathbb{Z}$ ao invés de $\{0, \ldots, p^n - 1\}$.

2.2. Algebricamente via o limite inverso

Para dar a $\mathbb{Z}_p$ uma estrutura de anel particularmente simples, em (2.1) substituímos o conjunto $\{0, \ldots, p^n - 1\}$ pelo anel quociente $\mathbb{Z}/p^n\mathbb{Z}$; de forma que $\mathbb{Z}_p$ herde as operações deste anel.

Definição 2.2. Seja

$$\mathbb{Z}_p := \left\{ (a_0, a_0 + a_1 p, a_0 + a_1 p + a_2 p^2, \ldots) \in \prod_{n=1,2,\ldots} \mathbb{Z}/p^n\mathbb{Z} \right\},$$

com a_0, a_1, a_2, $\ldots \in \{0, \ldots, p - 1\}$.

Observação. O leitor que se pergunta porque p é sempre primo, mencionamos aqui que se $n = pq$ é composto por dois primos p e q, então $\mathbb{Z}_n = \mathbb{Z}_p \times \mathbb{Z}_q$ pelo Teorema Chinês dos Restos (vide Exercício 2.4).

Recordemos a construção de $\mathbb{Z}/p^n\mathbb{Z}$ no Apêndice B como conjunto $\{0, \ldots, p^n - 1\}$ cuja adição e multiplicação é definida pelo resto da divisão por p^n. Ao compararmos a definição de $\mathbb{Z}_p$ nesta subseção com a na subseção anterior, observamos que a única mudança foi dar uma estrutura de anel ao conjunto $\{0, \ldots, p^n - 1\}$, ou seja, passamos a vê-lo como o anel $\mathbb{Z}/p^n\mathbb{Z}$. Ressaltamos que ao invés de definir a adição e multiplicação da maneira mais evidente, isto é,

- permitir $\mathbb{Z}$ em vez de $\{0, \ldots, p^{n-1}\}$ na n-ésima coordenada com expansões p-ádicas finitas arbitrariamente grandes, que permanecem truncamentos dos seus sucessores, porém, agora de múltiplos parcelas, para fazer as adições e multiplicações em cada coordenada em $\mathbb{Z}$;

- agora as adições e multiplicações em cada coordenada n são feitas módulo p^n e os truncamentos unicamente do último termo.

Revela-se que as duas maneiras são equivalentes (pelo isomorfismo de anéis que projeta, na n-ésima coordenada, $\mathbb{Z} \twoheadrightarrow \mathbb{Z}/p^n\mathbb{Z}$). Como a segunda é mais simples, deixaremos a primeira atrás e trabalharemos unicamente com a segunda.

Podemos reformular Definição 2.2 sem usar expansões p-ádicas: Temos o homomorfismo

$$\begin{aligned} \rho_n \colon \mathbb{Z}/p^{n+1}\mathbb{Z} &\to \mathbb{Z}/p^n\mathbb{Z} \\ x &\mapsto r(x) \end{aligned}$$

onde $x = qp^n + r(x)$ com $r(x)$ em $\{0, \dots, p^n - 1\}$; isto é, ρ_n associa cada inteiro em $\{0, \dots, p^{n+1} - 1\}$ ao seu resto na divisão por p^n. Com efeito, olhando as expansões p-ádicas finitas,

$$a_0 + a_1 p^1 + \cdots + a_{n-1}p^{n-1} + a_n p^n \mapsto a_0 + a_1 p^1 + \cdots + a_{n-1}p^{n-1},$$

isto é, ρ_n suprime o último termo. Sejam

$$\cdots \to \mathbb{Z}/p^{n+1}\mathbb{Z} \xrightarrow{\rho_n} \mathbb{Z}/p^n\mathbb{Z} \to \cdots \to \mathbb{Z}/p^3\mathbb{Z} \xrightarrow{\rho_2} \mathbb{Z}/p^2\mathbb{Z} \xrightarrow{\rho_1} \mathbb{Z}/p\mathbb{Z} \to 0,$$

as supressões sucessivas dos últimos termos. Definimos $\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}$ onde

$$\varprojlim_{n=1,2,\dots} \mathbb{Z}/p^n\mathbb{Z} := \{(s_1, s_2, s_3, \dots) \in \prod_{n=1,2,\dots} \mathbb{Z}/p^n\mathbb{Z} : \rho_1(s_2) = s_1, \rho_2(s_3) = s_2, \dots\},$$

é chamado de *limite inverso* dos $\mathbb{Z}/p\mathbb{Z}$, $\mathbb{Z}/p^2\mathbb{Z}$, ... Esta construção revela $\mathbb{Z}_p$ como anel (discutida em mais detalhes em Exercício 2.3): Como ρ_1, ρ_2, ... são homomorfismos, $\mathbb{Z}_p$ é um subanel do produto infinito dos anéis $(\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p^2\mathbb{Z}) \times \cdots$, onde a adição e multiplicação é definida coordenada a coordenada. O seu elemento neutro da adição é $0 = (0, 0, \dots)$ e o da multiplicação $1 = (1, 1, \dots)$.

Em particular, todo inverso aditivo ou multiplicativo (caso exista, vide Exercício 2.5) em $\mathbb{Z}_p$ tem uma expansão p-ádica: Por exemplo, verificado por adição ou multiplicação em cada coordenada:

- Temos
$$-1 = (p - 1, p^2 - 1, \dots)$$
em $\mathbb{Z}_p$. Ao expandirmos $p^n - 1 = (p-1) + (p-1)p + \cdots + (p-1)p^{n-1}$ na base p,
$$-1 = (p-1) + (p-1)p + (p-1)p^2 + \cdots. \tag{2.2}$$
Por exemplo, $-1 = 1 + 2 + 2^2 + \cdots$ em $\mathbb{Z}_2$.

- Para $p > 2$, vale $2 = (2, 2, \ldots)$ e

$$1/2 = ((p+1)/2, (p^2+1)/2, (p^3+1)/2, \ldots)$$

em $\mathbb{Z}_p$. Ao expandirmos na base p,

$$(p^n+1)/2 = [(p-1)/2] \cdot p^{n-1} + \cdots + [(p-1)/2] \cdot p + (p+1)/2,$$

obtemos

$$1/2 = (p+1)/2 + [(p-1)/2] \cdot p + [(p-1)/2] \cdot p^2 + \cdots .$$

Por exemplo, $1/2 = 2 + 3 + 3^2 + 3^3 \cdots$ em $\mathbb{Z}_3$.

Neste ponto podemos dar a demonstração omitida de Teorema 1.1 que trata das soluções de equações polinomiais com coeficientes inteiros, por exemplo, $P(X) = X^p - X = 0$ para p primo. Verifica-se que qualquer x em $\mathbb{Z}/p\mathbb{Z}$ é solução desta equação. A partir delas, é possível calcular iterativamente outras nos demais anéis $\mathbb{Z}/p^2\mathbb{Z}$, $\mathbb{Z}/p^3\mathbb{Z}$, ... (confira Exercício 2.6). Por exemplo, $(p-1, p^2-1, p^3-1, \ldots)$ é uma tal sequência de soluções.

Teorema (1.1). *Seja* P *um polinômio em* $\mathbb{Z}[X_1, \ldots, X_d]$. *A equação* $P(x) \equiv 0 \mod p^n$ *tem solução para todo* $n \in \mathbb{N}$ *se, e somente se,* $P(x) = 0$ *tem solução em* $\mathbb{Z}_p$.

Demonstração: Contentemo-nos com o caso proto-típico $d = 1$ (para assim evitar o uso de sub-índices). Demonstremos primeiro que a existência de uma solução em $\mathbb{Z}_p$ implica a em todos os $\mathbb{Z}/p^n\mathbb{Z}$:

Satisfaça $x = (x_1, x_2, \ldots)$ em $\mathbb{Z}_p = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$ a equação

$$P(x) = P((x_1, x_2, \ldots)) = (P(x_1), P(x_2), \ldots) = 0$$

onde a segunda igualdade vale porque todas as operações + e · são definidas coordenada a coordenada. Isto é, temos para todo n em $\mathbb{N}$ um x_n em $\mathbb{Z}/p^n\mathbb{Z}$ que satisfaz $P(x_n) = 0$ em $\mathbb{Z}/p^n\mathbb{Z}$.

Demonstremos finalmente que a existência de uma solução em todos os $\mathbb{Z}/p^n\mathbb{Z}$ implica a em $\mathbb{Z}_p$: Satisfaçam x_n para todo n em $\mathbb{N}$ a equação $P(x_n) = 0$ em $\mathbb{Z}/p^n\mathbb{Z}$.

Por conveniência, levantamos (x_n) a uma sequência com entradas em $\mathbb{Z}$ que denotamos da mesma maneira. Filtramos (x_n), isto é, escolhemos, passo a passo, as entradas de uma subsequência (y_n) de (x_n) que satisfaz $y_{n+1} \equiv y_n \mod p^n$ para cada n; isto é, tal que $(\bar{y}_n)$ em $\varprojlim \mathbb{Z}/p^n\mathbb{Z} = \mathbb{Z}_p$ onde $\bar{y}_n$ é o valor de y_n sob a aplicação $\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$:

1. Como $\mathbb{Z}/p\mathbb{Z}$ é finito e $\{x_n\}$ é infinito, existe uma infinitude de números em $\{x_n\}$, denotados por $\{x_n^{(1)}\}$, que satisfazem $x_n^{(1)} \equiv y \bmod p$ para certo y em $\mathbb{Z}/p\mathbb{Z}$. Fixamos um tal y_1 em $\{x_n^{(1)}\}$; em particular, $P(y_1) \equiv 0 \bmod p$.

2. Da mesma maneira, como $\mathbb{Z}/p^2\mathbb{Z}$ é finito e $\{x_n^{(1)}\}$ é infinito, existe uma infinitude de números em $\{x_n^{(1)}\}$, denotados por $\{x_n^{(2)}\}$, que satisfazem $x_n^{(2)} \equiv y \bmod p^2$ para certo y em $\mathbb{Z}/p^2\mathbb{Z}$. Fixamos um tal $y_2 = x_n^{(2)}$ (onde n é suficientemente grande para $P(y_2) \equiv 0 \bmod p^2$); em particular, $y_2 \equiv y_1 \bmod p$.

3. $\cdots$

Por isso, a sequência (y_n) assim iterativamente construída satisfaz $y_{n+1} \equiv y_n \bmod p^n$, isto é $y = (y_n)$ em $\varprojlim \mathbb{Z}/p^n\mathbb{Z} = \mathbb{Z}_p$, e vale $P(y) \equiv P(y_n) \equiv 0 \bmod p^n$ para todo n, isto é, $P(y) = 0$ em $\mathbb{Z}_p$. □

Exercício 2.3. Verifique que $\mathbb{Z}_p$ é um anel (pela sua construção como limite inverso).

Dica. É a interseção dos núcleos dos homomorfismos

$$\rho_n(\pi_{n+1}) - \pi_n \colon \prod_i \mathbb{Z}/p^i\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$$

onde $\pi_n \colon \prod_i \mathbb{Z}/p^i\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ é a projeção a n-ésima coordenada para $n = 1, 2, \ldots$.

Exercício 2.4. Verifique que se $n = pq$ é composto por dois primos p e q, então $\mathbb{Z}_n = \mathbb{Z}_p \times \mathbb{Z}_q$.

Dica. Usa a construção como limite inverso, e o o Teorema Chinês dos Restos para cada fator.

Exercício 2.5. Usando a definição $\mathbb{Z}_p = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$, mostra que a é invertível em $\mathbb{Z}_p$ se, e somente se, p não divide a, ou, expresso pela expansão p-ádica, se, e somente se, o primeiro coeficiente a_0 não é zero.

Dica. Mostra que a é invertível em $\mathbb{Z}/p^n\mathbb{Z}$ se, e somente se, p não divide a. A este fim, considera o endomorfismo do grupo aditivo $\mathbb{Z}/p^n\mathbb{Z}$ dado pela multiplicação por a, isto é

$$\begin{aligned} a\cdot \colon \mathbb{Z}/p^n\mathbb{Z} &\to \mathbb{Z}/p^n\mathbb{Z} \\ x &\mapsto a \cdot x \end{aligned}$$

é sobrejetor se, e somente se, a é uma unidade. Como $\mathbb{Z}/p^n\mathbb{Z}$ é finito, é sobrejetor se, e somente se, é injetor. É injetor? (Para isto, mostra que um número primo p [= um número só dividido por 1 e ele mesmo] divide um produto ab se, e somente se, p divide ou a ou b pelo Teorema do Algoritmo de Euclides Estendido sobre o maior divisor comum.)

Exercício 2.6. Mostrar iterativamente que para qualquer a em $\{0,\dots,p-1\}$ existem a_n em $\mathbb{Z}$ tal que $a_n \equiv a \bmod p$ e ${a_n}^{p-1} \equiv 1 \bmod p^n$ para $n > 1$.

Dica. Para $n = 2$, usar a fórmula binomial

$$(a+px)^{p-1} = \sum_{i=0,\dots,p-1} \binom{p-1}{i} (px)^i a^{p-1-i},$$

e reduzir módulo p^2.

2.3. Analiticamente via o valor absoluto p-ádico

Em $\mathbb{R}$, uma expansão decimal $a_0 + a_1 10^{-1} + \cdots$ com a_0, a_1, ... em $\{0,1,\dots,9\}$ converge com respeito ao valor absoluto usual $|\cdot|$. Sobre $\mathbb{Z}$, definiremos um valor absoluto $|\cdot|_p$ tal que $a_0 + a_1 p^1 + \cdots$ convirja com respeito ao $|\cdot|_p$.

Seja A um anel. Um *valor absoluto* em A é uma aplicação $|\cdot| \colon A \to [0,\infty[$ com valores reais não-negativos tal que

- (Hausdorff) $|x| = 0$ se, e somente se, $x = 0$,
- (Multiplicatividade) $|xy| = |x||y|$, e
- (Desigualdade Triangular) $|x+y| \leq |x| + |y|$.

Ao invés de valor absoluto, usa-se também o termo *norma*. Chamamos um anel munido de uma tal norma de anel *normado*.

Definição. Seja

$$|\cdot|_p \colon \mathbb{Z} \to \mathbb{R}_{\geq 0}$$

a norma *p-ádica* definida por $|0|_p := 0$ e

$$|a|_p := p^{-v_p(a)}, \text{ onde } p^{v_p(a)} \text{ é a maior potência de } p \text{ que divide } a.$$

Exemplo.

- Para $p = 2$ temos que $|8|_2 = |2^3|_2 = 2^{-3}$ e $|9|_2 = |9 \cdot 2^0|_2 = 2^{-0} = 1$;

- para $p = 3$ temos que $|8|_3 = |3^0 \cdot 8|_3 = 3^{-0} = 1$ e $|9|_3 = |3^2|_3 = 3^{-2}$.

A norma p-ádica $|\cdot|_p$ mede quantas vezes p aparece na fatoração de um número inteiro. Contra-intuitivamente para quem se acostumou à norma usual, a norma p-ádica *diminui* quando a potência de p *cresce*; um número inteiro grande (com respeito à norma usual) pode ter norma p-ádica pequena.

Nota. A contra-intuição da norma p-ádica $|\cdot|_p$, em comparação à norma usual, é que ela é *não-arquimediana*, isto é

$$\underbrace{|1 + \cdots + 1|}_{n \text{ vezes}} \leq 1 \quad \text{para todo } n.$$

Apesar da sua natureza aparentemente exótica, Teorema 3.1 mostra que as únicas normas sobre $\mathbb{Q}$ são a norma usual e as normas p-ádicas.

Os números p-ádicos são relativamente recentes, em comparação aos números reais; foram introduzidos há cerca de cem anos por Kurt Hensel: Em analogia a $\mathbb{R}$, que consiste de todos os limites de $\mathbb{Q}$ para a norma usual $|\cdot|$, vamos definir $\mathbb{Z}_p$ como o conjunto dos limites de $\mathbb{Z}$ para a norma p-ádica $|\cdot|_p$. Formalmente, $\mathbb{Z}_p$ é o *completamento* de $\mathbb{Z}$ para $|\cdot|_p$ e o qual definiremos agora em geral, em passos.

Definição. Seja X um conjunto. Uma aplicação $\mathrm{d}\colon \mathrm{X} \times \mathrm{X} \to [0, \infty[$ é uma *função distância* ou *métrica* se

- $\mathrm{d}(x, y) = 0$ se, e somente se, $x = y$,
- $\mathrm{d}(x, y) = \mathrm{d}(y, x)$, e
- $\mathrm{d}(x, z) \leq \mathrm{d}(x, y) + \mathrm{d}(y, z)$.

Um *espaço métrico* é um par (X, d) de um conjunto X e uma função distância d como acima.

Uma aplicação f entre espaços métricos é *uniformemente contínua* se

$$\text{para todo } \epsilon > 0, \text{ existe } \delta > 0 \text{ tal que } \mathrm{d}(x, y) < \delta \text{ implica } \mathrm{d}(f(x), f(y)) < \epsilon.$$

Como as aplicações que respeitam a estrutura entre espaços topológicos são as aplicações *contínuas* e entre espaços vetoriais as aplicações *lineares*, aqui, entre espaços métricos, supomos todas as aplicações *uniformemente contínuas*.

Se $|\cdot|$ é uma norma sobre um anel, então

$$\mathrm{d}(x, y) := |x - y|$$

é uma função distância.

Definição. Uma sequência (x_n) em um espaço métrico é dita de *Cauchy* se para todo $\epsilon > 0$, existe N tal que $\mathrm{d}(x_n, x_m) < \epsilon$ para todos os $n, m > \mathrm{N}$. Um espaço métrico é *completo* se toda sequência de Cauchy converge.

Em particular, toda sequência que converge é Cauchy.

Teorema 2.7 (Completamento). *Para todo espaço métrico* X *existe um (único) espaço métrico completo* $\widehat{\mathrm{X}}$ *com uma aplicação (uniformemente contínua)* $\mathrm{X} \to \widehat{\mathrm{X}}$ *tais que para qualquer espaço métrico completo* Y *com uma aplicação (uniformemente contínua)* $\mathrm{X} \to \mathrm{Y}$*, existe uma aplicação (uniformemente contínua)* $\widehat{\mathrm{X}} \to \mathrm{Y}$ *que a fatora, isto é*

$$\begin{array}{ccc} \mathrm{X} & \longrightarrow & \mathrm{Y} \\ \downarrow & \nearrow & \\ \widehat{\mathrm{X}} & & \end{array}$$

Demonstração: Definimos o conjunto

$$\widehat{\mathrm{X}} := \{ \text{ sequências de Cauchy em X } \}/\sim$$

(e no qual X se injeta pelas sequências constantes) onde a relação de equivalência $\sim$ é definida por $(x_n) \sim (y_n)$ se $\mathrm{d}(x_n, y_n) \to 0$. A função distância $\hat{\mathrm{d}}$ é definida por

$$\hat{\mathrm{d}}([(x_n)], [(y_n)]) := \lim \mathrm{d}(x_n, y_n)$$

para representantes (x_n) e (y_n) das classes de equivalência $[(x_n)]$ e $[(y_n)]$. Observe que é bem-definida, isto é, independente dos representantes das classes de equivalência. Se (x_n) é uma sequência de Cauchy (de classes de equivalência de sequências de Cauchy) com $x_n = [(x_{n,m} : m \in \mathbb{N})]$, então ela converge a sequência diagonal $x = (x_{n,n} : n \in \mathbb{N})$. (O leitor é convidado a convencer-se da existência da aplicação $\widehat{\mathrm{X}} \to \mathrm{Y}$.) □

Observação. Uma abordagem a construir o espaço métrico $\mathbb{R}$ é como completamento de $\mathbb{Q}$ para a função distância usual; pergunta ao teu professor de cálculo ou vide [RRS11]. Se completamos $\mathbb{Q}$ a $\mathbb{R}$, observamos que

- a função distância tem imagem em $\mathbb{Q}_{\geq 0} = \{x \in \mathbb{Q} : x \geq 0\}$, e
- a função distância sobre $\mathbb{R} = \widehat{\mathbb{Q}}$ presica de ser definida, como ainda não construímos $\mathbb{R}$, por

$$\hat{\mathrm{d}}([(x_n)], [(y_n)]) := [(\mathrm{d}(x_n, y_n))].$$

Observe que $[(\mathrm{d}(x_n, y_n))] = \lim \mathrm{d}(x_n, y_n)$ pela definição do limite como sequência diagonal (dos representantes) das sequências constantes cujas entradas todas são $\mathrm{d}(x_n, y_n)$.

Revela-se por construção

- que a aplicação $\mathrm{X} \to \widehat{\mathrm{X}}$ é injetora, e
- que X é *denso* em $\widehat{\mathrm{X}}$; isto é, para todo $\hat{x}$ em $\widehat{\mathrm{X}}$ existem x_1, x_2, ... em X tal que $x_1, x_2, \ldots \to \hat{x}$, ou
  - mais formalmente: para qualquer $\epsilon > 0$ e $\hat{x}$ em $\widehat{\mathrm{X}}$ existe x em X tal que $\mathrm{d}(\hat{x}, x) < \epsilon$;
  - informalmente: todos os elementos no completamento são limites do espaço completado.

Demonstração:

- Sejam x e y em X e $(x_n) = (x, x, \ldots)$ e $(y_n) = (y, y, \ldots)$ representantes dos seus valores sob $\mathrm{X} \to \widehat{\mathrm{X}}$. Vale $(x_n) = (y_n)$ se, e tão-somente se, $\hat{\mathrm{d}}((x_n), (y_n)) = \mathrm{d}(x, y) = 0$, isto é, $x = y$.
- Dado $\epsilon > 0$ e $\bar{x}$ em $\widehat{\mathrm{X}}$ representado por uma sequencia de Cauchy $x = (x_n)$ em X, escolhe N tal que $\mathrm{d}(x_n, x_m) < \epsilon$ para todos os $n, nm > \mathrm{N}$. Seja $y := (x_\mathrm{N}, x_\mathrm{N}, \ldots)$ a sequência constante e $\bar{y}$ a sua classe residual em $\widehat{\mathrm{X}}$. Logo,

$$\hat{\mathrm{d}}(\bar{x}, \bar{y}) = \lim_n \mathrm{d}(x_n, y) = \lim_n \mathrm{d}(x_n, x_\mathrm{N}) \leq \max\{\mathrm{d}(x_m, x_\mathrm{N}) : m \geq \mathrm{N}\} < \epsilon.$$

(Vide também Exercício 2.11.) □

Definição. O anel normado dos *números p-ádicos inteiros* é definido por

$$\mathbb{Z}_p := \text{ o completamento de } \mathbb{Z} \text{ com respeito à norma } p\text{-ádica } |\cdot|_p \ .$$

Em particular, como $|p^n|_p = p^{-n}$ e $|a_n|_p = 1$ para a_n em $\{1, \ldots, p-1\}$, vale $|a_n p^n|_p = p^{-n}$ e

$$|a_n p^n + a_{n+1} p^{n+1} + \cdots + a_\mathrm{N} p^\mathrm{N}|_p = p^{-n} \to 0 \quad \text{para } n \to \infty.$$

Isto é, a sequência $(a_0, a_0 + a_1 p, a_0 + a_1 p + a_2 p^2, \ldots)$ é Cauchy. Como $\mathbb{Z}_p$ é completo, ela converge, isto é, a série infinita $a_0 + a_1 p + a_2 p^2 + \cdots$ converge. Isto

é, em $\mathbb{Z}_p$, todo número $a = a_0 + a_1 p + a_2 p^2 + \cdots$ existe como limites de números em $\mathbb{Z}$, como em $\mathbb{R}$ todo número $b = b_0 + b_1 10^{-1} + b_2 10^{-2} + \cdots$ para b_0, b_1, ... em $\{0,1,\ldots,9\}$ existe como limites de números em $\mathbb{Q}$. (Por exemplo, Exercício 2.12 mostra como obter outras expansões p-ádicas por esta convergência.)

Uma vez que $\mathbb{Z}$ é denso em $\mathbb{Z}_p$, as funções $+$, $\cdot$ e $|\cdot|_p$ de $\mathbb{Z}$ uniformemente contínuas estendem-se ao completamento $\mathbb{Z}_p$. (Se $f : X \to Y$ é uma função de X cujo contra-domínio Y é completo, então $f(x) := \lim x_n$ para $x_n \to \hat{x}$ em $\widehat{X}$ estende f a $\widehat{X}$.) Por isso, $\mathbb{Z}_p$ é verdadeiramente um anel com um valor absoluto (e não somente um espaço métrico).

Observação (ou digressão para o dia-a-dia do matemático). A construção do completamento $\widehat{X}$ para X, como a de $\mathbb{Z}/m\mathbb{Z}$ para m em $\mathbb{N}$ (como a de $\mathbb{Q}$ a partir de $\mathbb{Z}$), é uma construção *universal* por classes residuais. No final só importa

- que $\widehat{X}$ seja o "menor" espaço métrico em que toda sequência de Cauchy converge, e
- que $\mathbb{Z}/m\mathbb{Z}$ seja o "menor" anel em que $m(= 1 + \cdots + 1) = 0$, e
- que $\mathbb{Q}$ seja o "menor" anel em que todo número inteiro é invertível.

Basta-nos saber que todos os limites de sequências que convergem já estão em $\widehat{X}$. A construção é teoricamente importante, mas, uma vez feita, é praticamente deixada de lado. Para $\mathbb{Z}_p$, o completamento de $\mathbb{Z}$ pelo valor absoluto p-ádico $|\cdot|_p$, trabalharemos sobre ele como sobre $\mathbb{R}$: Aproveitemos que $\mathbb{Z}_p$ sempre contêm todos os limites, mas não nos assanhemos pela sua construção explicita como completamento. (Como ninguém pensa em uma classe residual de números racionais ao ver um número real.)

A árvore binária da Figura 2.1 representa $\mathbb{Z}_2$ da seguinte maneira: Cada número p-ádico tem uma expansão p-ádica dada pelos seus coeficientes, sendo ou 0, ou 1, e conforme a estes coeficientes, pega, ou o ramo da esquerda, ou o da direita em uma bifurcação (ou nó) da árvore.

Reflitamos como se descreve o valor absoluto, a distância e as bolas sobre eles:

- Vale $\mathrm{d}(x,y)_p = |x - y|_p = p^{-v}$ onde v = o nível em que os dois ramos infinitos x e y bifurcam; em particular, vale $|x|_p = p^{-v}$ onde v = o nível da primeira bifurcação em que o ramo infinito vai à direita.

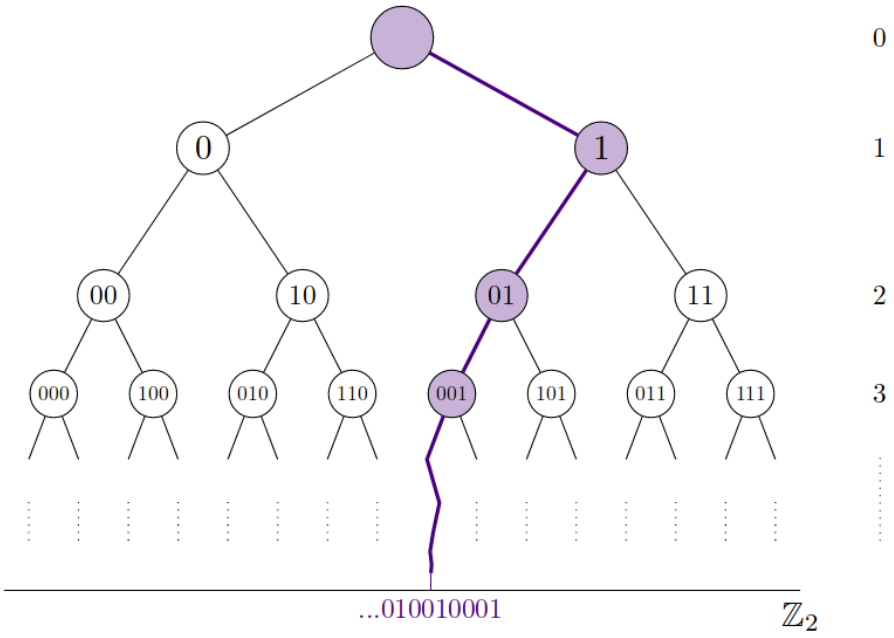


Figura 2.1: A árvore binária que representa $\mathbb{Z}_2$

- Uma bola corresponde a uma malha da árvore por $\mathrm{B}(x, p^{-n}) \mapsto x_0 + x_1 p + \cdots + x_n p^n$ se $x = x_0 + x_1 p + \cdots$, onde $\mathrm{B}(x, p^{-n})$ é a bola com centro x e raio p^{-n} e $x_0 + x_1 p + \cdots + x_n p^n$ é o ramo finito que leva à malha; a bola é dada por todos os ramos infinitos que passam pela malha correspondente. Observamos que duas bolas, ou se incluem, ou são disjuntas! Isto é, em termos topológicos, $\mathbb{Z}_p$ é *totalmente desconexo*.

Um elemento x em um anel A é uma *unidade*, ou *invertível*, se existe y, denotado por $y = x^{-1}$, tal que $xy = 1$. Um *corpo* é um anel em que todo elemento é invertível. Por exemplo, $\mathbb{Q}$ e $\mathbb{R}$ são corpos, e também, por Corolário B.4, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Seja A um anel comutativo sem divisores de 0, isto é, não existem x e y diferentes de 0 em A tais que $xy = 0$. Por exemplo, além dos corpos, $\mathbb{Z}$ e $\mathbb{Z}_p$ são tais anéis.

O *corpo das frações* Q de um tal A é o *menor* corpo que contém A (isto é, existe $\mathrm{A} \to \mathrm{Q}$ e, para qualquer outro corpo R com $\mathrm{A} \to \mathrm{R}$, existe $\mathrm{Q} \to \mathrm{A}$ que a fatore, isto é, tal que $\mathrm{A} \to \mathrm{R} = \mathrm{A} \to \mathrm{Q} \to \mathrm{R}$). Por exemplo, $\mathbb{Q}$ é o corpo das frações de $\mathbb{Z}$.

Ele é construído como conjunto por

$$\mathrm{Q} = \mathrm{A} \times \mathrm{A}/\sim$$

onde duas "frações" (x', y') e (x'', y'') são equivalentes se uma se simplifica a outra, isto é, $(x', y') \sim (x'', y'')$ se existe a em A tal que $a(x', y') = (ax', ay') = (x'', y'')$. A classe de equivalência de (x, y) em Q é denotada por x/y. Como anel, a adição e multiplicação é a de A em cada coordenada.

Definição. Seja

$$\mathbb{Q}_p := \text{ corpo das frações de } \mathbb{Z}_p \qquad \text{e} \qquad |x/y|_p := |x|_p/|y|_p$$

o corpo normado dos *números p-ádicos.*

Proposição 2.8. *O mergulho $\mathbb{Z} \to \mathbb{Z}_p$ induz para todo n em $\mathbb{N}$ um isomorfismo de anéis*

$$\mathbb{Z}/p^n\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}_p/p^n\mathbb{Z}_p.$$

Demonstração: Como a aplicação $\mathbb{Z} \to \mathbb{Z}_p \to \mathbb{Z}_p/p^n\mathbb{Z}_p$ tem núcleo $p^n\mathbb{Z}$, obtemos a injeção $\mathbb{Z}/p^n\mathbb{Z} \to \mathbb{Z}_p/p^n\mathbb{Z}$. Ela é sobrejetora se, e somente se, dado $\hat{x}$ em $\mathbb{Z}_p$ e n em $\mathbb{N}$, existe x em $\mathbb{Z}$ tal que $|\hat{x} - x|_p \leq p^{-n}$. Como $\mathbb{Z}$ é denso no seu completamento $\mathbb{Z}_p$, em particular tal x existe. □

Para um anel A, denote A^* o grupo multiplicativo das suas unidades. Por exemplo, $\mathbb{Z}^* = \{\pm 1\}$ e $\mathbb{Q}^* = \mathbb{Q} - \{0\}$.

Proposição 2.9. *Temos*

$$\mathbb{Z}_p^* = \mathbb{Z}_p - p\mathbb{Z}_p.$$

Isto é, x em $\mathbb{Z}_p$ é invertível se, e somente se, $|x| = 1$.

Demonstração: Seja x em $\mathbb{Z}_p$. Se x é invertível, isto é, existe y em $\mathbb{Z}_p$ tal que $xy = 1$, em particular $1 = |1|_p = |xy|_p = |x|_p|y|_p$ e logo $|x|_p = 1$.

Seja x em $\mathbb{Z}_p - p\mathbb{Z}_p$. Pela Proposição 2.8 e pelo Corolário B.4,

$$\mathbb{Z}_p/p\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$$

é um corpo, isto é, existe a em $\{1, \ldots, p-1\}$ tal que ax em $1 + p\mathbb{Z}_p$, isto é, $|ax - 1|_p < 1$. Pela série geométrica $1 + \alpha + \alpha^2 + \cdots \to 1/(1-\alpha)$, existe $(ax)^{-1}$, e em particular $x^{-1} = a(ax)^{-1}$.

(Para uma demonstração que $\mathbb{Z}_p^* = \mathbb{Z}_p - p\mathbb{Z}_p$ pela sua caraterização como limite projetivo, vide Exercício 2.5.) □

Recordemos que um *ideal* I em um anel A é um subconjunto I em A tal que

- para x e y em I, vale $x+y$ em I, e
- para x em I e y em A, vale xy em I.

Isto é, em comparação a um subanel, o ideal é fechado sob multiplicações por qualquer elemento do anel inteiro, não só do subanel.

Um ideal é *máximo* se não existe outro ideal, distinto do anel todo, que o contenha.

Corolário.

- *O único ideal máximo de* $\mathbb{Z}_p$ *é* $p\mathbb{Z}_p$.
- $\{$ *ideais em* $\mathbb{Z}_p$ $\} = \{\mathbb{Z}_p, p\mathbb{Z}_p, p^2\mathbb{Z}_p, \ldots\} \cup \{0\}$.
- $\mathbb{Q}_p = \mathbb{Z}_p[1/p] = \dot{\bigcup}_{n\in\mathbb{Z}} p^n\mathbb{Z}_p^* \cup \{0\}$.

Demonstração:

- Seja A um anel. O complemento $A - I$ de todo ideal I contém as unidades A^*. Logo, pela Proposição 2.9, se I é um ideal em $\mathbb{Z}_p$, logo $\mathbb{Z}_p - I \supseteq Z_p^* = Z_p - p\mathbb{Z}_p$, ou, equivalentemente, $I \subseteq p\mathbb{Z}_p$; isto é, $p\mathbb{Z}_p$ é o único ideal máximo.
- Seja x em $\mathbb{Z}_p$ com $|x| = p^{-n}$, isto é, $x = p^n a$ com $|a| = 1$. Pela Proposição 2.9, a em $\mathbb{Z}_p^*$, isto é, existe b em $\mathbb{Z}_p$ tal que $ab = 1$, ou, equivalentemente, $\mathbb{Z}_p x = p^{-n}\mathbb{Z}_p$. Concluímos que um ideal I em $\mathbb{Z}_p$ é gerado pelo seu elemento do mínimo valor, isto é, $I = p^n\mathbb{Z}_p$ onde p^n é a maior potência de p que divide todos os elementos em I.
- Pela Proposição 2.9, só falta inverter p em $\mathbb{Z}_p$ para poder inverter todos os elementos em $\mathbb{Z}_p$. O corpo $\mathbb{Q}_p$ é o menor anel em que todos os elementos em $\mathbb{Z}_p$ são invertíveis, logo $\mathbb{Q}_p = \mathbb{Z}_p[1/p]$. □

Por isso, se voltamos à definição de $\mathbb{Z}_p$ como conjunto de expansões p-ádicas infinitas,

$$\mathbb{Q}_p = \{p^{-n}a_{-n}p^{-n} + \cdots + a_0 + a_1p^1 + \cdots \text{ para } n \text{ em } \mathbb{N} \text{ e } a_{-n}, \ldots \in \{0, \ldots, p-1\}\}.$$

Corolário 2.10. *A aplicação*

$$\widehat{\mathbb{Z}} \xrightarrow{\sim} \varprojlim \mathbb{Z}/p^n\mathbb{Z}$$

é um isomorfismo de anéis.

Demonstração: A injetividade vale porque $x \mapsto 0$ se, e somente se, $|x|_p = 0$.

Para a sobrejetividade: Seja $(\bar{x}_n)$ em $\widehat{\mathbb{Z}}$ e seja (x_n) uma sequência com entradas em $\mathbb{Z}$ que a levanta; isto é, $x_{n+1} \equiv x_n \bmod p^n$, ou, equivalentemente, $|x_{n+1} - x_n|_p \leq p^{-n}$ para todo n. Então x_n é em particular uma sequência de Cauchy e, como $\widehat{\mathbb{Z}}$ é completo, converge a x. Vale $x \mapsto (\bar{x}_n)$. □

Observação. Duas séries iguais com entradas racionais podem convergir a limites diferentes em $\mathbb{Q}_p$ e $\mathbb{R}$. Por exemplo, com $\binom{x}{n} := x(x-1)\cdots(x-n+1)/n!$, a série dada por

$$\sqrt[2]{16/9} = (1+7/9)^{1/2} = \sum_{n\geq 0} \binom{1/2}{n} (7/9)^n$$

converge à raiz positiva $4/3$ em $\mathbb{R}$ e a raiz negativa em $\mathbb{Q}_7$ (vide [Kob84, Capítulo IV, Prova do Non-Theorem 1])!

Exercício 2.11. Usando a construção explicita da demonstração de Teorema 2.7, mostra que X (= as sequências constantes) é denso em $\widehat{\mathrm{X}}$.

Dica. Usando a definição da função distância em $\widehat{\mathrm{X}}$, dado $\epsilon > 0$, observa que para qualquer sequência de Cauchy (x_n), vale $\hat{\mathrm{d}}(x_n, x) < \epsilon$ para $x = x_{n_0}$ com n_0 um índice suficientemente grande.

Exercício 2.12. Usando a definição de $\mathbb{Z}_p$ como completamento de $\mathbb{Z}$, expande $2/3$ em $\mathbb{Z}_5$.

Dica. Escreva $\frac{2}{3} = \frac{4}{6} = 4 \cdot \frac{1}{1-(-5)}$ e utiliza que a série geométrica para $\frac{1}{1-(-5)} = 1 - 5 + 5^2 - 5^3 + \cdots$ converge em $\mathbb{Z}_5$.

O *diâmetro* de um subconjunto A num espaço métrico X é dado por

$$\operatorname{dia} \mathrm{A} := \sup\{\mathrm{d}(a,b) : a, b \in \mathrm{A}\}.$$

Exercício 2.13.

(i) Se A^- é o fecho de A, então $\operatorname{dia} \mathrm{A} = \operatorname{dia} \mathrm{A}^-$.

(ii) Se A é compacto, então o diâmetro de A é a maior distância entre os pontos em A.

Equipemos o plano p-ádico $\mathbb{Q}_p \times \mathbb{Q}_p$ com a norma $\|\cdot\|$ definida por $\|(x,y)\| = \max\{|x|_p, |y|_p\}$.

Exercício 2.14. Seja A um conjunto em $\mathbb{Q}_p \times \mathbb{Q}_p$. Mostra

$$\text{dia}(\text{A}) = \inf\{\text{dia}(\text{B}) : \text{B} \supseteq \text{A disco }\}.$$

Dica. Como $\text{dia}(\text{A}) = \text{dia}(\text{A}^-)$ para A^- o fecho de A, podemos supor A fechado. É suficiente mostrar que existe um disco B de diâmetro dia(A) que contém A: Como A é fechado e é contido em um disco, o qual é em particular compacta, segue que A é compacto. Logo, existem pontos a' e a'' em A cuja distância entre eles é

$$\text{d}(a', a'') = \text{dia}(\text{A})$$

Seja B um disco com a' e a'' no seu bordo. Para todo ponto fora do disco, uma das duas distâncias, ou a distância dele ao ponto a', ou a dele ao ponto a'', é pela desigualdade triangular forte maior do que o diâmetro do bolo. Isto é, por definição do diâmetro como distância máxima entre os pontos de A, todos os pontos fora de B estão fora de A; isto é, A é contido em B.

Exercício.

(i) Esta igualdade vale para um conjunto A em $\mathbb{R} \times \mathbb{R}$?

(ii) Caso não valha, qual é a constante mínima $\text{C} > 1$ tal que

$$\text{C} \cdot \text{dia}(\text{A}) \geq \inf\{\text{dia}(\text{B}) : \text{B} \supseteq \text{A disco }\}.$$

Dica. Considera um triângulo equilátero!

3. O Teorema de Ostrowski

Definição. Duas funções de distância são *(Cauchy-)equivalentes* se elas têm as mesmas sequências de Cauchy. Isto é, duas funções de distância d' e d'' sobre um conjunto X são equivalentes se para todo $\epsilon > 0$ existe δ tais que, para todos os x, y em X, se $d'(x,y) < \delta$ então $d''(x,y) < \epsilon$ e se $d''(x,y) < \delta$ então $d'(x,y) < \epsilon$

Observação (Exercício 3.2). Dados dois valores absolutos $|\cdot|'$ e $|\cdot|''$ sobre um corpo, as métricas induzidas são (Cauchy-)equivalentes se, e somente se, existe um $c > 0$ tal que $|\cdot|'' = |\cdot|'^{c}$.

O valor absoluto $|\cdot|$ é *trivial* se $|0| = 0$ e $|\cdot| = 1$ senão.

Teorema 3.1 (Ostrowski). *Todo valor absoluto não-trivial sobre $\mathbb{Q}$ é (Cauchy-)equivalente*

- *ou ao valor absoluto usual $|\cdot|$,*
- *ou a um valor absoluto p-ádico $|\cdot|_p$ para um número primo p.*

Demonstração: Seja $|\cdot|$ um valor absoluto sobre $\mathbb{Q}$. Distinguimos dois casos, o caso *arquimediano* e *não-arquimediano*. Como $|-1| = 1$ e $|x/y| = |x|/|y|$, basta verificar sobre $\mathbb{N}$ que existe $\alpha > 0$ tal que $|\cdot|^\alpha$ é ou igual ao valor absoluto usual, ou a um $|\cdot|_p$ para p um número primo.

Caso arquimediano: Existe um n em $\mathbb{N}$ tal que $|n| > 1$.

Seja n_0 o menor tal n. (Por exemplo, se $|\cdot|$ é o valor absoluto usual, então $n_0 = 2$.) Como $|n_0| > 1$, existe $\alpha > 0$ tal que $|n_0| = n_0^\alpha$.

Expanda n na base n_0, isto é

$$n = a_0 + a_1 n_0 + \cdots + a_s n_0^s \quad \text{com } a_0, \ldots, a_s \text{ em } \{0, \ldots, n_0 - 1\}.$$

Segue

$$\begin{aligned} |n| &= |a_0 + a_1 n_0 + \cdots + a_s n_0^s| \\ &\leq |a_0| + |a_1||n_0| + \cdots + |a_s||n_0|^s \end{aligned}$$

Como $a_0, \ldots, a_s < n_0$, pela escolha de n_0, vale $|a_0|, \ldots, |a_s| \leq 1$. Segue

$$\begin{aligned} |a_0| + |a_1||n_0| + \cdots + |a_s||n_0|^s &\leq 1 + n_0^\alpha + \cdots + n_0^{\alpha s} \\ &= n_0^{s\alpha}(1 + n_0^{-\alpha} + \cdots + (n_0^{-\alpha})^s) \end{aligned}$$

Esta soma é limitada pela série geométrica: Pondo $c = n_0^{-\alpha} < 1$, vale

$$\begin{aligned} 1 + n_0^{-\alpha} + \cdots + (n_0^{-\alpha})^s &= 1 + c + \cdots + c^s \\ &\leq 1 + c + c^2 + \cdots = \frac{1}{1-c} =: \mathrm{C} \end{aligned}$$

com $\mathrm{C} = \mathrm{C}(\alpha, n_0)$ independente de n. Como $n_0^s \leq n$, vale $|n| \leq n^\alpha \mathrm{C}$. Segue, para todo N em $\mathbb{N}$

$$|n|^{\mathrm{N}} = |n^{\mathrm{N}}| \leq (n^\alpha)^{\mathrm{N}} \mathrm{C},$$

extraindo a raiz de índice N

$$|n| \leq n^\alpha \sqrt[\mathrm{N}]{\mathrm{C}},$$

e como isto vale para N arbitrariamente grande

$$|n| \leq n^\alpha. \tag{$*$}$$

Para ver a desigualdade $|n| \geq n^\alpha$ oposta a $(*)$, observe

$$|n_0^{s+1}| = |n_0^{s+1} - n + n| \leq |n_0^{s+1} - n| + |n|,$$

e consequentemente, pela desigualdade obtida $(*)$,

$$|n| \geq |n_0^{s+1}| - |n_0^{s+1} - n| \geq n_0^{\alpha(s+1)} - (n_0^{s+1} - n)^\alpha.$$

Como $n_0^s \leq n \leq n_0^{s+1}$, segue

$$\begin{aligned} |n| &\geq (n_0^{s+1})^\alpha - (n_0^{s+1} - n_0^s)^\alpha \\ &= (n_0^{s+1})^\alpha \left(1 - (1 - \frac{1}{n_0})^\alpha\right) \geq n^\alpha \mathrm{D} \end{aligned}$$

com $\mathrm{D} = \mathrm{D}(\alpha, n_0) := 1 - (1 - \frac{1}{n_0})^\alpha$ independente de n. Como acima, segue $|n| \geq n^\alpha$ concluindo com $(*)$ que $n = n^\alpha$ e assim o caso arquimediano.

Caso não-arquimediano: Para todos os n em $\mathbb{N}$ vale $|n| \leq 1$.

Como $|\cdot|$ não é trivial, existe n em $\mathbb{N}$ tal que $n < 1$. Seja n_0 o menor tal n. Como $|\cdot|$ é multiplicativo, necessariamente $n_0 = p$ primo.

Proposição: Para todo número primo $q \neq p$ vale $|q| = 1$.

Demonstração: Caso contrário, existe N tal que $|q^{\mathrm{N}}|, |p^{\mathrm{N}}| < 1/2$. Pelo Teorema A.1, os números q^{N} e p^{N} são relativamente primos se, e somente se, $\langle q^{\mathrm{N}} \rangle + \langle p^{\mathrm{N}} \rangle = \langle 1 \rangle$, isto é, existem m, n tais que

$$mq^{\mathrm{N}} + np^{\mathrm{M}} = 1.$$

Segue a contradição

$$1 = |1| = |m||q^{\mathrm{N}}| + |n||p^{\mathrm{M}}| < 1/2 + 1/2 = 1.$$

Seja x em $\mathbb{N}$. Se $x = p_1^{x_1} \cdots p_r^{x_r}$ é a fatoração em números primos e $p_{i_0} = p$, então $|x| = |p_{i_0}|^{x_{i_0}}$. Isto é, $|x| = c^{v_p(x)}$ com $c = |p|$ e $v_p(x)$ = o maior n tal que p^n divide x.

Se escolhemos α tal que $c = |p| = |p|_p^{\alpha} = p^{-\alpha}$, isto é $\alpha = -\log_p |p|$, então concluímos que $|x| = |x|_p^{\alpha}$. □

Recordemo-nos de que um anel, ou corpo, com um valor absoluto um anel, ou corpo, é chamado *normado.*

Corolário. *Temos*

$$\{ \textit{ completamentos (normados) do corpo } \mathbb{Q} \}$$
$$= \{\mathbb{R}\} \cup \{ \textit{ todos os } \mathbb{Q}_p \textit{ para } p \textit{ um número primo } \}.$$

Demonstração: Segue da definição do completamento de um espaço métrico. □

Exercício 3.2. Dados dois valores absolutos $|\cdot|'$ e $|\cdot|''$ sobre um corpo, demonstra que as métricas induzidas sejam (Cauchy-)equivalentes se, e somente se, existe um $c > 0$ tal que $|\cdot|'' = |\cdot|'^{c}$.

Dica. Se são equivalentes, então use a multiplicatividade para mostrar que para todo x vale $|x|' < 1$ se, e somente se, $|x|'' < 1$. Use a multiplicatividade para concluir a existência de tal constante $c > 0$.

4. Corpos Completos

Seja $\mathbf{K}$ um corpo. Recordemos a *característica* de um corpo como o n em $\mathbb{N}$ que gera o núcleo de $\mathbb{Z} \to \mathbf{K}$.

4.1. A propriedade da não-arquimedianidade

Definição. A função $|\cdot| \colon \mathbf{K} \to [0, \infty[$ é um *valor absoluto não-arquimediano* se ela satisfaz

- $|x| = 0$ se, e somente se, $x = 0$,
- $|xy| = |x||y|$, e
- $|x + y| \leq \max\{|x|, |y|\}$.

Em comparação a um valor absoluto geral, a desigualdade triangular satisfeita por um valor absoluto não-arquimediano é mais forte, a *desigualdade triangular mais forte* ou *ultramétrica*.

Definição. A função $v \colon \mathbf{K} \to]-\infty, \infty]$ é uma *valoração* se ela satisfaz

- $vx = -\infty$ se, e somente se, $x = 0$,
- $v(xy) = vx + vy$, e
- $v(x + y) \geq \min\{vx, vy\}$

Se $|\cdot|$ é um valor absoluto não-arquimediano, então para qualquer $c > 1$, a função $v := \log_c |\cdot|$ é uma valoração; e, vice-versa, se v é uma valoração, então para qualquer $c < 1$ a função $|\cdot| := c^{v(\cdot)}$ é um valor absoluto não-arquimediano.

Exemplo.

- O corpo $\mathbf{K} = \mathbb{Q}_p$ com valoração $v(x) = n$ se $x = p^n \frac{x'}{x''}$ com $p \nmid x', x''$.
- O corpo $\mathbf{K} = \mathbb{F}_p((t))$ com valoração $v(x) = n$ se $x = a_n t^n + \cdots$ com $a_n \neq 0$.

Poderíamos concluir pelos exemplos que o valor absoluto $|\cdot|$ necessita um contorno da definição, mas é mais próximo do intuito (adquirido do cálculo real), enquanto a valoração é mais próxima da definição, ao custo do intuito.

Proposição 4.1 (Propriedades não-arquimedianas).

- *Se* $x = x_1 + \cdots + x_n$ *e existe* i_0 *em* $\{1, \ldots, n\}$ *tal que* $|x_{i_0}| > |x_i|$ *para* $i \neq i_0$, *então* $|x| = |x_{i_0}|$*;*
- *Se* $x_n \to x$, *então* $|x_n| = |x|$ *para* n *suficientemente grande.*
- *Se* $\mathbf{K}$ *é completo, então* $\sum_n x_n$ *converge se, e somente se,* $x_n \to 0$.

Demonstração:

- Observa que (o caso $n = 2$) se x, y em $\mathbf{K}$ e $|x| > |y|$, então $|x + y| \leq |x|$ e

$$|x| = |x + y - y| \leq \max\{|x + y|, |y|\},$$

  então, como $|y| < |x|$, segue $|x| \leq |x + y|$. Concluímos $|x| = |x + y|$.

- Se $\epsilon < |x|$, então $|x - x_n| < \epsilon$ para n suficientemente grande; logo, pelo primeiro item, $|x_n| = |x|$.
- Pela desigualdade triangular mais forte,

$$|x_m + \cdots + x_{\mathrm{M}}| \leq \max\{|x_m|, \ldots, |x_{\mathrm{M}}|\} \to 0.$$

4.2. Notações

Seja $\mathbf{K}$ um corpo com valor absoluto não-arquimediano.

- O *anel dos inteiros* $\mathcal{O}_{\mathbf{K}} := \{x \text{ em } \mathbf{K} \text{ tal que } |x| \leq 1\}$,
- o *ideal máximo* $\mathfrak{m}_{\mathbf{K}} := \{x \text{ em } \mathbf{K} \text{ tal que } |x| < 1\}$, e
- o *corpo residual* $\mathbf{k}_{\mathbf{K}} := \mathcal{O}_{\mathbf{K}}/\mathfrak{m}_{\mathbf{K}}$.

Notamos que $|x| = 1$ se, e somente se, $|x^{-1}| = 1$, e por isso a união disjunta $\mathcal{O}_{\mathbf{K}} = \mathcal{O}_{\mathbf{K}}^* \cup \mathfrak{m}_{\mathbf{K}}$; isto é, $\mathfrak{m}_{\mathbf{K}}$ é o ideal máximo local.

A valoração é *discreta* se $v(\mathbf{K}^*)$ é discreta em $\mathbb{R}$; se, e somente se, existe um c em $\mathbb{R}$ tal que $v(\mathbf{K}^*) = c\mathbb{Z}$; se, e somente se, existe um $\pi_{\mathbf{K}}$ em $\mathbf{K}$ que gera $\mathfrak{m}_{\mathbf{K}}$.

Tal $\pi_{\mathbf{K}}$ é um *uniformizador* de $\mathbf{K}$.

Exemplo.

- Se $\mathbf{K} = \mathbb{Q}_p$, então $\mathcal{O}_{\mathbf{K}} = \mathbb{Z}_p$, $\mathfrak{m}_{\mathbf{K}} = p\mathbb{Z}_p$ e $\mathbf{k}_{\mathbf{K}} = \mathbb{F}_p$ (e $\pi_{\mathbf{K}} = p$), e
- Se $\mathbf{K} = \mathbb{F}_p((t))$, então $\mathcal{O}_{\mathbf{K}} = \mathbb{F}_p[[t]]$, $\mathfrak{m}_{\mathbf{K}} = t\mathbb{F}_p[[t]]$ e $\mathbf{k}_{\mathbf{K}} = \mathbb{F}_p$ (e $\pi_{\mathbf{K}} = t$).

4.3. Corpos Locais

Um corpo com uma valoração v é *local* se

- o espaço métrico induzido é completo, e
- a valoração v é discreta, e
- o corpo residual $\mathbf{k_K}$ é finito.

Exemplo. Os corpos $\mathbb{Q}_p$ e $\mathbb{F}_p((t))$ são completos.

Teorema. *Temos*

$$\{ \textit{ corpos locais } \} = \{ \textit{ extensões finitas de } \mathbb{Q}_p \} \cup \{ \textit{ extensões finitas de } \mathbb{F}_p((t)) \}.$$

Proposição 4.2. *Seja*

$$\hat{\mathbf{k}}_\mathbf{K} = \{\textit{representantes de } \mathbf{k_K} \textit{ em } \mathcal{O}_\mathbf{K}\}$$

e $\pi_0, \pi_1, \ldots$ *em* $\mathcal{O}_\mathbf{K}$ *tais que* $v(\pi_0) = 0$, $v(\pi_1) = 1$, ...

Se $\mathbf{K}$ *é local, então para todo* x *em* $\mathcal{O}_\mathbf{K}$ *existem* x_0, x_1, ... *em* $\hat{\mathbf{k}}_\mathbf{K}$ *tais que* $x = x_0\pi_0 + x_1\pi_1 + \cdots$.

Demonstração: Seja $s\colon \mathcal{O}_\mathbf{K} \to \hat{\mathbf{k}}_\mathbf{K}$ a aplicação que fatora através de $\mathcal{O}_\mathbf{K} \to \mathbf{k_K}$. Ponha $x_0 = s(x/\pi_0)$, para obter $x = x_0\pi_0 + x_1'\pi_1$. Ponha $x_1 = s(x_1'/\pi_1)$, e assim por diante. □

Obtemos a seguinte generalização de Corolário 2.10, de $\mathbb{Z}_p$ a anéis locais:

Proposição. *Se* $\mathbf{K}$ *é local, então*

$$\mathcal{O}_\mathbf{K} \xrightarrow{\sim} \varprojlim \mathcal{O}_\mathbf{K}/\pi_\mathbf{K}^n\mathcal{O}_\mathbf{K}.$$

Demonstração: A injetividade vale porque $x \mapsto 0$ se, e somente se, $|x| = 0$. Para a sobrejetividade: Seja $(\bar{x}_n)$ uma sequência no lado esquerdo e seja (x_n) uma sequência que a levanta. Então $x_{n+1} \equiv x_n \bmod \pi_\mathbf{K}^n$ para todo n, isto é, $|x_{n+1} - x_n| \leq p^{-n}$; em particular, x_n é uma sequência de Cauchy e, como $\mathbf{K}$ é completo, converge a x, e vale $x \mapsto (\bar{x}_n)$. □

Corolário 4.3. *O anel topológico* $\mathcal{O}_\mathbf{K} = \varprojlim \mathcal{O}_\mathbf{K}/\pi_\mathbf{K}^n\mathcal{O}_\mathbf{K}$ *é compacto.*

Demonstração: Se $\mathbf{k_K}$ é finito, então $\mathcal{O}_\mathbf{K}/\pi_\mathbf{K}^n\mathcal{O}_\mathbf{K}$ é finito para todo n. Em particular compacto. Pelo Teorema de Tychonoff (Teorema C.11), $\mathcal{O}_\mathbf{K} = \varprojlim \mathcal{O}_\mathbf{K}/\pi_\mathbf{K}^n\mathcal{O}_\mathbf{K}$ (como subconjunto fechado do produto compacto) é compacto. □

O seguinte teorema surpreende, porque a partir de uma propriedade inteiramente topológica, a compacidade local, nasce um valor absoluto. Este valor absoluto é dado pela *medida de Haar* que existe sobre qualquer grupo topológico localmente compacto.

Teorema 4.4. *Temos*

$$\{ \textit{ corpos topológicos localmente compactos } \} = \{\mathbb{R}, \mathbb{C}\} \cup \{ \textit{ corpos locais } \}.$$

5. Método de Newton

Recordemos o *Método de Newton* sobre os números reais, antes de derivar o sobre um corpo não-arquimediano $\mathbf{K}$ tal como os números p-ádicos. A intuição geométrica sobre $\mathbb{R}$ fornece fórmulas que valem igualmente, até certo ponto, sobre $\mathbf{K}$.

5.1. Sobre $\mathbb{R}$

Dada uma função derivável $f : \mathbb{R} \to \mathbb{R}$, o *método de Newton* aproxima iterativamente um zero de f, isto é, define $x_1, x_2, \ldots$ em $\mathbb{R}$ com $x_n \to x$ tal que $f(x) = 0$. Geometricamente:

Pegamos um ponto (apropriado) x_1,

(i) olhamos o seu valor $f(x_1)$ sob f, e

(ii) a tangente em $f(x_1)$ ao grafo de f.

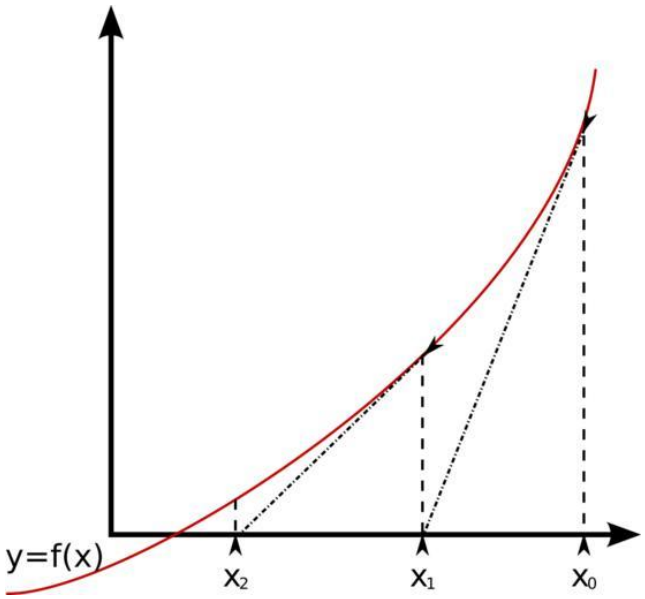
Esta tangente intersecta o eixo-x em um ponto x_2.

(i) Olhamos o seu valor $f(x_2)$ sob f, e

(ii) a tangente em $f(x_2)$ ao grafo de f;

e assim por diante.

Moralmente, dado um ponto x_0, a tangente t_{x_0} é a reta que aproxima f em x_0 o máximo. Por isso, o zero x_1 de t_{x_0}, que é facilmente computável graças a sua forma simples, aproxima o zero x de f. Em x_1, de novo, a tangente t_{x_1} é a reta que aproxima f em x_1 o máximo; de fato, como x_1 é mais próximo do zero x de f, a nova tangente t_{x_1} aproxima f mais do que t_{x_0} em volta de x; em particular, o x_2 de t_{x_1} é mais próximo do zero x de f do que o zero x_1 de t_{x_0}.



Como função no argumento h, na iteração n, a tangente t_n em $f(x_n)$ é definida por

$$t_n(x_n + h) := f(x_n) + f'(x_n)h,$$

logo, pondo $h = x - x_n$,

$$t_n(x) := f(x_n) + f'(x_n)(x - x_n).$$

Como $0 = t(x_{n+1})$, obtemos

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_{n+1})}.$$

Em outras palavras, o método de Newton é a iterada aplicação do *operador de Newton* $\mathrm{N} = \mathrm{N}_f : \mathbb{R} \to \mathbb{R}$ definido por

$$x \mapsto x - \frac{f(x)}{f'(x)}.$$

De fato, se $x_n := \mathrm{N}^n(x_0) \to x$, então $x = \mathrm{N}(x) = x - \frac{f(x)}{f'(x)}$, logo $f(x) = 0$.

A única condição é o encontro de um ponto de partida x_0 tal que a sequência $x_1, x_2, \ldots$ converge. Esta não é garantida, mas vale a convergência local quadrática:

Se f é duas vezes diferenciável, então pela Fórmula de Taylor

$$f(x + h) = f(x) + f'(x)h + \mathrm{R}''f(x + h, h)h^2$$

com $\mathrm{R}''f(x_h, h) \to f''(x)/2$ para $h \to 0$. Para $x + h = x_{n+1}$ e $x = x_n$, e porque $h = \frac{f(x_n)}{f'(x_n)}$, obtemos

$$\begin{aligned} f(x_{n+1}) &= f(x_n) + f'(x_n)(x_{n+1} - x_n) + \mathrm{R}''f(x_{n+1}, x_n)(x_{n+1} - x_n)^2 \\ &= \mathrm{R}''f(x_{n+1}, x_n)\left(\frac{f(x_n)}{f'(x_n)}\right)^2. \end{aligned}$$

Em particular,

$$|f(x_{n+1})| \leq \frac{|\mathrm{R}''f(x_{n+1}, x_n)|}{|f'(x_n)|^2} \cdot |f(x_n)|^2. \tag{5.1}$$

Pondo

- $\mathrm{M}'' := \sup |\mathrm{R}''f(x + h, x)|$,
- $m' := \inf |f'(x)| > 0$,

e $\mathrm{C} := \mathrm{M}''/m'^2$, obtemos

$$|f(x_{n+1})| \leq \mathrm{C} \cdot |f(x_n)|^2$$

Pondo $d_n = \mathrm{C} \cdot |f(x_n)|$, iterativamente

$$d_{n+1} \leq d_n^2 \leq d_{n-1}^{2\cdot 2} \leq d_{n-1}^{2\cdot 2\cdot 2} \leq \ldots \leq d_0^{2^n},$$

isto é,

$$\mathrm{C} \cdot |f(x_n)| \leq (\mathrm{C} \cdot |f(x_0)|)^{2^n}.$$

Concluímos que se $|f(x_0)| < 1/\mathrm{C} = m'^2/\mathrm{M}''$, então $f(x_n) \to 0$ quadraticamente. Por exemplo, se $|f(x_0)| \leq 10^{-1} \cdot m'/\mathrm{M}''$, então o número dos algarismos decimais nulos de $f(x_0)$, $f(x_1)$, $f(x_2)$, ... duplica a cada iteração.

Vemos aqui a convergência do método aplicado três vezes à função $f(x) = x^2 - 2$ com o seu zero $x_0 = \sqrt{2}$:

Iteração	Zero aproximativo	Erro
0	1	−0,4142135
1	1,5	0,0857864
2	1,41666667	0,0024531
3	1,41421569	0,0000021

5.2. Sobre um corpo não-arquimediano

Seja $\mathbf{K}$ um corpo *não-arquimediano*, isto é, um corpo equipado com um valor absoluto não-arquimediano que é não-trivial e torna $\mathbf{K}$ um espaço métrico completo, e seja $\mathcal{O}_{\mathbf{K}}$ seu anel dos números inteiros.

Por um lado, se a característica de $\mathbf{K}$ é 0, então a mesma definição de diferenciabilidade (iterada) como sobre $\mathbb{R}$ permite a mesma demonstração como sobre $\mathbb{R}$ da convergência do polinômio de Taylor. Por outro lado, se a característica de $\mathbf{K}$ é positiva, então os coeficientes do polinômio de Taylor 1, $1/2!$, $1/3!$, ... não são definidos, e precisa de uma definição da diferenciabilidade (iterada) mais rígida do que a sobre $\mathbb{R}$ para demonstrar a convergência do polinômio de Taylor.

Nesta Seção 5.2 consideramos apenas funções polinomiais, cujo polinômio de Taylor se define ad hoc: Seja A um anel e $f(\mathrm{X}) = a_d\mathrm{X}^d + \cdots + a_0$ em $\mathrm{A}[\mathrm{X}]$. Para $i \geq 0$, ponha

$$f^{<i>}(\mathrm{X}) := \binom{d}{i} a_d \mathrm{X}^{d-i} + \cdots + \binom{i}{i} a_i \quad \text{em } \mathrm{A}[\mathrm{X}].$$

De fato, $f^{<i>} = f^{(i)}/i!$ se $i!$ é invertível. Em particular, $f^{<i>} = f^{(i)}$ para $i = 0, 1, \ldots$ Vale a Fórmula de Taylor

$$f(\mathrm{X}+\mathrm{Y}) = f(\mathrm{X}) + f^{<1>}(\mathrm{X})\mathrm{Y} + \cdots + f^{<d>}(\mathrm{X})\mathrm{Y}^d.$$

Para $i \geq 0$, denotamos

$$\begin{aligned} \mathrm{R}^{<i>} f(\mathrm{X}+\mathrm{Y},\mathrm{X}) &= f(\mathrm{X}+\mathrm{Y}) - [f(\mathrm{X}) + \cdots + f^{<i-1>}(\mathrm{X})\mathrm{Y}^{i-1}] \\ &= f^{<i>}(\mathrm{X}) + f^{<i+1>}(\mathrm{X})\mathrm{Y} + \cdots + f^{<d>}(\mathrm{X})\mathrm{Y}^{d-i}. \end{aligned}$$

Em particular, $\mathrm{R}^{<i>} f(\mathrm{X}+\mathrm{Y},\mathrm{X}) = \mathrm{R}^i f(\mathrm{X}+\mathrm{Y},\mathrm{X})$ para $i = 0, 1, 2$.

Teorema 5.1. *Seja f em $\mathcal{O}_{\mathbf{K}}[\mathrm{X}]$. Se existe um x_0 em $\mathcal{O}_{\mathbf{K}}$ tal que*

$$|f(x_0)| < |f'(x_0)|,$$

então existe x em $\mathcal{O}_{\mathbf{K}}$ tal que $f(x) = 0$. Além disso, se $\lambda < 1$ é tal que

$$|f(x_0)| \leq \lambda |f'(x_0)|^2,$$

então existe um único x tal que $|x - x_0| \leq \lambda |f'(x_0)|$.

Demonstração: Como em Seção 5.1, chegamos à Equação (5.1), dando

$$|f(x_{n+1})| \leq \frac{|\mathrm{R}^2 f(x_{n+1}, x_n)|}{|f'(x_n)|^2} \cdot |f(x_n)|^2.$$

Como $f\colon \mathcal{O}_{\mathbf{K}} \to \mathcal{O}_{\mathbf{K}}$, segue $\sup\{|\mathrm{R}^2 f(x+h,x)|\} \leq 1$ (isto é, na notação de Seção 5.1 que $\mathrm{M}' = 1$),

$$|f(x_{n+1})| \leq \frac{|f(x_n)|^2}{|f'(x_n)|^2}. \tag{$*$}$$

Deduzamos por indução a partir de $(*)$ e pela hipótese $|f(x_0)| \leq \lambda |f'(x_0)|^2$ que

$$\frac{|f(x_n)|}{|f'(x_n)|} \leq \lambda^{2^n} |f'(x_0)|, \tag{$\dagger$}$$

resultando em

$$|f'(x_n)| = |f'(x_0)| \tag{$\ddagger$}$$

para todo $n = 0, 1, 2, \ldots$ Demonstremos primeiro $(\ddagger)$, a partir de $(\dagger)$, e depois $(\dagger)$: Notamos que $(\dagger)$ implica

$$|x_{n+1} - x_n| = \frac{|f(x_n)|}{|f'(x_n)|} \leq \lambda^{2^n} |f'(x_0)|$$

para $n = 0, 1, \ldots$, e consequentemente

$$\begin{aligned} |x_n - x_0| &\leq \max\{|x_1 - x_0|, \ldots, |x_n - x_{n-1}|\} \\ &\leq \max\{\lambda, \ldots, \lambda^{2^n}\} |f'(x_0)| < |f'(x_0)|. \end{aligned}$$

Pela Fórmula de Taylor,

$$f'(x+h) = f'(x) + h\mathrm{R}' f(x+h, x),$$

e porque $\sup\{|\mathrm{R}' f(x+h,x)|\} \leq 1$ (dado $f\colon \mathcal{O}_{\mathbf{K}} \to \mathcal{O}_{\mathbf{K}}$),

$$|f'(x+h) - f'(x)| \leq |h|.$$

Em particular,

$$|f'(x_n) - f'(x_0)| \leq |x_n - x_0| < |f'(x_0)|,$$

então, pela forte desigualdade triangular, $|f'(x_n)| = |f'(x_0)|$.

Demonstremos (†): Pela hipótese,

$$\frac{|f(x_0)|}{|f'(x_0)|} \leq \lambda |f'(x_0)|,$$

obtemos por ‡ para $n = 1$ que $|f'(x_1)| = |f'(x_0)|$. Por esta igualdade, pela equação $(*)$, pela desigualdade obtida acima (= a hipótese da indução) e pela hipótese,

$$\begin{aligned}\frac{|f(x_1)|}{|f'(x_1)|} &= \frac{|f(x_1)|}{|f'(x_0)|} \\ &\leq \left(\frac{|f(x_0)|}{|f'(x_0)|}\right)^2 \frac{1}{|f'(x_0)|} \leq \lambda^2 |f'(x_0)|.\end{aligned}$$

Iterativamente

$$|x_{n+1} - x_n| = \frac{|f(x_n)|}{|f'(x_n)|} \leq \lambda^{2^n} |f'(x_0)|.$$

Em particular, $x_0, x_1, x_2, \ldots$ é uma sequência de Cauchy, e por $(*)$ e (†),

$$f(x_{n+1}) = \frac{|f(x_n)|}{|f'(x_n)|} \leq \lambda^{2^n} |f'(x_0)| \to 0.$$

Logo, existe x em $\mathcal{O}_{\mathbf{K}}$ tal que $x_n \to x$ e $f(x) = 0$.

Suponhamos que $x = x', x''$ ambos satisfazem $f(x) = 0$ e $|x - x_0| \leq \lambda |f'(x_0)|$. Como

$$f(x'') = f(x') + f'(x')(x'' - x') + \mathrm{R}^2 f(x'', x')(x'' - x')^2,$$

se $x'' \neq x'$, então, porque $f : \mathcal{O}_{\mathbf{K}} \to \mathcal{O}_{\mathbf{K}}$ e por divisão com $x'' - x' \neq 0$,

$$|f'(x')| \leq |x'' - x'| \leq \max\{|x'' - x_0|, |x' - x_0|\} \leq \lambda |f'(x_0)| < |f'(x_0)|,$$

em contradição a $|f'(x)| = |f'(x_0)|$ para todos os n (como consequência de (‡)). □

Seja $\mathbf{k}_{\mathbf{K}}$ o corpo residual de $\mathcal{O}_{\mathbf{K}}$. Em particular, as condições do teorema são satisfeitas se $|f'(x_0)| = 1$, isto é, se $\bar{x}_0$ é uma raiz simples de $\overline{f(\mathrm{X})}$ em $\mathbf{k}_{\mathbf{K}}[\mathrm{X}]$. Por exemplo, $\mathrm{X}^p - \mathrm{X}$ tem p raízes simples em $\mathbb{F}_p$, e por isso p raízes em $\mathbb{Z}_p$. Isto é, todas as raízes da unidade de ordem $p - 1$ são contidas em $\mathbb{Z}_p$.

5.3. Lema de Hensel

Seja $\mathbf{K}$ um corpo *não-arquimediano* (isto é, um corpo equipado com um valor absoluto não-arquimediano que é não-trivial e torna $\mathbf{K}$ um espaço métrico completo) e seja $\mathcal{O}_\mathbf{K}$ seu anel dos números inteiros. Seja $\mathfrak{m}_\mathbf{K}$ o ideal máximo de $\mathcal{O}_\mathbf{K}$ e $\mathbf{k}_\mathbf{K} = \mathcal{O}_\mathbf{K}/\mathfrak{m}_\mathbf{K}$ o corpo residual.

Teorema 5.2 (Hensel). *Seja f em $\mathcal{O}_\mathbf{K}[\mathrm{X}]$. Se existem $\bar{g}$ e $\bar{h}$ em $\mathbf{k}_\mathbf{K}[\mathrm{X}]$ tais que eles são relativamente primos e*

$$\bar{f} = \bar{g}\bar{h},$$

então existem g e h em $\mathcal{O}_\mathbf{K}[\mathrm{X}]$ tais g, h reduzem a $\bar{g}, \bar{h}$ módulo $\mathfrak{m}_\mathbf{K}$, o grau de g é o grau de $\bar{g}$, e

$$f = gh.$$

Demonstração: Seja $d = \deg(f)$, $m = \deg(g)$, então $d - m \geq \deg(h)$. Sejam g_0, h_0 polinômios tais que $g_0 \equiv g$ e $h_0 \equiv h$ módulo $\mathfrak{m}_\mathbf{K}$ e $\deg(g_0) = m$, $\deg(h_0) \leq d - m$. Como g e h são relativamente primos, existem polinômios a, b em $\mathcal{O}_\mathbf{K}[\mathrm{X}]$ tais que $ag_0 + bh_0 \equiv 1 \bmod \mathfrak{m}_\mathbf{K}$ Entre os coeficientes dos dois polinômios $g - g_0h_0$ e $ag_0 + bh_0 - 1$ em $\mathfrak{m}_\mathbf{K}[\mathrm{X}]$, seja π um tal com valor absoluto mínimo.

Determinamos polinômios g e h da forma

$$g = g_0 + p_1\pi + p_2\pi^2 + \cdots$$
$$h = h_0 + q_1\pi + q_2\pi^2 + \cdots$$

onde $p_1, p_2, \ldots$ e $q_1, q_2, \ldots$ são polinômios de grau $< m$ respectivamente $\leq d - m$.

A este fim, determinamos sucessivamente polinômios

$$g_{n-1} = g_0 + p_1\pi + p_2\pi^2 + \cdots + p_{n-1}\pi^{n-1}$$
$$h_{n-1} = h_0 + q_1\pi + q_2\pi^2 + \cdots + p_{n-1}\pi^{n-1}$$

tais que

$$f \equiv g_{n-1}h_{n-1} \bmod \pi^n$$

Os limites de $g_1, g_2, \ldots$ e $h_1, h_2, \ldots$ serão os procurados polinômios g e h.

Para $n = 1$, esta congruência vale por escolha de g_0 e h_0. Para $n > 1$, como

$$g_n = g_{n-1} + p_n\pi^n \quad \text{e} \quad h_n = h_{n-1} + q_n\pi^n$$

a condição imposta a g_n e h_n se reduz a

$$f - g_{n-1}h_{n-1} \equiv (g_{n-1}p_n + h_{n-1}q_n)\pi^n \bmod \pi^{n+1}$$

Dividindo por π^n, equivalentemente

$$g_{n-1}q_n + h_{n-1}p_n \equiv g_0q_n + h_0p_n \equiv f_n \bmod \pi$$

onde $f_n = \pi^{-n}(f - g_{n-1}h_{n-1})$ em $\mathcal{O}_{\mathbf{K}}[X]$. Como $g_0a + h_0b \equiv 1 \bmod \pi$, vale

$$g_0af_n + h_0bf_n \equiv f_n \bmod \pi$$

Gostaríamos de pôr $q_n = af_n$ e $p_n = bf_n$, mas os graus são eventualmente grandes demais. Por isso, escrevemos

$$bf_n = qg_0 + p_n$$

com $\deg(p_n) < \deg(g_0) = m$. Como $g_0 \equiv \bar{g} \bmod \mathfrak{m}_{\mathbf{K}}$ e $\deg(g_0) = \deg(\bar{g})$, o coeficiente o mais alto de g_0 é uma unidade; por isso q em $\mathcal{O}_{\mathbf{K}}[X]$ e obtemos

$$g_0(af_n + h_0q) + h_0p_n \equiv f_n \bmod \pi.$$

Omitindo do polinômio $af_n + h_0q$ todos os coeficientes divisíveis por π, obtemos um polinômio q_n tal que $g_0q_n + h_0p_n \equiv f_n \bmod \pi$ e o qual, tomando em conta $\deg(f_n) \leq d$, $\deg(g_0) = m$ e $\deg(h_0p_n) < (d-m)+m = d$, satisfaz $\deg(q_n) \leq d-m$, como preciso. □

Para $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ em $\mathbf{K}[X]$, ponha

$$|f(x)| := \max\{|a_0|, \ldots, |a_n|\}.$$

Corolário 5.3 (Kurschak). *Seja $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ em $\mathbf{K}[X]$. Se f é irredutível e $a_0, a_n \neq 0$, então $|f| = \max\{|a_0|, |a_n|\}$.*

Demonstração: Após escalamento, podemos supor que $|f| = 1$. Seja r mínimo com $|a_r| = 1$. Logo,

$$f(x) \equiv x^r(a_r + \cdots + a_nx^{n-r}) \bmod \mathfrak{m}_{\mathbf{K}}.$$

Se $r \neq 0, n$, então f é redutível por Teorema 5.2. □

6. Extensões Finitas

Dois valores absolutos $|\cdot|''$ e $|\cdot|''$ são equivalentes se as suas funções de distância induzidas são *Cauchy-* ou *topologicamente* equivalentes. Enquanto estas duas noções de equivalência são distintas em *geral* (por exemplo, o $\exp\colon \mathbb{R} \to]0,\infty[$ é um isomorfismo topológico contudo o seu domínio $\mathbb{R}$ é completo e o seu contradomínio é incompleto, em particular não Cauchy-equivalente), recordemo-nos de que elas são iguais quando as funções de distância são induzidas de um valor absoluto.

Proposição 6.1. *Dois valores absolutos $|\cdot|'$ e $|\cdot|''$ são equivalentes se, e somente se, existe $\alpha > 0$ tal que $|\cdot|'' = |\cdot|'^{\alpha}$.*

Definição. Seja V um espaço vetorial sobre um corpo normado $\mathbf{K}$. Uma *norma* sobre V é uma aplicação $\|\cdot\|\colon \mathrm{V} \to [0,\infty[$ tal que

- $\|v\| = 0$ se, e somente se, $v = 0$,
- $\|\lambda v\| = |\lambda|\|v\|$ para todo λ em $\mathbf{K}$ e v em V, e
- $\|v + w\| \leq \|v\| + \|w\|$.

Uma norma $\|\cdot\|$ (sobre um espaço vetorial V) induz uma função distância d (sobre $\mathrm{V} \times \mathrm{V}$) por $\mathrm{d}(x,y) := \|x - y\|$.

Proposição. *Duas normas $\|\cdot\|'$ e $\|\cdot\|''$ são (Cauchy-)equivalentes se, e somente se, existe $c \leq 1 \leq \mathrm{C}$ tal que*

$$c\|\cdot\|' \leq \|\cdot\|'' \leq \mathrm{C}\|\cdot\|'.$$

Demonstração: A identidade é uma aplicação linear sobre um espaço normado, logo é limitada. □

Teorema 6.2. *Seja* V *é um espaço vetorial sobre um corpo normado completo. Se* V *tem dimensão finita, então todas as normas sobre ele são equivalentes e* V *é completo.*

Demonstração: Seja $e_1, \ldots, e_d$ uma base de V e $\|\cdot\|_{\max}$ a norma do máximo definida por

$$\|\lambda_1 e_1 + \cdots + \lambda_d e_d\|_{\max} = \max\{|\lambda_1|, \ldots, |\lambda_d|\}.$$

Seja $\|\cdot\|$ uma norma sobre V. Mostramos por indução em $n = \dim \mathrm{V}$ que $\|\cdot\|$ é equivalente a $\|\cdot\|_{\max}$.

Certamente vale para $n = 1$.

Seja $n > 1$. Temos

$$\|\lambda_1 e_1 + \cdots + \lambda_d e_d\| \leq |\lambda_1|\|e_1\| + \cdots + |\lambda_d|\|e_d\| \leq C \cdot \|\lambda_1 e_1 + \cdots + \lambda_d e_d\|_{\max}$$

com $C = \|e_1\| + \cdots + \|e_d\|$. Para concluir, precisamos de mostrar que existe $D > 0$ tal que

$$\|\cdot\|_{\max} \leq D\|\cdot\|.$$

Caso não, existe para todo n em $\mathfrak{n}$ um x_n em V tal que $|x_n|_{\max} > n\|x_n\|$. Em particular, após escalamento, uma sequência (x_n) em V tal que $x_n \to 0$ para $\|\cdot\|$ e $\|x_n\|_{\max} \geq 1$ para todos os n em $\mathbb{N}$. Seja $x_n = x_{n,1}e_1 + \cdots + x_{n,d}e_d$.

Como $\|x_n\|_{\max} \geq 1$, existe para cada n em $\mathbb{N}$ um i_n em $\{1,\ldots,d\}$ tal que $|x_{n,i_n}| \geq 1$. Existe i em $\{1,\ldots,d\}$ e uma subsequência $(x_{n_m} : m)$ tal que $i_{n_m} = i$ para todos os m em $\mathbb{N}$. Suponhamos que $i = 1$, e denotamos (x_{n_m}) por (x_n). Ponhamos $y_n = x_{n,1}^{-1}x_n = e_1 + \cdots$ e seja W o subespaço gerado por $e_2,\ldots,e_d$.

Como $y_n \to 0$ e $\|e_1\|$ é constante, a sequência $(y_n - e_1 : n)$ é uma sequência de Cauchy. Como W é completo (pela hipótese da indução), existe y em W tal que $y_n \to y + e_1$. Como $y_n \to 0$, logo $e_1 = -y$ em W, o que contradiz a definição de W. □

Corolário 6.3. *Seja* **K** *um corpo normado e* **L** *uma extensão de* **K**. *Se* **K** *é completo e* **L** *é finita, então o valor absoluto sobre* **K** *se estende unicamente.*

Demonstração: Seja $|\cdot|$ um valor absoluto sobre **L**. Por Teorema 6.2, qualquer outra norma, em particular valor absoluto, sobre o espaço vetorial finito **L** sobre **K** é equivalente a $|\cdot|$. Por Proposição 6.1, existe $\alpha > 0$ tal que ele é igual a $|\cdot|^\alpha$. Como ele é igual a $|\cdot|$ sobre **K**, ele é igual a $|\cdot|$ sobre **L**. □

Dada uma extensão finita **L** sobre **K**, a *norma* sobre **L** definida por

$$N_{\mathbf{L}/\mathbf{K}}\,\alpha := \det \alpha\cdot \tag{6.1}$$

onde $\alpha\cdot$ é o endomorfismo **K**-linear sobre **L** definida pela multiplicação por α.

Se $\mathbf{K} = \mathbb{Q}_p$ e $\mathbf{L} = \mathbb{Q}_p[\sqrt[2]{p}]$, então $|\sqrt[2]{p}| = \sqrt[2]{|p|}$ pela multiplicatividade. Se x em **L** é uma polinômio sobre **K** avaliado em $\sqrt[2]{p}$, a determinação de $|x|$ é mais complicada: No nosso exemplo $\mathbf{K} = \mathbb{Q}_p$ e $\mathbf{L} = \mathbb{Q}_p[\sqrt[2]{p}] = \mathbb{Q}_p 1 \oplus \mathbb{Q}_p\alpha$ para $\alpha = \sqrt[2]{p}$, calculamos que $\alpha\cdot$ é dada, nesta base de **L**, por

$$\begin{pmatrix} & p \\ 1 & \end{pmatrix};$$

logo $\det \alpha = -p$ e $N\,\alpha = \sqrt[2]{|p|}$.

Teorema 6.4. *Seja* $\mathbf{K}$ *um corpo valorado e* $\mathbf{L}$ *uma extensão de* $\mathbf{K}$. *Se* $\mathbf{K}$ *é completo e* $\mathbf{L}$ *é finita (de dimensão n), então o valor absoluto* $|\cdot|_{\mathbf{K}}$ *de* $\mathbf{K}$ *se estende a um valor absoluto* $|\cdot|_{\mathbf{L}}$ *sobre* $\mathbf{L}$ *por*

$$|x|_{\mathbf{L}} := \sqrt[n]{|\mathrm{N}_{\mathbf{L}/\mathbf{K}}\, x|_{\mathbf{K}}}$$

e $\mathbf{L}$ *é igualmente completo.*

Demonstração: Seja $\mathcal{O}_{\mathbf{L}}$ o anel dos inteiros de $\mathbf{K}$ e $\mathcal{O}$ o seu fecho integral em $\mathbf{L}$, isto é,

$$\mathcal{O} := \{\alpha \in \mathbf{L} : \text{ existe } f(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \cdots + a_0 \text{ em } \mathcal{O}_{\mathbf{K}}[\mathrm{X}] \text{ tal que } f(\alpha) = 0\}$$

Mostramos que

$$\mathcal{O} = \{\alpha \in \mathbf{L} : \mathrm{N}_{\mathbf{L}/\mathbf{K}}\, x \text{ em } \mathcal{O}_{\mathbf{K}}\}. \tag{$*$}$$

Se α em $\mathcal{O}$, então, porque invariante pelos mergulhos no fecho algébrico, o elemento $\mathrm{N}_{\mathbf{L}/\mathbf{K}}\, \alpha$ é em $\mathcal{O}_{\mathbf{K}}$.

Vice-versa, seja α em $\mathbf{L}$ tal que $\mathrm{N}_{\mathbf{L}/\mathbf{K}}\, \alpha$ em $\mathcal{O}_{\mathbf{K}}$. Seja $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0$ o polinômio mínimo de α em $\mathbf{K}[x]$.

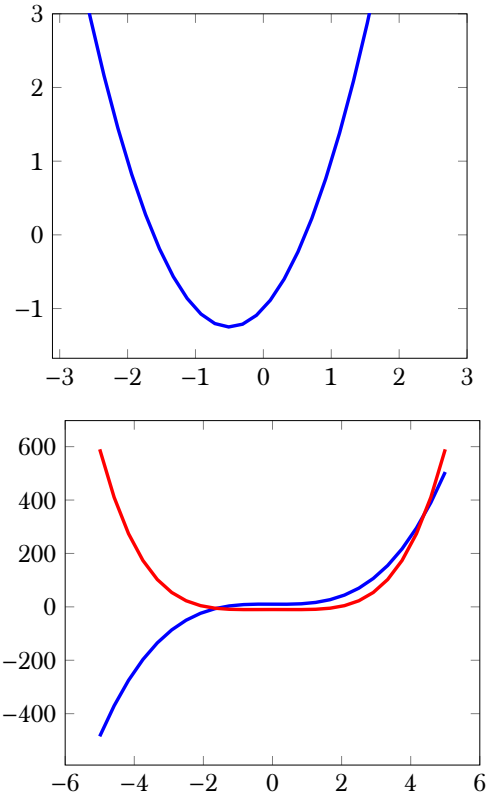
Como $\mathrm{N}_{\mathbf{L}/\mathbf{K}}\, \alpha = a_0^m$ em $\mathcal{O}_{\mathbf{K}}$, logo $|a_0| \leq 1$, isto é, a_0 em $\mathcal{O}_{\mathbf{K}}$. Por Corolário 5.3, todos os coeficientes têm valor absoluto ≤ 1, isto é, $f(x)$ em $\mathcal{O}_{\mathbf{K}}$, isto é, α em $\mathcal{O}$.

Verificamos que

$$|x| := \sqrt[n]{|\mathrm{N}_{\mathbf{L}/\mathbf{K}}\, x|} \tag{$**$}$$

é um valor absoluto sobre $\mathbf{L}$: Vale $|x| = 0$ se, e somente se, $x = 0$, e pela multiplicatividade de N, vale $|xy| = |x||y|$. Para $|x + y| \leq \max\{|x|, |y|\}$, basta após divisão por x ou y, mostrar que $|x| \leq 1$ implica $|x + 1| \leq 1$. Isto é, por $(*)$, se x em $\mathcal{O}$ então $x + 1$ em $\mathcal{O}$. Então $(**)$ define um valor absoluto, e sua restrição sobre $\mathbf{K}$ recupera o valor absoluto original. □

Por Corolário 6.3, este valor absoluto é o único sobre $\mathbf{L}$ que estende o sobre $\mathbf{K}$.



7. Teoria de Galois p-ádica

Definição. Um *polinômio* é uma expressão obtida pelas operações + e · sobre uma incógnita X e $\mathbb{Q}$.

Ele pode ser escrito da forma

$$a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$$

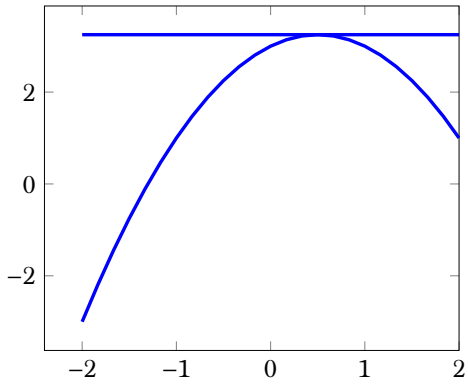
com a_n, a_{n-1}, …, a_0 em $\mathbb{Q}$; fornece uma função $f : \mathbb{R} \to \mathbb{R}$. Por exemplo, a função polinomial $f(x) = x^2 + x - 1$ tem a seguinte curva:

Frequentemente interessa o ponto

- em que duas tais curvas se intersectam, ou
- em que uma tal curva atinge seu máximo:

Achar as coordenadas destes pontos reduz-se à resolução de uma equação polinomial

$$f(X) = X^n + a_{n-1} X^{n-1} + \cdots + a_0 = 0.$$



Isto é, queremos calcular as *raízes* de f, os números $r_1, \ldots, r_n$ tais que $f(r_1)$, $\ldots$, $f(r_n) = 0$.

Questão. *Há uma fórmula para calcular as raízes de f?*

7.1. Soluções em grau menor

Quanto maior o grau n do polinômio, tanto mais engenhosidade requerida para calcular a raiz:

- (Completamento do Quadrado) Se $n = 2$, isto é $x^2 + px + q = 0$, então

$$x^2 + px + q = (x + p/2)^2 - p^2/4 + q$$

  e obtemos

$$x = -p/2 \pm \sqrt{p^2/4 - q}. \qquad (*)$$

- (Método de Cardan) Se $n = 3$, isto é $x^3 + ax^2 + bx + c = 0$, então

  (i) Substitua x por $\tilde{x} = x + h$ com $h = -a/3$, para obtermos

$$\tilde{x}^3 + p\tilde{x} + q = x^3 + ax^2 + bx + c.$$

  (ii) Substitua $\tilde{x}$ por $x' + x''$ tal que $x'x'' = -p/3$, para obter

$$x'^3 + x''^3 + (x' + x'')(3x'x'' + p) + q = x'^3 + x''^3 + q.$$

  (iii) Ponha $X' = x'^3$ e $X'' = x''^3$, para obtermos

$$X' + X'' = -q \qquad \text{e} \qquad X'X'' = -p^3/27.$$

Como

$$(X+\alpha)(X+\beta) = X^2 + (\alpha+\beta)X + \alpha\beta,$$

segue que os valores X' e X'' são as soluções de

$$X^2 - qX - p^3/27 = 0.$$

A fórmula $(*)$ para $n = 2$ nos dá para $\tilde{x} = \sqrt[3]{X'} + \sqrt[3]{X''}$,

$$\tilde{x} = \sqrt[3]{-q/2 + \sqrt{q^2/4 + p^3/27}} + \sqrt[3]{-q/2 - \sqrt{q^2/4 + p^3/27}}.$$

- Se $n = 4$, isto é $x^4 + \cdots = 0$, então o *Método de Ferrari* mostra como reduzir a uma equação polinomial de grau 3.

Então toda equação polinomial de grau 2, 3 ou 4 tem soluções que se exprimem

- pelos seus coeficientes $a_0, a_1, \ldots$ e números racionais,
- sujeitos às operações $+, \cdot$ e $\sqrt[n]{\cdot}$ (para $n = 2, 3, 4$)

Questão. *Há uma fórmula dando as raízes para $n = 5$?*

7.2. Permutações de Raízes

Para raízes $r_1, \ldots, r_n$, o seu *Corpo de Números* é

$$\begin{aligned}&\mathbb{Q}(r_1, \ldots, r_n)\\ :=&\{ \text{ todos os números obtidos por } + \text{ e } \cdot \text{ sobre } \mathbb{Q} \text{ e } r_1, \ldots, r_n\}\end{aligned}$$

Por exemplo, para $f(X) = X^4 - 2$ com raízes $\{\pm\sqrt[4]{2}, \pm\sqrt{-1}\sqrt[4]{2}\}$, estes números têm a forma

$$\begin{aligned}\mathbb{Q}(\sqrt{-1}\sqrt[4]{2}) = &\mathbb{Q} \quad \oplus\mathbb{Q}\sqrt[4]{2} \quad \oplus\mathbb{Q}\sqrt[4]{2}^2 \quad \oplus\mathbb{Q}\sqrt[4]{2}^3\\ &\oplus\mathbb{Q}\sqrt{-1} \quad \oplus\mathbb{Q}\sqrt{-1}\sqrt[4]{2} \quad \oplus\mathbb{Q}\sqrt{-1}\sqrt[4]{2}^2 \quad \oplus\mathbb{Q}\sqrt{-1}\sqrt[4]{2}^3,\end{aligned}$$

um espaço vetorial de dimensão 8 sobre $\mathbb{Q}$.

Definição (Corpo Radical). Um corpo de números $\mathbb{Q}(\alpha_1, \ldots, \alpha_m)$ é *radical* se, para cada $i = 1, \ldots, m$, existe s_i tal que

$$\alpha_i^{s_i} \text{ em } \mathbb{Q}(\alpha_1, \ldots, \alpha_{i-1}).$$

Por exemplo

$$r = \sqrt[2]{\sqrt[3]{2} + 5 - \sqrt[2]{12}}.$$

é no corpo de números radical

$$\mathbb{Q}(\sqrt[3]{2}, \sqrt[2]{12}, \sqrt[2]{\sqrt[3]{2} + 5 - \sqrt[2]{12}}).$$

Observação (Radical = Formulável). As raízes de um polinômio são num corpo de números radical se, e tão-somente se, são dadas por uma fórmula.

Notamos que,

- o corpo radical pode ser maior que o gerado pelas raízes;
- em particular os geradores podem diferir das raízes.

Questão. *Como as raízes revelam a radicalidade?*

Recordemo-nos de que um *automorfismo* é uma aplicação

- injetora cujo domínio iguala a sua imagem (= *auto*), e
- que respeita as operações + e · (= *homomorfismo*).

Definição (Grupo de Galois). Sejam $r_1, \ldots, r_n$ as raízes de um polinômio irredutível em $\mathbb{Q}[X]$. O seu *Grupo de Galois* é

$$\mathrm{Gal}(\mathbb{Q}(r_1, \ldots, r_n)/\mathbb{Q}) := \{ \text{ todas as } \textit{permutações} \text{ das raízes } r_1, \ldots, r_n \text{ que } \textit{se estendem a automorfismos sobre } \mathbb{Q}(r_1, \ldots, r_n) \ \}$$

Por exemplo para $f(X) = X^4 - 2$ e as suas raízes

$$\{\pm\sqrt[4]{2}, \pm\sqrt{-1}\sqrt[4]{2}\},$$

toda permutação σ que respeita + e · satisfaz

- $\sigma(-\cdot) = -\sigma(\cdot)$,
- $\sigma(\sqrt{-1}) = \pm\sqrt{-1}$,

Logo há 8 permutações no Grupo de Galois dadas

- por † em $\{\pm 1, \pm\sqrt{-1}\}$ dado por $\sqrt[4]{2} \mapsto \dagger\sqrt[4]{2}$, e

$$\left|\begin{array}{c|c|c|c} \sqrt[4]{2} & -\sqrt[4]{2} & \sqrt{-1}\sqrt[4]{2} & -\sqrt{-1}\sqrt[4]{2} \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \dagger\sqrt[4]{2} & -\dagger\sqrt[4]{2} & *\sqrt{-1}\dagger\sqrt[4]{2} & -*\sqrt{-1}\dagger\sqrt[4]{2} \end{array}\right|$$

- por $*$ em $\{\pm 1\}$ dado por $\sqrt{-1}\sqrt[4]{2} \mapsto *\sqrt{-1}\sqrt[4]{2}$,

da forma que as permutações são dadas pela tabela

Examinamos o corpo radical $\mathbb{Q}(\sqrt[n]{\alpha})$ que é incluso no corpo

$$\mathbb{Q}(\sqrt[n]{\alpha}, \zeta_n) \quad \text{onde } \zeta_n \text{ é uma raiz de 1 de ordem } n$$

gerado pelas raízes do polinômio $f(X) = X^n - \alpha$.

O Grupo de Galois G' de $\mathbb{Q}(\zeta_n)$ sobre $\mathbb{Q}$ é descrito por

$$\begin{aligned} G' &\hookrightarrow (\mathbb{Z}/n\mathbb{Z})^* \\ \sigma &\mapsto k \quad \text{determinado por } \sigma(\zeta) = \zeta^k, \text{ e} \end{aligned}$$

o Grupo de Galois G'' de $\mathbb{Q}(\sqrt[n]{\alpha}, \zeta_n)$ sobre $\mathbb{Q}(\zeta_n)$, isto é, que fixa todo elemento em $\mathbb{Q}(\zeta_n)$, é descrito por

$$\begin{aligned} G'' &\hookrightarrow \mathbb{Z}/n\mathbb{Z} \\ \sigma &\mapsto k \quad \text{determinado por } \sigma(\alpha) = \zeta^k\alpha. \end{aligned}$$

Os monomorfismos $G' \hookrightarrow \mathbb{Z}/n\mathbb{Z}^*$ e $G'' \hookrightarrow \mathbb{Z}/n\mathbb{Z}$ unem-se a

$$\mathrm{Gal}(\mathbb{Q}(\sqrt[n]{\alpha}, \zeta_n)/\mathbb{Q}) \hookrightarrow \begin{pmatrix} \mathbb{Z}/n\mathbb{Z}^* & \mathbb{Z}/n\mathbb{Z} \\ & 1 \end{pmatrix}$$

Recordemo-nos da notação $\mathbb{F}_p$ para o corpo finito de p elementos; explicitamente dado por $\mathbb{Z}/p\mathbb{Z}$.

Teorema (Galois). *Seja p um número primo e f em $\mathbb{Q}[X]$ de grau p. Existe uma fórmula para os zeros de f se, e tão-somente se, o Grupo de Galois dos zeros de f é contido em $\begin{pmatrix} \mathbb{F}_p^* & \mathbb{F}_p \\ & 1 \end{pmatrix}$*

Para $p = 5$ e $f(X) = X^5 - X + 1$, todas as permutações das raízes $r_1, \ldots, r_5$ respeitam $+$ e $\cdot$. Isto é, o Grupo de Galois é

$$\{ \text{ todas as permutações de } \mathbb{F}_5 \},$$

o qual não é um subgrupo de $\begin{pmatrix} \mathbb{F}_5^* & \mathbb{F}_5 \\ & 1 \end{pmatrix}$. Logo, não há fórmula.

7.3. Grupo de Galois

Um elemento α na extensão E de um corpo F é *algébrico* se existe $P(X) \in F[X]$ tal que $P(\alpha) = 0$. Entre todos os tais P com $P(\alpha) = 0$ existe um único polinômio $M(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ de grau mínimo (ou, equivalentemente, irredutível) e cujo coeficiente dominante é igual a 1, o *polinômio mínimo* de α. A extensão E de um corpo F é *algébrica* se todo elemento é algébrico; equivalentemente, se é gerada por elementos algébricos. Em particular, E é uma extensão algébrica finitamente gerada de F, se, e tão-somente se, o espaço vetorial E tem dimensão finita sobre F; chamemos uma tal extensão de extensão *finita* e denote $[E : F] = \dim_F E$. Se E é gerado por um elemento α e M(X) o seu polinómio mínimo, então $F[\alpha] = F[X]/M(X)F[X]$; em particular, $\dim_F E =$ grau de M(X).

Um polinómio P(X) sobre um corpo **K** é *separável* se se todos os zeros (em uma extensão normal de **K**) de P são diferentes.

Proposição 7.1. *Um polinômio* P(X) *é separável se, e tão-somente se,* P(X) *e* P′(X) *são relativamente primos.*

Demonstração: Se P(X) é separável, então para qualquer zero α, pelo produto $P(X) = (X - \alpha) \cdots Q(X)$, observamos que α não é um zero de P′(X).

Se P(X) não é separável, então $P = (X - \alpha)^2 Q(X)$, e α é um zero de P′(X) também pela regra de produto da derivação. □

Observação 7.2. O maior divisor comum de P e P′ pode ser computado pelo algoritmo de Euclides.

Proposição 7.3. *Os polinômios irredutíveis sobre um corpo k são todos separáveis se, e tão-somente se, a característica de k é* 0. *Mais precisamente, se* **K** *tem característica* 0*, então todo polinômio é separável, e se* **K** *tem característica* $p > 0$*, então o polinômio é separável se, e tão-somente se, em um polinómio em* X^p.

Demonstração: Seja P um polinómio irredutível. Por Proposição 7.1, P é separável se, e tão-somente se, P e P′ são relativamente primos. Caso contrário, P|P′ porque P é irredutível. Logo, como o grau de P′ é menor do que o de P, temos $P' = 0$. Se $P' = 0$, então P e P′ não são relativamente primos, então P não é separável. Isto é, P é separável se, e tão-somente se, $P' \neq 0$. Temos $P' = 0$ se, e tão-somente se, a característica é $p > 0$ e P é um polinômio em X^p. □

Uma elemento α da extensão E algébrica de um corpo F é *separável* se o polinômio mínimo M(X) de α é separável.

Observação 7.4. Se a característica de F é 0, então toda extensão é separável.

A extensão E de um corpo F é *separável* se todo elemento é separável; equivalentemente, se é gerada por elementos separáveis.

Teorema 7.5 (Teorema do Elemento Primitivo). *Uma extensão finita é separável se, e tão-somente se, é gerada por um único elemento separável.*

Proposição 7.6. *A extensão* E *finita de um corpo* F *é separável, se, e tão-somente se, para toda extensão normal* N *que contém* E *existem* $[\mathrm{E} : \mathrm{F}]$ *homomorfismos* $\mathrm{E} \hookrightarrow \mathrm{N}$ *que fixam* F.

Demonstração: Por indução, basta olhar o caso que α seja um elemento separável que gera E. Todo automorfismo que fixa F tem de enviar α a um zero do polinômio mínimo M de α, e o valor de α determina o automorfismo. Logo, existem $\leq \mathrm{E} : \mathrm{F}$ homomorfismos $\mathrm{E} \hookrightarrow \mathrm{N}$. Se E é separável, então existem $\mathrm{E} : \mathrm{F}$ zeros diferentes, e todo zero como valor de α determina um homomorfismo $\mathrm{E} \hookrightarrow \mathrm{N}$. □

Uma extensão E algébrica de um corpo F é *normal* se $\mathrm{P}(\mathrm{X}) \in \mathrm{F}[\mathrm{X}]$ e α em E tais que $\mathrm{P}(\alpha) = 0$, então todos os zeros de $\mathrm{P}(\mathrm{X})$ são em E. Uma extensão de *Galois* é uma extensão normal e separável.

A formulação e demonstração do Teorema Fundamental da Teoria de Galois segue as dadas por Emil Artin nas suas Notre Dame lectures. Notemos que E é Galois:

Teorema 7.7 (Teorema Fundamental da Teoria de Galois). *Seja* F *um corpo. Se* E *é uma extensão de Galois de* F*, isto é,* $\mathrm{E} = \mathrm{F}(a_1, \ldots, a_n)$ *é gerado por elementos distintos tais que* $(\mathrm{X} - a_1) \cdots (\mathrm{X} - a_n)$ *em* $\mathrm{F}[\mathrm{X}]$*, então*

- *o grupo* $\mathrm{G} = \mathrm{Aut}(\mathrm{E}/\mathrm{F})$ *é finito,*
- *são aplicações mutuamente inversas*

$$\begin{aligned} \{\textit{sub-extensões } \mathrm{E}/\mathrm{S}/\mathrm{F}\} &\tilde{\longrightarrow} \{\textit{subgrupos de } \mathrm{G}\} \\ \mathrm{S} &\mapsto \mathrm{Aut}(\mathrm{E}/\mathrm{S}) \\ \mathrm{E}^{\mathrm{H}} &\mapsfrom \mathrm{H} \end{aligned}$$

- *vale* $\mathrm{E} : \mathrm{S} = \#\,\mathrm{Aut}(\mathrm{E}/\mathrm{S})$ *e* $\#\mathrm{H} = \mathrm{E} : \mathrm{E}^{\mathrm{H}}$.

Demonstração: Demonstremos que as aplicações são mutuamente inversas, isto é

$$S = E^{\mathrm{Aut}(E/S)} \quad \text{e} \quad H = \mathrm{Aut}(E/E^H).$$

Como as inclusões valem sempre, bastaria mostrar que as cardinalidades são iguais, isto é,

$$E : S = E : E^{\mathrm{Aut}(E/S)} \quad \text{e} \quad \#H = \mathrm{Aut}(E/E^H).$$

Em vez de demonstrar estas igualdades, demonstremos

$$E : S = \#\,\mathrm{Aut}(E/S) \quad \text{e} \quad \#H = E : E^H \qquad (*)$$

que as implica por

$$E : S = \#\,\mathrm{Aut}(E/S) = E : E^{\mathrm{Aut}(E/S)} \quad \text{e} \quad \#H = E : E^H = \#\,\mathrm{Aut}(E : E^H);$$

e, além disso mostra à terceira (em particular, à primeira) parte da proposição! Logo, basta de demonstrar $(*)$.

Demonstremos $E : S = \#\,\mathrm{Aut}(E/S)$! Por indução, basta de demonstrar que se $\mathbf{K} = S(a_1, \ldots, a_{i-1})$ e $\mathbf{L} = \mathbf{K}(a_i)$ para $i \leq n$, então existem exatamente $\mathbf{L} : \mathbf{K}$ extensões do mergulho $\phi\colon \mathbf{K} \to E$ a $\mathbf{L}$. Ou, equivalentemente, que (a imagem sob ϕ d') o polinômio mínimo P de $a = a_i$ sobre $\mathbf{K}$ tem exatamente $\mathbf{L} : \mathbf{K}$ raízes. Como $\mathbf{L} = \mathbf{K}[X]/P\mathbf{K}[X]$, logo $\mathbf{L} : \mathbf{K}$ = grau de P, isto vale se, e tão-somente se, P tem nenhuma raiz dupla. Isto vale porque P, o polinômio mínimo de a_i, divide o polinômio $(X - a_1) \cdots (X - a_n)$ sem raiz dupla.

Demonstremos $\#H = E : E^H$! Como $H \subseteq \mathrm{Aut}(E/E^H)$ e $\#\,\mathrm{Aut}(E/E^H) = [E : E^H]$ pela primeira parte, basta mostrar $E : E^H = \dim_{E^H} E \leq \#H$. Isto é, quaisquer $b_1, \ldots, b_n$ em E para $n > \#H$ são linearmente dependentes sobre E^H; isto é,

$$b^\perp \cap (E^H)^n \neq 0$$

para $b = (b_1, \ldots, b_n)$ e $\cdot^\perp$ os vetores ortogonais em E^n a $\cdot$ com respeito ao produto escalar.

Se e em $(E^H)^n$ é ortogonal a b, então $he = e$ é ortogonal a hb; logo

$$b^\perp \cap (E^H)^n = (Hb)^\perp \cap (E^H)^n.$$

Seja x um dos vetores diferentes de zero $Hb^\perp$ com o número máximo de entradas iguais a zero. Seja $x_k \neq 0$ uma entrada diferente de zero; após escalamento, suponhamos $x_k = 1$.

Se x não fosse em $(E^H)^n$, isto é, existisse um h em H tal que $y := x - hx \neq 0$, então y em $Hb^\perp$ e $y_k = 0$; em contradição à escolha de x. □

7.4. Utilidade do Números p-ádicos

Como $\mathbb{Q}_p$ é completo, as propriedades algébricas são de um ponto de vista da Teoria dos Números mais fáceis do que as de $\mathbb{Q}$: Recordemo-nos de que o Grupo de Galois de uma extensão $\mathbf{E}$ de um corpo $\mathbb{F}$ consiste de todos os automorfismos do corpo $\mathbf{E}$ que fixam $\mathbb{F}$. Pela definição de $\mathbb{Q}_p$ como completamento de $\mathbb{Q}$, a inclusão $\mathbb{Q} \subseteq \mathbb{Q}_p$ é densa. Logo, toda aplicação contínua sobre $\mathbb{Q}_p$ é determinada pelos seus valores sobre $\mathbb{Q}$, isto é,

$$\mathrm{Aut}(\bar{\mathbb{Q}}_p) = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p) \hookrightarrow \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}).$$

Porém, ao contrário de $\mathbb{R}$, o completamento de $\mathbb{Q}$ para $|\cdot|$, cujo grupo de Galois absoluto

$$\mathrm{Gal}(\bar{\mathbb{R}}/\mathbb{R}) = \mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{\mathrm{id}, \bar{\cdot}\}$$

é finito, o grupo de Galois absoluto $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ de $\mathbb{Q}_p$, embora mais fácil do que o de $\mathbb{Q}$, continua a ser infinito.

8. Extensões Ciclotómicas

Seja **K** um corpo. Denote

$$\mu_n := \{x \text{ em } \mathbf{K} \text{ tais que } x^n = 1\}.$$

Pela iterada divisão com resto um polinômio P de grau n tem $\leq n$ raízes: pela divisão com resto $P = Q(X - \alpha) + R$ com grau $R < 1$, isto é, constante; como $P(\alpha) = 0$, necessariamente $R = 0$. Como $x^n = 1$ se e tão-somente se $P(x) = 0$ para $P(X) = X^n - 1$, vale

$$\#\{x \text{ em } \mathbf{K}^* \text{ tal que } x^n = 1\} \leq n$$

Se a característica de **K** é $p \nmid n$ ou 0, então

$$\#\mu_n = n$$

pois a derivada nX^{n-1} é relativamente primo a $X^n - 1$.

Proposição 8.1. *Todo grupo* G *tal que, para toda ordem n,*

$$\#\{g \text{ em } G \text{ tais que } g^n = 1\} \leq n \quad (\dagger)$$

é cíclico, *isto é, gerado por um elemento.*

Demonstração: Seja x um elemento em G de *ordem* (= o menor número $n > 0$ tal que $x^n = 1$) máxima. Se $n < \#G$, então há por (†) um y em G cuja ordem não divide n. Então a ordem de $z = xy$ é $> n$, em *contradição à escolha* de n. □

Corolário 8.2. *O grupo* μ_n *é cíclico.*

Demonstração: por Proposição 8.1, porque se x tem ordem n, se, e tão-somente se é zero de $P(X) = X^n - 1$; e $P(X)$ tem $\leq n$ = grau P zeros. □

8.1. Teoria de Galois

Um elemento α na extensão E de um corpo F é *algébrico* se existe $P(X) \in F[X]$ tal que $P(\alpha) = 0$. Entre todos os tais P com $P(\alpha) = 0$ existe um único polinômio $M(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ de grau mínimo (ou, equivalentemente, irredutível) e normalizado (isto é, cujo coeficiente dominante é igual a 1), o *polinômio mínimo* de α.

Observação. Observemos que, se M é um polinómio mínimo, então todo zero α de M tem M como polinómio mínimo, porque M anula α, é irredutível e normalizado. Logo, pela unicidade, M o polinómio mínimo de α.

8.2. Injetividade

Seja ζ_0 um gerador de μ_n e P o seu polinômio mínimo.

Lema 8.3. *Todos os zeros de* P *são geradores de* μ_n.

Demonstração: Se ζ não é gerador de μ_n, então ζ é zero de um polinômio $X^d - 1$ para $d \mid n$. Logo, o seu polinómio mínimo M divide $X^d - 1$. Se $P(\zeta)$ fosse 0, então, pela observação, $P = M|X^d - 1$. Em particular, todos os zeros de P estariam em μ_d; contradição a ζ_0 ser zero de P! Logo $P(\zeta) \neq 0$. □

Seja $\mathbb{Z}/n\mathbb{Z}$ o "menor" anel tal que $n \mapsto 0$, isto é, se A tal que $n \mapsto 0$, então existe $\mathbb{Z}/n\mathbb{Z} \to A$. Explicitamente, $\mathbb{Z}/n\mathbb{Z} = \{x + n\mathbb{Z} : n \in \mathbb{Z}\}$. Seja

$$\mathbb{Z}/n\mathbb{Z}^* = \{\text{todas as unidades em } \mathbb{Z}/n\mathbb{Z}\}.$$

e seja a *função de Euler* $\phi\colon \mathbb{N} \to \mathbb{N}$ dada por

$$n \mapsto \#\mathbb{Z}/n\mathbb{Z}^*.$$

Proposição 8.4 (Injetivade do homomorfismo). *Temos o mergulho*

$$\begin{aligned} \mathrm{Gal}(\mathbf{K}(\mu_n)/\mathbf{K}) &\hookrightarrow \mathbb{Z}/n\mathbb{Z}^* \\ \sigma &\mapsto k \text{ onde } \sigma(\zeta) = \zeta^k \end{aligned}$$

para um gerador ζ *de* μ_n.

Demonstração: Se P é o polinômio mínimo de ζ, então todo automorfismo σ em $\mathrm{Gal}(\mathbf{K}(\mu_n)/\mathbf{K})$ permuta os zeros de P, isto é, por Lema 8.3, permuta os geradores de μ_n. Como $\mu_n = \mathbb{Z}/n\mathbb{Z}$, e

$$\{\text{geradores de } \mathbb{Z}/n\mathbb{Z}\} = \mathbb{Z}/n\mathbb{Z}^*,$$

temos

$$\mathbb{Z}/n\mathbb{Z}^* \xrightarrow{\sim} \{\text{ geradores de } \mu_n\}$$

dado por $k \mapsto \zeta^k$ para um gerador ζ de μ_n. Logo, o homomorfismo é bem-definido.

É injetor porque ζ gera μ_n e por isso todo homomorfismo σ de $\mathbf{K}(\mu_n)$ que fixa $\mathbf{K}$ é determinado pela imagem $\sigma(\zeta)$. □

8.3. Sobrejetividade

Recordemo-nos de que $\mathbb{Z}$ é um domínio euclidiano, em particular, existe o maior divisor comum. Como $\mathbb{Z}^* = \{\pm 1\}$, definimo-lo pelo maior divisor comum positivo.

Definição 8.5. Um polinómio não-nulo $f(x)$ em $\mathbb{Z}[x]$ é *primitivo* se o maior divisor comum dos seus coeficientes é (associado a) 1.

Lema 8.6. *Sejam f e g em $\mathbb{Z}[x]$. Se f e g são primitivos, então fg é primitivo.*

Demonstração: Por contraposição: Seja fg não primitivo, isto é, existe p em $\mathbb{Z}$ que divide todos os coeficientes de fg, isto é, tal que

$$\overline{fg} = \bar{f}\bar{g} = 0 \in \mathbb{Z}/p\mathbb{Z}[x]$$

onde $\overline{fg}$, $\bar{f}$ e $\bar{g}$ denotem as reduções de fg, f e g módulo p, isto é, os polinómios cujos coeficientes são as reduções módulo p dos coeficientes de fg, f e g. Logo, como $\mathbb{Z}/p\mathbb{Z}$ é um domínio íntegro, ou $\bar{f}$, ou $\bar{g}$ é zero. Em particular, não ambos, f e g são primitivos. □

Proposição 8.7. *Seja f em $\mathbb{Z}[x]$. Se f é não-constante e primitivo, e se $f(x) = g(x)h(x)$ para $g(x)$ e $h(x)$ em $\mathbb{Q}[x]$, então $g(x)$ e $h(x)$ em $\mathbb{Z}[x]$.*

Demonstração: Seja $f = gh$ com f, g em $\mathbb{Q}[x]$. Logo, existem d e e em $\mathbb{Z}$ tais que

$$d \cdot f(x) = e g_0(x) h_0(x)$$

com g_0 e h_0 em $\mathbb{Z}[x]$ primitivos. Por Lema 8.6 $g_0(x)h_0(x)$ é primitivo. Como $f(x)$ é primitivo e não-constante, necessariamente $d = e$. Isto é, $f(x)$ é redutível em $\mathbb{Z}[x]$. □

Corolário (Lema de Gauss). *Seja f em $\mathbb{Z}[x]$. Se f não é constante e primitivo, então f é (ir)redutível em $\mathbb{Z}[x]$ se, e tão-somente se, f é (ir)redutível em $\mathbb{Q}[x]$.*

Demonstração: Se f é redutível em $\mathbb{Z}[x]$, então a fortiori em $\mathbb{Q}[x]$.

A implicação inversa é Proposição 8.7. □

Seja ζ uma raiz primitiva em μ_n e $m(x)$ o seu polinómio mínimo em $\mathbb{Q}[x]$. Por Lema 8.3,

$$\{ \text{ zeros de } m(x)\} \subseteq \{ \text{ raízes primitivas em } \mu_n\}.$$

Por Proposição 8.4, temos

$$\deg m(x) = \mathbb{Q}(\zeta) : \mathbb{Q} = \#\operatorname{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \leq \#\mathbb{Z}/\mathbb{Z}^* = \phi(n).$$

Temos #{ raízes de $m(x)$ } $\leq \deg m(x)$. Mostraremos

$$\#\{ \text{ raízes de } m(x) \} \geq \phi(n)$$

para concluir $\mathbb{Q}(\zeta) : \mathbb{Q} = \phi(n)$.

Lema 8.8. *Tem-se*

$$\prod_{\zeta \in \mu_n - \{1\}} 1 - \zeta = n.$$

Demonstração: Tem-se

$$\prod_{\zeta \in \mu_n - \{1\}} X - \zeta = X^n - 1/X - 1 = X^{n-1} + \cdots + X + 1.$$

Ao substituirmos X por 1, obtemos o resultado. □

Lema 8.9. *Seja p um número primo e ζ em μ_n. Se $p \nmid n$, então $\zeta^i \equiv 1 \bmod p$ implica $\zeta^i = 1$.*

Demonstração: Por contraposição, mostremos equivalentemente que se $\zeta^i - 1 \neq 0$, então $p \nmid \zeta^i - 1$. Como $p \nmid n$, por Lema 8.8,

$$\prod_{\zeta \in \mu_n - \{1\}} (1 - \zeta) = n \not\equiv 0 \bmod p.$$

Logo $1 - \zeta \not\equiv 0 \bmod p$ para todo $\zeta \neq 1$ em μ_n; em particular, para todo ζ^i. □

Como $m(x)$ é mínimo, $m(x) | X^n - 1$. Isto é, existe $h(x)$ em $\mathbb{Q}[x]$ tal que $X^n - 1 = m(x)h(x)$.

Proposição 8.10. *Seja ζ uma raiz de $m(x)$ e $X^n - 1 = m(x)g(x)$ e p um número primo. Se ζ^p é uma raiz de $h(x)$, então $p | n$.*

Demonstração: Como $X^n - 1$ é primitivo e não-constante, pelo Lema de Gauss, mais exatamente Proposição 8.7, $m(x)$ e $g(x)$ em $\mathbb{Z}[x]$. Se $h(\zeta^p) = 0$, então

$$0 = h(\zeta^p) \equiv h^{(p)}(\zeta^p) \equiv h(\zeta)^p \bmod p$$

onde $h^{(p)}(x)$ é o polinómio cujos coeficientes são as p-ésimas potências dos de $h(x)$. Logo, $X^n - 1 = \bar{m}(x)\bar{g}(x)$ em $\mathbb{Z}/p\mathbb{Z}[x]$ tem a raiz dupla ζ; equivalentemente, $n(\zeta^{n-1} - 1) \equiv 0 \bmod p$.

Como ζ é uma raiz de $m(x)$, em particular primitiva por Lema 8.3, em particular, $\zeta^{n-1} \neq 1$. Se $p \nmid n$, então, por Lema 8.9, $\zeta^{n-1} \not\equiv 1 \bmod p$. Como $\mathbb{Z}/p\mathbb{Z}$ é domínio íntegro, segue $n \equiv 0 \bmod p$; isto é, $p|n$; contradição a $p \nmid n$! Logo $p|n$. □

Corolário 8.11. *Temos*

$$\#\{ \textit{ raízes de } m(x) \} = \phi(n).$$

Demonstração: Seja ζ uma raiz de $m(x)$. Por Proposição 8.10, para todo primo p que não divide n, também ζ^p é uma raiz de $m(x)$. Como tais p geram o grupo multiplicativo $\mathbb{Z}/n\mathbb{Z}^*$ e

$$\mathbb{Z}/n\mathbb{Z}^* \xrightarrow{\sim} \{\text{raízes primitivas em } \mu_n\},$$

logo

$$\{ \text{ raízes de } m(x) \} = \{\text{raízes primitivas em } \mu_n\}.$$

Corolário 8.12. *Temos*

$$\mathbb{Q}(\zeta) : \mathbb{Q} \geq \phi(n).$$

Demonstração: Temos

$$\mathbb{Q}(\zeta) : \mathbb{Q} = \deg m(x) \geq \#\{ \text{ raízes de } m(x) \} = \phi(n);$$

a última igualdade por Corolário 8.11. □

Corolário 8.13 (Sobrejetividade do homomorfismo sobre $\mathbb{Q}$). *O homomorfismo entre grupos*

$$\begin{aligned} \mathrm{Gal}(\mathbb{Q}(\mu_n)/\mathbb{Q}) &\hookrightarrow \mathbb{Z}/n\mathbb{Z}^* \\ \sigma &\mapsto k \textit{ onde } \sigma(\zeta) = \zeta^k \end{aligned}$$

para um gerador ζ de μ_n é sobrejetor.

Demonstração: Por Proposição 8.4 o homomorfismo é injetor. Por Corolário 8.12 e pelo Teorema Fundamental da Teoria de Galois, o lado esquerdo tem cardinalidade $\phi(n)$ = a cardinalidade do lado direito. Logo, o homomorfismo é bijetor. □

8.4. Cálculo da Função de Euler

Para calcular $\phi(n)$, decomponhamos n nos seus fatores primos.

Definição 8.14. Seja A um anel. Para dois elementos a e b,

- um *maior divisor comum* $d = \mathrm{mdc}(a,b)$ é um divisor de a e b, tal que, se e é outro divisor de a e b, então d divide e.
- um *menor múltiplo comum* $m = \mathrm{mmc}(a,b)$ é um múltiplo de a e b, tal que, se n é outro múltiplo de a e b, então n é múltiplo de m.

Um elemento a em A é um *divisor de zero* se existe b em A não-nulo tal que $ab = 0$. A é um *domínio íntegro* se não tem divisores de zero.

Dois elementos a e b em A são *associados* se existe ϵ em A^* tal que $b = \epsilon a$.

Observação 8.15. Seja A um anel. Se A é um domínio íntegro, então o maior divisor comum e o menor múltiplo comum de dois elementos a e b em A é univocamente determinado exceto associação, isto é,

- dois maiores divisores d' e d'' comuns são associados, e
- dois menores múltiplos m' e m'' comuns são associados.

Demonstração: Se d' e d'' são dois maiores divisores, então, por definição, $d'|d''$ e $d''|d'$, isto é, existem u e v em A tal que $d' = uvd'$ e $d'' = uvd''$. Como A não tem divisores de zero, ou d' e d'' são nulos, ou $uv = 1$; em ambos os casos, d' e d'' são associados.

Da mesma maneira para dois menores múltiplos comuns. □

Um anel A é um *domínio euclidiano* se existe uma *função de grau* $v\colon A \to \mathbb{N} \cup \{-\infty\}$ tal que $v(0) = -\infty$ e para todo a e b não-nulo com $v(b) \leq v(a)$ existem q e r tais que

$$a = bq + r \quad \text{com } v(r) < v(b).$$

Exemplo 8.16. Anéis euclidianos são

(i) o anel $\mathbb{Z}$ com $v = |\cdot|$,

(ii) o anel $\mathbb{Z}[i]$ com $v(a + bi) = a^2 + b^2$, e

(iii) o anel polinomial A[X] para um domínio íntegro A com $v(f)$ definido pelo grau de f.

Lema 8.17. *Se* A *é um domínio euclidiano, então* A *é um domínio principal.*

Demonstração: Seja I um ideal em A e i_0 em I um elemento não-nulo em I de grau mínimo. Para todo $a = i$ em I e $b = i_0$ existem q e r tais que

$$a = qb + r \quad \text{com } v(r) < v(b).$$

Em particular, r em I. Como $v(r) < v(b)$ e $v(b) = v(i_0)$ é mínimo, $r = 0$. Isto é, $\mathrm{I} = \langle i_0 \rangle$. □

O maior divisor comum e o menor múltiplo comum de dois elementos a e b em um anel A não sempre existe. Porém, se A é um domínio de fatoração única, então existe, sim. Um domínio principal é em particular um de fatoração única; a seguinte Proposição 8.18 mostra diretamente que o maior divisor comum e o menor múltiplo comum em um domínio principal sempre existe:

Proposição 8.18. *Seja* A *um anel. Se* A *é um domínio principal, então*

$$\langle a \rangle + \langle b \rangle = \langle \mathrm{mdc}(a,b) \rangle \quad e \quad \langle a \rangle \cap \langle b \rangle = \langle \mathrm{mmc}(a,b) \rangle$$

Demonstração: Como A é um domínio principal, existe d em A tal que $\langle a \rangle + \langle b \rangle = \langle d \rangle$. Em particular, $d|a, b$. Logo,

$$\langle \mathrm{mdc}(a,b) \rangle \subseteq \langle d \rangle = \langle a \rangle + \langle b \rangle$$

Como $\mathrm{mdc}(a,b)$ pertence a $\{d \in \mathrm{A} : d|a, b\}$,

$$\langle a \rangle + \langle b \rangle \subseteq \langle \mathrm{mdc}(a,b) \rangle.$$

Concluímos

$$\langle \mathrm{mdc}(a,b) \rangle \subseteq \langle a \rangle + \langle b \rangle \subseteq \langle \mathrm{mdc}(a,b) \rangle,$$

logo $\langle a \rangle + \langle b \rangle = \mathrm{mdc}(a,b)$.

Semelhantemente: Como A é um domínio principal, existe m em A tal que $\langle a \rangle \cap \langle b \rangle = \langle m \rangle$. Em particular, $a, b|m$. Logo,

$$\langle \mathrm{mmc}(a,b) \rangle \supseteq \langle m \rangle = \langle a \rangle \cap \langle b \rangle$$

Como $\mathrm{mmc}(a,b)$ pertence a $\{m \in \mathrm{A} : a, b|m\}$,

$$\langle a \rangle \cap \langle b \rangle \supseteq \langle \mathrm{mmc}(a,b) \rangle.$$

Concluímos

$$\langle \mathrm{mmc}(a,b) \rangle \subseteq \langle a \rangle \cap \langle b \rangle \subseteq \langle \mathrm{mmc}(a,b) \rangle,$$

logo $\langle a \rangle \cap \langle b \rangle = \mathrm{mmc}(a,b)$. □

Nota. O maior divisor comum de dois números em um domínio euclidiano pode ser calculado explicitamente pelo *Algoritmo de Euclides Estendido*; vide Teorema A.1.

Dois ideais I e J são *co-primos* (ou *relativamente primos*) se todo ideal que contêm I e J necessariamente contém 1, isto é, se $\mathrm{I}+\mathrm{J}=\mathrm{A}$.

Lema 8.19. *Seja* A *um anel e sejam* I *e* J *ideais em* A. *Se* I *e* J *são relativamente coprimos, então*

$$\mathrm{I}\cap\mathrm{J}=\mathrm{IJ}.$$

Demonstração: Basta demonstrar que se $\mathrm{I}+\mathrm{J}=\mathrm{A}$, então

$$\mathrm{I}\cap\mathrm{J}\subseteq\mathrm{IJ}.$$

Seja a em $\mathrm{I}\cap\mathrm{J}$ e $1=i+j$. Como ai e aj são em IJ, logo $a=ai+aj$ é em IJ. □

Teorema 8.20 (Teorema Chinês dos Restos). *Seja* A *um anel e sejam* I *e* J *ideais em* A. *Se* I *e* J *são relativamente primos, então*

$$\mathrm{A}/\mathrm{IJ}\xrightarrow{\sim}\mathrm{A}/\mathrm{I}\times\mathrm{A}/\mathrm{J}.$$

Demonstração: A aplicação é injetora, porque $\mathrm{A}\to\mathrm{A}/\mathrm{I}\times\mathrm{A}/\mathrm{J}$ tem núcleo $\mathrm{I}\cap\mathrm{J}=\mathrm{IJ}$ por Lema 8.19.

A aplicação é sobrejetora, porque I e J são relativamente primos se, e tão-somente se, $\mathrm{I}+\mathrm{J}=\mathrm{A}$, isto é, existem i em I e j em J tal que $i+j=1$. Como a imagem é um ideal sobre A, é suficiente mostrar que os seus geradores $(1,0)$ e $(0,1)$ são valores. Calculamos

$$j\equiv i+j=1 \bmod \mathrm{I} \quad \text{e} \quad j\equiv 0 \bmod \mathrm{J}$$

e

$$i\equiv 0 \bmod \mathrm{I} \quad \text{e} \quad i\equiv i+j=1 \bmod \mathrm{J}.$$

Corolário 8.21 (Teorema Chinês dos Restos para os Inteiros). *Se m e n são coprimos, isto é* $\mathrm{mdc}(m,n)=1$, *então*

$$\mathbb{Z}/mn\mathbb{Z}=\mathbb{Z}/m\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z}.$$

Demonstração: Como $\mathbb{Z}$ é um domínio euclidiano, em particular, um domínio principal, as condições de Proposição 8.18 são satisfeitas e

$$\langle m\rangle\langle n\rangle=\langle \mathrm{mmc}(m,n)\rangle=\langle mn\rangle \quad \text{e} \quad \langle m\rangle+\langle n\rangle=\langle \mathrm{mdc}(m,n)\rangle=\langle 1\rangle.$$

Por Teorema B.5

$$\mathbb{Z}/\langle mn\rangle\mathbb{Z}=\mathbb{Z}/\langle m\rangle\langle n\rangle\mathbb{Z}\xrightarrow{\sim}\mathbb{Z}/\langle m\rangle\mathbb{Z}\times\mathbb{Z}/\langle n\rangle\mathbb{Z}.$$

Corolário 8.22. *Se $n = p_1^{e_1} \cdots p_n^{e_n}$ é a decomposição de n em fatores primos, então*

$$\mathbb{Z}/n\mathbb{Z}^* \xrightarrow{\sim} \mathbb{Z}/p_1^{e_1}\mathbb{Z}^* \times \cdots \times \mathbb{Z}/p_n^{e_n}\mathbb{Z}^*.$$

Demonstração: Por indução, usando que os produtos cujos fatores são dados por dois conjuntos de números primos têm um divisor comum se, e tão-somente se, a interseção dos dois conjuntos não é vazia, obtemos por Corolário 8.21

$$\mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/p_1^{e_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_n^{e_n}\mathbb{Z}.$$

Como π é um isomorfismo de anéis, em particular é um homomorfismo multiplicativo; logo

$$\mathbb{Z}/n\mathbb{Z}^* \xrightarrow{\sim} \mathbb{Z}/p_1^{e_1}\mathbb{Z}^* \times \cdots \times \mathbb{Z}/p_n^{e_n}\mathbb{Z}^*.$$

Lema 8.23. *Um elemento x em $\mathbb{Z}/p^n\mathbb{Z}$ é uma unidade se, e tão-somente se, p não divide x. Isto é,*

$$\mathbb{Z}/p^n\mathbb{Z}^* = \{x + y \in \mathbb{Z}/p^n\mathbb{Z} : x = 1, \ldots, p-1 \text{ e } y = 1, \ldots, p^{n-1}\}.$$

Demonstração: Um elemento x é uma unidade em um anel A se, e tão-somente se, o endomorfismo dado pela multiplicação $m\colon a \mapsto x \cdot a$ é sobrejetor. Se A é finito, então m é sobrejetor se, e tão-somente se, m é injetor. O endomorfismo dado pela multiplicação $m\colon a \mapsto x \cdot a$ é injetor se, e tão-somente se, x não divide 0.

Como $A = \mathbb{Z}/p^n\mathbb{Z}$ é finito, e x divide 0 se, e tão-somente se, p divide x, concluímos que x é uma unidade se, e tão-somente se, p não divide x. □

Corolário 8.24. *Se $n = p_1^{e_1} \cdots p_n^{e_n}$ é a decomposição de n em fatores primos, então*

$$\phi(n) = (p_1 - 1)p_1^{e_1-1} \cdots (p_n - 1)p_n^{e_n-1}.$$

8.5. Teoria do Corpo de Classes p-ádico

A *Teoria do Corpo de Classes* classifica as representações de dimensão 1 de $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. Todas elas fatoram através do seu quociente abeliano máximo. Vamos descrevê-lo:

Teorema (Kronecker-Weber). *A extensão* abeliana máxima *de $\mathbb{Q}_p$ (= a maior extensão em $\overline{\mathbb{Q}}_p$ cujo Grupo de Galois é abeliano) é $\mathbb{Q}_p(\mu)$ com*

$$\mu = \bigcup_{n\in\mathbb{N}} \mu_n \qquad e \qquad \mu_n = \{ \text{ todos os } \zeta \text{ em } \overline{\mathbb{Q}}_p \text{ tal que } \zeta^n = 1\}$$

as raízes da unidade. Equivalentemente, com $\cdot^{\mathrm{ab}}$ o maior quociente abeliano,

$$\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)^{\mathrm{ab}} = \mathrm{Gal}(\mathbb{Q}_p(\mu)/\mathbb{Q}_p)$$

Com

$$\mu_{p^\infty} = \bigcup_{n\in\mathbb{N}} \mu_{p^n} \quad \text{e} \quad \mu_{\nmid p} = \bigcup_{n\in\mathbb{N}} \mu_{p^n-1},$$

vale

$$\mathbb{Q}_p(\mu) = \mathbb{Q}_p(\mu_{p^\infty}) \otimes \mathbb{Q}_p(\mu_{\nmid p}).$$

Tem-se

$$\begin{aligned} \mathrm{Gal}(\mathbb{Q}_p(\mu_{p^n})/\mathbb{Q}_p) &\xrightarrow{\sim} \mathbb{Z}/p^n\mathbb{Z}^* \\ \sigma &\mapsto k \quad \text{com } \sigma(\zeta) = \zeta^k \text{ para } \zeta \text{ gerador de } \mu_{p^n} \end{aligned}$$

e

$$\begin{aligned} \mathrm{Gal}(\mathbb{Q}_p(\mu_{p^n-1})/\mathbb{Q}_p) &\xrightarrow{\sim} \mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) \xrightarrow{\sim} \mathbb{Z}/n\mathbb{Z} \\ \sigma &\mapsto k \quad \text{com } \sigma = \phi^k \text{ para } \phi = \cdot^p \text{ Frobenius} \end{aligned}$$

Logo

$$\mathrm{Gal}(\mathbb{Q}_p(\mu)/\mathbb{Q}_p) \xrightarrow{\sim} \mathbb{Z}_p^* \times \widehat{\mathbb{Z}} = \widehat{\mathbb{Q}}_p^*$$

Definimos o *Grupo de Weil* $\mathrm{Weil}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ pela imagem inversa

$$\begin{array}{ccc} \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)^{\mathrm{ab}} & \xrightarrow{\sim} & \mathbb{Z}_p^* \times \widehat{\mathbb{Z}} = \widehat{\mathbb{Q}}_p^* \\ \bigcup & & \bigcup \\ \mathrm{Weil}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)^{\mathrm{ab}} & \xrightarrow{\sim} & \mathbb{Z}_p^* \times \mathbb{Z} = \mathbb{Q}_p^* \end{array}$$

A imagem inversa de $\mathbb{Q}_p = \mathbb{Z}_p^* \times \widehat{\mathbb{Z}}$ sob o isomorfismo $\mathrm{Gal}(\mathbb{Q}_p(\mu)/\mathbb{Q}_p) \xrightarrow{\sim} \mathbb{Z}_p^* \times \widehat{\mathbb{Z}} = \widehat{\mathbb{Q}}_p^*$ é o *Grupo de Weil* $\mathrm{Weil}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. Seja $\mathbf{K}$ um *corpo de números p-ádicos*, isto é, uma extensão finita de $\mathbb{Q}_p$ e seja V um $\mathbf{K}$-espaço vetorial de dimensão finita. Investiguemos as ações sobre V quando $\dim \mathrm{V} = 1$:

Teorema (Corpo de Classes). *Para* $\mathbf{K}$ *corpo de números p-ádicos,*

$$\mathrm{Gal}(\overline{\mathbf{K}}/\mathbf{K})^{\mathrm{ab}} \xrightarrow{\sim} \widehat{\mathbf{K}}^*.$$

Corolário (Langlands para $\dim \mathrm{V} = 1$). *Dado um corpo topológico e* V *um espaço vetorial de dimensão* 1*, há um espaço vetorial* B *tal que "naturalmente"*

$$\left\{ \begin{array}{c} \textit{representações contínuas} \\ \mathrm{Weil}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \curvearrowright \mathrm{V} \end{array} \right\} \quad \overleftrightarrow{\sim} \quad \left\{ \begin{array}{c} \textit{representações contínuas} \\ \mathrm{GL}_1(\mathbb{Q}_p) \curvearrowright \mathrm{B} \end{array} \right\}$$

9. Números complexos

Para um corpo $\mathbf{K}$, denota

$$\overline{\mathbf{K}} := \text{ o fecho algébrico de } \mathbf{K};$$

isto é, o corpo dado pela reunião de todas as extensões finitas de $\mathbf{K}$.

Seja $\mathbf{K}$ um corpo não-arquimediano.

Corolário (de Teorema 6.4). *O valor absoluto sobre* $\mathbf{K}$ *estende-se de maneira única ao seu fecho algébrico* $\overline{\mathbf{K}}$.

Então, $\overline{\mathbf{K}}$ é um corpo normado (por um valor absoluto não-arquimediano). Logo, veremos que não sempre $\overline{\mathbf{K}}$ é um corpo não-arquimediano porque pode ser incompleto! Por exemplo, $\overline{\mathbb{Q}}_p$ não é completo.

Seja $\mathbf{k}$ o corpo residual de $\mathbf{K}$ e $\overline{\mathbf{k}}$ o corpo residual de $\overline{\mathbf{K}}$.

Proposição. *O corpo residual* $\overline{\mathbf{k}}$ *de* $\overline{\mathbf{K}}$ *é o fecho algébrico de* $\mathbf{k}$. *Vale* $|\overline{\mathbf{K}}^*| = \sqrt[\infty]{|\mathbf{K}^*|}$, *isto é,*

$$|\overline{\mathbf{K}}^*| = \{ \textit{ todos os } r \textit{ em }]0,\infty[\textit{ para que existe } n \textit{ em } \mathbb{N} \textit{ tal que } r^n \textit{ em } |\mathbf{K}^*| \}.$$

Demonstração: Para ver que $\overline{\mathbf{k}}$ é algébrico sobre $\mathbf{k}$, dado $\bar{x}$ em $\overline{\mathbf{k}}$, basta reduzir o polinômio (tal que o maior valor absoluto dos seus coeficientes é 1) de um levantamento x a $\overline{\mathbf{K}}$. Da mesma maneira, para ver que $\overline{\mathbf{k}}$ é algebricamente fechado, dada um polinômio $\bar{F}(X)$ em $\overline{\mathbf{k}}[X]$, basta reduzir a raiz de um levantamento $F(X)$ do polinômio a $\overline{\mathbf{K}}[X]$.

Para ver $|\overline{\mathbf{K}}^*| = \sqrt[\infty]{|\mathbf{K}^*|}$, mostramos ambas inclusões: Para $|\overline{\mathbf{K}}^*| \supseteq \sqrt[\infty]{|\mathbf{K}^*|}$, basta notar que para todo x em $\mathbf{K}$ e n em $\mathbb{N}$, existe y em $\overline{\mathbf{K}}$ tal que $y^n = x$. Por Teorema 6.4, vale $|\overline{\mathbf{K}}^*| \subseteq \sqrt[\infty]{|\mathbf{K}^*|}$. □

Por exemplo, para $\mathbf{K} = \mathbb{Q}_p$, temos $\overline{\mathbf{k}} = \overline{\mathbb{F}}_p$ e $|\overline{\mathbf{K}}^*| = p^{\mathbb{Q}}$.

Fato 9.1. *O fecho algébrico* $\overline{\mathbb{Q}}_p$ *de* $\mathbb{Q}_p$ *é incompleto.*

Seja

$$\mathbb{C}_p := \text{ o completamento do fecho algébrico } \mathbb{Q}_p$$

Lema 9.2. *Seja* $\mathbf{K}$ *completo e* f *em* $\mathbf{K}[X]$. *Se existem* $\lambda_1, \lambda_2, \ldots$ *em* $\mathbf{K}$ *tal que* $f(\lambda_1), f(\lambda_2), \ldots \to 0$, *então existe* λ *em* $\mathbf{K}$ *tal que* $f(\lambda) = 0$.

Demonstração: Podemos supor que o coeficiente do índice mais alto de f seja 1. Em particular, existem $\xi_1, \ldots, \xi_n$ em $\overline{\mathbf{K}}$ tal que $f = (x - \xi_1) \cdots (x - \xi_n)$.

Seja λ em $\overline{\mathbf{K}}$. Como $|f(\lambda)| = |\lambda - \xi_1| \cdots |\lambda - \xi_n|$, existe n_i em $\{1, \ldots, n\}$ tal que $|\lambda - x_{n_i}| \leq |f(\lambda)|$. Como $\{1, \ldots, n\}$ é finito e $\mathbb{N}$ infinito, existe n_0 em $\{1, \ldots, n\}$ e um subconjunto infinito I em $\mathbb{N}$ tal que $|\lambda_i - \xi_{n_0}| \leq |f(\lambda_i)|$ para todos os i em I. Como $f(\lambda_i) \to 0$, em particular $(\lambda_i - \xi_{n_0} : i \in \mathrm{I})$ converge a 0, isto é $(\lambda_i : i \in \mathrm{I})$ a ξ_{n_0}. Como $\mathbf{K}$ é completo, $\lim_{i \in \mathrm{I}} \lambda_i = \xi_{n_0}$ em $\mathbf{K}$. Isto é, f tem uma raiz em $\mathbf{K}$. □

Teorema ([Sch84, Theorem 17.1]). *O completamento do fecho algébrico de* $\mathbf{K}$ *é algebricamente fechado.*

Demonstração: Vamos mostrar que se $\mathbf{L}$ é um corpo algebricamente fechado e denso em um corpo completo $\mathbf{K}$, então todo polinômio $f = a_0 + a_1 \mathrm{X} + \cdots + a_{n-1}\mathrm{X}^{n-1} + \mathrm{X}^n$ em $\mathbf{K}[\mathrm{X}]$ tem uma raiz em $\mathbf{K}$.

Seja $f_1, f_2, \ldots$ uma sequência de polinômios em $\mathbf{L}[\mathrm{X}]$ de grau n que converge a f; isto é, cada coeficiente converge.

Como $\mathbf{L}$ é algebricamente fechado, obtemos raízes $\lambda_1, \lambda_2, \ldots$ de $f_1, f_2, \ldots$ em $\mathbf{L}$. Para aplicar Lema 9.2 e assim concluir a demonstração, vamos mostrar que $f(\lambda_1), f(\lambda_2), \ldots \to 0$: Como

$$f(\lambda_i) = f(\lambda_i) - f_i(\lambda_i) = [a_0 - a_0(i]) + [a_1 - a_1(i])\lambda_i + + \cdots + [a_{n-1} - a_{n-1}(i])\lambda_i^{n-1},$$

e os coeficientes de f_i convergem aos de f, obtemos que

$$f(\lambda_i) \to 0,$$

pois $|\lambda_1|, |\lambda_2|, \ldots$ é limitado pela continuidade da radiciação de um polinômio nos seus coeficientes: Para $f = a_0 + a_1 \mathrm{X} + \cdots + a_{n-1}\mathrm{X}^{n-1} + \mathrm{X}^n$ e λ uma raiz de f, vale

$$|\lambda| \leq \max\{\sqrt[n]{|a_0|},\ \sqrt[n-1]{|a_1|}, \ldots, |a_{n-1}|\}.$$

□

Corolário. *O corpo* $\mathbb{C}_p$ *não-arquimediano • é algebricamente fechado, • é de dimensão infinita sobre* $\mathbb{Q}_p$*, • tem como corpo residual* $\mathbf{k}_{\mathbb{C}_p} = \overline{\mathbb{F}}_p$*, e • tem como grupo de valores absolutos* $|\mathbb{C}_p^*| = p^{-\mathbb{Q}}$.

Demonstração: A primeira propriedade acabamos de demonstrar. As outras seguem porque, por definição, $\overline{\mathbb{Q}}_p$ é denso em $\mathbb{C}_p$, e são satisfeitas por $\overline{\mathbb{Q}}_p$. □

10. Definição da Diferenciabilidade

Recordemo-nos de que por causa da desigualdade triangular forte $\mathbb{Q}_p$ é topologicamente desconexo. Estudamos este fenômeno mais geralmente. Seja doravante **K** um tal corpo completo *não-Arquimediano.*

10.1. Funções diferenciáveis sobre os números reais

Vejamos primeiro a situação clássica sobre $\mathbb{R}$. Seja $X \subseteq \mathbb{R}$ um intervalo aberto e $f : X \to \mathbb{R}$.

Definição. Uma função f é $\mathscr{C}^1$ no ponto $x_0 \in X$ se

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exista. Declaramos que f é $\mathscr{C}^1$ se f é $\mathscr{C}^1$ em todos os pontos $x_0 \in X$ e é contínua.

Proposição 10.1. *Seja* X *compacto. O espaço* $\mathscr{C}^1(X, \mathbb{R})$ *com a norma*

$$\|f\|_{\mathscr{C}^1} = \max\{\|f\|_{\sup}, \|f'\|_{\sup}\}$$

é completo.

Demonstração: A prova habitual usa o teorema fundamental do cálculo. □

Se $\mathbb{R}$ é substituído por um corpo **K** não-arquimediano, então esta proposição é incorreta. Por isso vamos mudar a definição de derivabilidade para este enuncio ficar correto.

10.2. Patologias sobre os números p-ádicos

A priori, poderíamos definir a condição de diferenciabilidade sobre **K** como sobre $\mathbb{R}$:

Definição. Uma função $f : X \to \mathbf{K}$ sobre um subconjunto X de **K** é *diferenciável* no ponto a em X se existe $f'(a)$ em **K** tal que

$$\frac{f(x_n) - f(a)}{x_n - a} \to f'(a)$$

para toda sequência (x_n) em X tal que $x_n \to a$.

Contudo, como a topologia de $\mathbf{K}$ é totalmente desconexa, não existe *nenhum* equivalente do Teorema do Valor Intermediário, e portanto nem

- do Teorema do Valor Médio, e nem
- do Teorema Fundamental do Cálculo.

Estes dois teoremas estão no centro da Teoria de Cálculo. Por exemplo,

- O Teorema do Valor Médio mostra que a diferenciabilidade parcial contínua implica a diferenciabilidade total, e
- o Teorema Fundamental do Cálculo mostra que o espaço das funções diferenciáveis é completo com respeito à norma natural.

Por isso, encontramos entre outras as patologias seguintes:

- O espaço vetorial das funções deriváveis com a sua norma natural não é mais completo (isto é, um espaço de Banach).
- Existe uma função ([Sch84, Example 26.6])
  - cuja derivada é invertível em todo lugar, mas
  - ela mesma é localmente invertível em nenhum lugar;
- A função

$$\begin{aligned} f : \mathbb{Z}_p &\to \mathbb{Z}_p \\ \sum_{n\in\mathbb{N}} a_n p^n &\mapsto \sum_{n\in\mathbb{N}} a_n p^{2n} \end{aligned}$$

  é
  - injetora, mas a sua função derivada é 0 em todo lugar,
  - é infinitamente diferenciável mas o seu polinômio de Taylor de grau > 1 não converge;

  Demonstração: Como $|f(x+h) - f(x)| = |h|^2$, é uma função diferenciável cuja derivada f' é zero; em particular, é infinitamente diferenciável. Porém, a sua expansão do polinômio de Taylor do grau 2 não converge, por exemplo, em $a = 0$. Isto é, seja

$$\mathrm{T}(h) = f(0) + f'(0)h + f''(0)h^2 = 0$$

o polinômio de Taylor de f expandido em 0 de grau 2, e

$$\mathrm{R}(h) = f(h) - \mathrm{T}(h) = f(h)$$

o seu resto. Então $|\mathrm{R}(h)|/|h|^2 = 1$ para todo h (tal que $x+h$ e x sejam no domínio de f). Em particular, se $h \to 0$, então $|\mathrm{R}(h)|/|h|^2 \not\to 0$. □

10.3. Funções diferenciáveis sobre os números p-ádicos

Visto que não existe o teorema do valor intermediário com todas suas consequências, restringimos a condição de derivabilidade: para obter um equivalente da Proposição 10.1 e excluir estas patologias:

Definição. Sejam $\mathrm{X} \subseteq \mathbf{K}$ aberto e $f\colon \mathrm{X} \to \mathbf{K}$. Então f é $\mathscr{C}^1$ no ponto $a \in \mathrm{X}$ se o limite

$$\lim_{(x,y)\to(a,a)} f^{]1[}(x,y) \quad \text{com } f^{]1[} = \frac{f(x)-f(y)}{x-y} \quad \text{para } x,y \text{ distintos}$$

existe. Então f é $\mathscr{C}^1$ se f é $\mathscr{C}^1$ em todos os pontos $a \in \mathrm{X}$ ou igualmente se $f^{]1[}$ estende a uma função $f^{[1]}$ contínua.

Para uma função com valores reais, o Teorema do Valor Médio mostra que as condições de diferenciabilidade arquimediano e não-arquimediana são equivalentes. Com efeito, livramo-nos por esta definição da dependência do Teorema do Valor Médio.

Proposição 10.2. *A função $f \in \mathscr{C}^1(\mathrm{X},\mathbb{R})$ se, e somente se, a função*

$$f^{]1[}(x,y) = \frac{f(x)-f(y)}{x-y},$$

definida para todos $x,y \in \mathrm{X}$ desiguais, estende-se a uma função $f^{[1]}\colon \mathrm{X}\times\mathrm{X} \to \mathbb{R}$ contínua.

Demonstração: A direção $\Longleftarrow$ é evidente. Na outra direção, se $(x,y) \to (a,a) \in \mathrm{X}\times\mathrm{X}$. Então

$$f^{]1[}(x,y) = f'(\xi) \to f'(a) = f^{[1]}(a,a) \quad \text{com } \xi \in [x,y]$$

onde a primeira igualdade provém do TVM. A segunda por causa da continuidade de f'. Isto é suficiente para concluir que $f^{[1]}$ é contínua em todos os lugares como $f^{[1]}(\mathrm{X}\times\mathrm{X}) \subseteq \overline{f^{]1[}(\{(x,y) \in \mathrm{X}\times\mathrm{X} \text{ diferentes}\})}$ pela construção. □

Damos uma demonstração diferente da Proposição 10.1 acima que já indica como podemos proceder no caso arquimediano (em ausência do Teorema Fundamental do Cálculo):

Corolário. *Seja* X *compacto. O espaço* $\mathscr{C}^1(\mathrm{X},\mathbb{R})$ *é completo.*

Demonstração: Como visto acima pelo Teorema do valor médio, a norma $\|f\| = \max\{\|f\|_{\sup}, \|f^{[1]}\|_{\sup}\}$ é igual a norma

$$\|f\|_{\mathscr{C}^1} = \max\{\|f\|_{\sup}, \|f'\|_{\sup}\}.$$

Então, quanto à primeira norma, esta proposição é evidente. □

Agora fica a questão de como iterar a noção de diferenciabilidade: Como definir uma função duas vezes diferenciável? Observamos que neste caso $f^{[1]}$ é uma função em duas variáveis, ao contrário da função f' no caso real, e não podemos iterar esta definição diretamente. Então é necessário estudar o caso de muitas variáveis para definir a diferenciação repetida de uma função de uma variável só!

10.4. Funções r-vezes diferenciáveis sobre espaços p-ádicos vetoriais

Sejam V e **E** dois espaços de Banach sobre **K** e X um subconjunto aberto de V.

Definição da diferenciabilidade de grau 1.

Recordemo-nos da definição de uma função derivável de múltiplos argumentos.

Definição. Sejam V e **E** espaços vetoriais, $\mathrm{X} \subseteq \mathrm{V}$ aberto e $f : \mathrm{X} \to \mathbf{E}$. Então f é $\mathscr{C}^1$ no ponto $a \in \mathrm{X}$ se existe uma aplicação linear contínua A tal que para todos $\varepsilon > 0$ existe $\mathrm{U} \ni a$ aberto em X tal que

$$f(x+h) - f(x) = \mathrm{A}\cdot h + \mathrm{R}(x+h, x)$$

onde o resto satisfaz $\|\mathrm{R}(x+h,x)\| \leq \varepsilon\|h\|$ para todos $x+h, x \in \mathrm{U}$.

Diferenciabilidade iterada.

Para "iterar" a definição de diferenciabilidade não-arquimediana, introduzamos coordenadas sobre V pela escolha duma base (ordenada) $(e_1, \ldots, e_d)$ de V.

Definição. A *diferença dividida* $f^{]1[}(x+h,x)$ duma função $f : \mathrm{X} \to \mathbf{E}$ no ponto $x+h,x$ em X com h em $\mathbf{K}^{*d}$ é a aplicação linear A definida por

$$\mathrm{A}e_k := frac f(x+h_1e_1+\cdots+h_{k-1}e_{k-1}+h_ke_k) - f(x+h_1e_1+\cdots+h_{k-1}e_{k-1})h_k$$

para todo $k = 1,\ldots,d$. A função f é uma $\mathscr{C}^1$-*função* se $f^{]1[}$ se estende a uma função contínua $f^{[1]} : \mathrm{X}\times\mathrm{X} \to \mathbf{E}$.

Isto não permite diretamente obter uma definição de diferenciabilidade geral, mas possibilita uma boa perspectiva como proceder em geral.

Comparemos o domínio e codomínio de $f^{[1]}$ com os de f:

- O domínio $\mathrm{X}^{[1]} := \mathrm{X}\times\mathrm{X}$ de $f^{[1]}$ é contido no espaço vetorial $\mathrm{V}^{[1]} = \mathrm{V}\times\mathrm{V}$ com uma base (ordenada) canônica, como o domínio X de f, e
- o codomínio $\mathbf{E}^{[1]} := \mathrm{Hom}_{\mathbf{K}}(\mathrm{V},\mathbf{E})$ de $f^{[1]}$ é um espaço de Banach, como o codomínio $\mathbf{E}$ de f.

Podemos então "iterar" a definição de diferenciabilidade não-arquimediano se aplicamos a condição à $f^{[1]}$ no lugar de f:

Definição. Dizemos que $f : \mathrm{X} \to \mathbf{E}$ é $\mathscr{C}^2$ se f é $\mathscr{C}^1$ e $f^{[1]} : \mathrm{X}\times\mathrm{X} \to \mathrm{Hom}_{\mathbf{K}}(\mathrm{V},\mathbf{E})$ é $\mathscr{C}^1$. Geralmente, f é $\mathscr{C}^n$ se f é $\mathscr{C}^{n-1}$ e $f^{[n-1]}$ é $\mathscr{C}^1$.

Com esta definição, podemos compreender melhor as propriedades destas funções. Pois esta definição é complicada e não tínhamos até este momento muita teoria sobre a diferenciabilidade, mesmo a verificação das propriedades naturais exige muita atenção.

Diferenciabilidade fracionária. Na teoria das representações de grupos de Lie p-ádicos, a seguinte noção tem um papel importante: Seja $r \geq 0$ um número real; escrevamos $r = \nu + \rho$ com ν inteiro e ρ no intervalo $[0,1[$.

Definição. Sejam X e Y espaços métricos e $f : \mathrm{X} \to \mathrm{Y}$. A função f é uma $\mathscr{C}^\rho$-*função* se para todo ponto a em X e todo $\epsilon > 0$ existe uma vizinhança U em volta de a contida em X tal que

$$d(f(x),f(y)) \leq \epsilon d(x,y)^\rho$$

para todos os x e y em U.

Sejam V um espaço vetorial e $\mathbf{E}$ um espaço de Banach. Seja X um subconjunto aberto de V e $f : \mathrm{X} \to \mathbf{E}$. Apliquemos a condição de diferenciabilidade fracionária à diferença dividida iterada de f:

Definição. A função f é uma $\mathscr{C}^r$-*função* se é uma $\mathscr{C}^\nu$-função e $f^{[\nu]}$ é uma $\mathscr{C}^\rho$-função.

11. Funções Diferenciáveis de uma Variável

No caso de uma variável podemos dar descrições mais fáceis das funções r-vezes diferençáveis para $r \geq 0$.

11.1. Diferenças divididas iteradas

A definição anterior é bem adequada para questões conceituais. Para computações, a definição do livro-texto (vide [Sch84, Seção 26ff.]) é mais apropriada:

Schikhof observou que a diferença dividida $f^{]1[}$ é uma função simétrica; como tal, é diferenciável se, e somente se, é parcialmente diferenciável na sua primeira coordenada. Isso reduz, com o aumento do grau de diferenciabilidade ν, o crescimento exponencial do número de variáveis de $f^{]\nu[}$ a um crescimento linear no número de variáveis de uma diferença dividida $f^{>\nu<}$, que definimos abaixo:

Definição. Seja X um subconjunto de $\mathbf{K}$e $f : X \to \mathbf{E}$. Para $\nu \in \mathbb{N}$ põe

$$X^{<\nu>} = X^{\{0,\ldots,\nu\}} \quad \text{e} \quad X^{>\nu<} = \{(x_0,\ldots,x_\nu) : \text{se} u \neq j \text{então} x_i \neq x_j\}.$$

A *ν-ésima diferença dividida* $f^{>\nu<} : X^{>\nu<} \to \mathbf{E}$ de uma função $f : X \to \mathbf{E}$ é indutivamente dada por $f^{>0<} := f$ e por $n \in \mathbb{N}$e$(x_0,\ldots,x_\nu) \in X^{>\nu<}$ por

$$f^{>\nu<}(x_0,\ldots,x_\nu) := \frac{f^{>\nu-1<}(x_0,x_2,\ldots,x_\nu) - f^{>\nu-1<}(x_1,x_2,\ldots,x_\nu)}{x_0 - x_1}.$$

A seguinte definição para $\rho = 0$ é dada em[Sch84, Seção 29], onde diferenciabilidade *integral* (isto é, para ν em $\mathbb{N}$) é definida. Isto é, uma função f é ν vezes diferenciável se $f^{>\nu<}$ se estende a uma função contínua em $X^{<\nu>}$.

Definição 11.1. Põe $r = \nu + \rho \in \mathbb{R}_{\geq 0}$. Seja X um subconjunto de $\mathbf{K}$ e $f : X \to \mathbf{E}$.

- A função f é $\mathscr{C}^r$ (ou *r vezes diferenciável*) em um ponto $a \in X$ se $f^{>\nu<} : X^{>\nu<} \to \mathbf{E}$ é $\mathscr{C}^\rho$ em $\vec{a} = (a,\ldots,a) \in X^{<\nu>}$.
- A função f é uma *$\mathscr{C}^r$-função* (ou uma função *r vezes diferenciável*) se f é $\mathscr{C}^r$ em todo a em X. Denote $\mathscr{C}^r(X,\mathbf{E})$ o conjunto de todas as $\mathscr{C}^r$-funções $f : X \to \mathbf{E}$.

Note que esta condição de diferenciabilidade é, mesmo para ordens mais altas, dada em pontos. Se a é um ponto de acumulação, então o valor $D^\nu f(a)$ ao qual $f^{>\nu<}$ se estende em $\vec{a}$, o derivado de f em a, é determinado univocamente.

Se $f^{(\nu)}$ é a ν-ésima derivada ordinária de f, então $\nu!\,\mathrm{D}^\nu f = f^{(\nu)}$ ([Sch84, Theorem 29.5]).

Contenha X nenhum ponto isolado. Então f é uma função $\mathscr{C}^r$ se e somente se $f^{>\nu<}$ estender para uma *única* $\mathscr{C}^\rho$-função $f^{<\nu>}\colon \mathrm{X}^{<\nu>} \to \mathbf{E}$ ([Nag11, Proposição 2.5]).

Cada função r vezes diferenciável é (por [Nag11, Lemma 2.3]) em particular s vezes diferenciáveis para cada $s \leq r$ não-negativo. Portanto, se X é compacto sem pontos isolados, então podemos equipar o espaço vetorial sobre $\mathbf{K}$ de $\mathscr{C}^r$-funções com a norma

$$\|f\|_{\mathscr{C}^r} := \max\{\|f^{[0]}\|_{\sup}, \ldots, \|f^{[\nu-1]}\|_{\sup}, \|f^{[\nu]}\|_{\mathscr{C}^\rho}\}.$$

11.2. Polinómio de Taylor

Damos uma condição de diferenciabilidade de apenas dois argumentos por polinômios de Taylor. (Enquanto a por diferenciais lineares iterados respectivamente por diferenças divididas iteradas, têm um crescimento exponencial respectivamente linear no número de variáveis ao aumentar o grau de diferenciabilidade ν).

Definição. Seja V um espaço vetorial $\mathbf{K}$-vetorial. Seja $\mathrm{Sym}^n_{\mathbf{K}}(\mathrm{V},\mathbf{E})$ o conjunto de todas as aplicações contínuas *simétricas* $\mathbf{K}$-lineares $\mathrm{M}\colon \mathrm{V}\times\cdots\times\mathrm{V} \to \mathbf{E}$ de n variáveis. Elas formam um espaço de $\mathbf{K}$-Banach não-arquimediano pela norma do operador

$$\|\mathrm{M}\| = \sup\{\|\mathrm{M}(x)\| : x \in \mathrm{V}^n \mathrm{com} \|x\| \leq 1\}$$

dada pelo supremo de M sobre a bola unitária de $\mathrm{V}\times\cdots\times\mathrm{V}$ para a norma do produto $\|v_1,\ldots,v_n\| = \max\{\|v_1\|,\ldots,\|v_n\|\}$.

Definição 11.2. Seja X um subconjunto aberto de V. A função $f\colon \mathrm{X} \to \mathbf{E}$ é uma função $\mathscr{C}^r_{\mathrm{T}}$ se houver funções $\mathrm{D}^n f\colon \mathrm{X} \to \mathrm{Sym}^n(\mathrm{V},\mathbf{E})$ para $n = 0,1,\ldots,\nu$ e $\mathrm{R}^v f\colon \mathrm{X}\times\mathrm{X} \to \mathbf{E}$ tal que

$$f(x+h) = \sum_{n=0,\ldots,\nu} \mathrm{D}^n f(x)(h,\ldots,h) + \mathrm{R}^\nu f(x+h,x)$$

e para cada a em X e $\varepsilon > 0$, existe uma vizinhança U em torno de a dentro de X tal que

$$\|\mathrm{R}^\nu f(x+h,x)\| \leq \varepsilon\|h\|^r \quad \text{para todos os } x+h, x \text{ em U.}$$

A norma. Seja $\mathscr{C}^r_{\mathrm{T}}(\mathrm{X},\mathbf{E})$ o espaço $\mathbf{K}$-vetorial de todas as $\mathscr{C}^r_{\mathrm{T}}$-funções $f\colon \mathrm{X}\to\mathbf{E}$. Por [Nag16, Corolário 2.5] as funções $\mathrm{D}^0f,\mathrm{D}^1f,\ldots,\mathrm{D}^\nu f$ são unicamente determinadas e diferenciáveis de grau r, $r-1,\ldots,$ ρ. Conseqüentemente

1. em particular, as funções $\mathrm{D}^0f,\mathrm{D}^1f,\ldots,\mathrm{D}^\nu f$ são contínuas, e
2. o temo do resto $\mathrm{R}^\nu f$ do polinômio de Taylor até grau ν converge como em Definição 11.2 se e somente se a função $\Delta^r f$, definida por

$$\Delta^r f(x,y)=\|\mathrm{R}^\nu f(x,y)\|/\|xy\|^r \quad \text{para todos os distintos} x,y \text{in} \mathrm{X},$$

se estende a uma função contínua $|\Delta^r f|\colon \mathrm{X}\times\mathrm{X}\to\mathbb{R}_{\geq 0}$ que desvanece na diagonal.

Assim, se X é um subconjunto aberto compacto de V, então existe uma norma bem definida $\|\cdot\|_{\mathscr{C}^r_{\mathrm{T}}}$ em $\mathscr{C}^r_{\mathrm{T}}(\mathrm{X},\mathbf{E})$ dado por $\|f\|_{\mathscr{C}^r_{\mathrm{T}}}:=\max\{\|\mathrm{D}^0f\|_{\sup},\ldots,\|\mathrm{D}_nf\|\cup\{\||\Delta^r f|\|_{\sup}\}$.

Necessidade. Cada função r vezes diferenciável pode ser aproximada localmente por sua expansão polinomial de Taylor até o grau ν: Mais exatamente, temos o seguinte critério de suficiência para uma função ser uma função $\mathscr{C}^r_{\mathrm{T}}$:

Proposição 11.3 ([Nag16, Corolário 3.6]). *Temos $\mathscr{C}^r(\mathrm{X},\mathbf{E})\subseteq\mathscr{C}^r_{\mathrm{T}}(\mathrm{X},\mathbf{E})$ e se X é um subconjunto compacto aberto de V, então a inclusão $\mathscr{C}^r(\mathrm{X},\mathbf{E})\hookrightarrow\mathscr{C}^r_{\mathrm{T}}(\mathrm{X},\mathbf{E})$ é um monomorfismo de espaços vetoriais normados.*

Propriedades de $\mathscr{C}^r$-funções.

Lema 11.4. *Seja $\mathrm{X}\subseteq\mathbf{K}$ um subconjunto, a algum ponto em X e $f\colon\mathrm{X}\to\mathbf{K}$ uma aplicação. Se f é $\mathscr{C}^r$ em a, então f é $\mathscr{C}^s$ em a para todo $s\leq r$.*

Demonstração: Se f for $\mathscr{C}^r$ em a, então claramente f será $\mathscr{C}^s$ em a para cada $\nu\leq s\leq r$. Por transitividade, é suficiente provar que f é $\mathscr{C}^s$ em r com $s=\nu-1+\eta$ para $\eta\in[0,1[$. Usamos a caracterização por **??**: Em $\mathrm{X}^{]\nu-1[}$ retém para $x_0,\tilde{x}_0\in\mathrm{X}$ por definição

$$\begin{aligned}&|f^{]\nu-1[}(x_0,x_1,\ldots,x_\nu)-f^{]\nu-1[}(\tilde{x}_0,x_1,\ldots,x_\nu)|\\ =&|x_0-\tilde{x}_0||f^{]\nu[}(x_0,\tilde{x}_0,x_1,\ldots,x_\nu)|\\ =&|x_0-\tilde{x}_0|^\eta|x_0-\tilde{x}_0|^{1-\eta}|f^{]\nu[}(x_0,\tilde{x}_0,x_1,\ldots,x_\nu)|.\end{aligned}$$

Se agora f for $\mathscr{C}^r$ em $a \in X$, então $f^{]\nu[}$ será $\mathscr{C}^0$ em $\vec{a}$ e em particular localmente limitada por uma constante $C > 0$.

Fixa $\varepsilon > 0$! Existe uma vizinhança $V \ni a$ em X com $|x - \tilde{x}|^{1-\eta} \leq \varepsilon/C$ para todo $x, \tilde{x} \in U$. Daí na vizinhança $U \cap X^{]\nu-1[}$ com $U := V^{[\nu-1]} \ni \vec{a}$ em $X^{[\nu]}$,

$$|f^{]\nu-1[}(x_0, x_1, \ldots, x_\nu) - f^{]\nu-1[}(\tilde{x}_0, x_1, \ldots, x_\nu)| \leq |x_0 - \tilde{x}_0|^\eta \varepsilon.$$

Por **??**, isso prova que f é $\mathscr{C}^s$ em a. □

Lema 11.5. *Seja* $X \subseteq \mathbf{K}$ *um subconjunto não vazio sem pontos isolados e mapeamento* $f : X \to \mathbf{K}$ *a. Suponha que, para* $r = \nu + \rho \in \mathbb{R}_{\geq 1}$, *o mapa* $f^{]\nu-1[} : X^{]\nu-1[} \to \mathbf{K}$ *é* $\mathscr{C}^\rho$ *em todo* $X^{[\nu-1]}$. *Então* $f^{]\nu[} : X^{]\nu[} \to \mathbf{K}$ *pode ser estendido para uma função* $\mathscr{C}^\rho$ - $f^{<\nu>} : X^{[\nu]} - \triangle X^{[\nu]} \to \mathbf{K}$.

Demonstração: Para $i, j \in \{0, \ldots, \nu\}$ com $i \neq j$ set

$$U_{ij} = \{(x_0, \ldots, x_\nu) \in X^{[\nu]} : x_i \neq x_j\}.$$

Então cada U_{ij} é aberto em $X^{[\nu]}$ e sua união é $X^{[\nu]} - \triangle X^{[\nu]}$. Por causa de nossa suposição em $f^{]\nu-1[}$, encontramos por **??** que $f^{]\nu-1[}$ se estende a um $\mathscr{C}^\rho$ - function $f^{[\nu-1]} : X^{[\nu-1]} \to \mathbf{K}$. Podemos, portanto, definir $h_{ij} : U_{ij} \to \mathbf{K}$ por

$$\begin{aligned} &h_{ij}(x_0, \ldots, x_\nu) \\ =&\frac{f^{[\nu-1]}(x_0, \ldots, \check{x}_j, \ldots, x_\nu) - f^{[\nu-1]}(x_0, \ldots, \check{x}_i, \ldots, x_\nu)}{x_i - x_j}; \end{aligned}$$

aqui os argumentos abaixo dos breves sendo omitidos. Pela simetria de $f^{]\nu[}$ e $f^{]\nu-1[}$, encontramos $x \in X^{]\nu[}$

$$\begin{aligned} f^{]\nu[}(x) &= f^{]\nu[}(x_i, x_j, x_2, \ldots, \overbrace{x_0}^{i\text{-th lugar}}, \ldots, \overbrace{x_1}^{j\text{-th place}}, \ldots, x_\nu) \\ &= [\quad f^{]\nu-1[}(x_i, x_2, \ldots, \overbrace{x_0}^{i\text{-th lugar}}, \ldots, \overbrace{x_1}^{j\text{-th place}}, \ldots, x_\nu) \\ &\quad - f^{]\nu-1[}(x_j, x_2, \ldots, \underbrace{x_0}_{i\text{-th lugar}}, \ldots, \underbrace{x_1}_{j\text{-th place}}, \ldots, x_\nu)]/[x_i - x_j] \\ &= \frac{f^{]\nu-1[}(x_0, x_2, \ldots, \overbrace{x_1}^{j\text{-th lugar}}, \ldots, x_\nu) - f^{]\nu-1[}(x_1, x_2, \ldots, \overbrace{x_0}^{i\text{-th lugar}}, \ldots, x_\nu)}{x_i - x_j} \end{aligned}$$

$$= \frac{f^{]\nu-1[}(x_0,\ldots,\check{x}_j,\ldots,x_\nu) - f^{]\nu-1[}(x_0,\ldots,\check{x}_i,\ldots,x_\nu)}{x_i - x_j} = h_{ij}(x);$$

portanto, cada h_{ij} estende $f^{]\nu[}$. Como $(x_i - x_j)^{-1}$ é uma função $\mathscr{C}^\rho$ - em U_{ij} e também $f^{[\nu-1]}$ em $X^{[\nu-1]}$, o mesmo vale para o nosso mapa h_{ij} por **??**(ii). Colamos essas funções juntas, colocando

$$f^{<\nu>}(x) = h_{ij}(x) \quad \text{se } x \in U_{ij}.$$

Então $f^{<\nu>}\colon X^{[\nu]} - \triangle X^{[\nu]} \to \mathbf{K}$ é uma função bem definida como todas as funções contínuas h_{ij} coincidem no subconjunto denso comum $X^{]\nu[}$ de seus domínios. Para $\mathscr{C}^\rho$ ser uma propriedade local, $f^{<\nu>}$ também é uma função $\mathscr{C}^\rho$. □

Proposição 11.6. *Seja* $X \subseteq \mathbf{K}$ *um subconjunto não vazio sem pontos isolados e mapeamento* $f\colon X \to \mathbf{K}$ *a. Então* $f \in \mathscr{C}^r(X,\mathbf{K})$ *se e somente se* $f^{]\nu[}\colon X^{]\nu[} \to \mathbf{K}$ *estende para* $\mathscr{C}^\rho$ *- função* $f^{[\nu]}\colon X^{[\nu]} \to \mathbf{K}$.

Demonstração: Em primeiro lugar, notamos que se $f^{]\nu[}$ se estende a uma função $\mathscr{C}^\rho$ - $f^{[\nu]}\colon X^{[\nu]} \to \mathbf{K}$, então será em particular $\mathscr{C}^\rho$ em $\vec{a} \in X^{[\nu]}$, de modo que nós só temos que mostrar a parte "somente se" . Isso é provado pela indução em ν. Para $\nu = 0$, isso vale por definição, então vamos supor que $\nu \geq 1$ e que a afirmação seja verdadeira para $n - 1$. Por Lema 11.4, encontramos $f \in C^{r-1}(X,\mathbf{K})$. Por nossa hipótese de indução, sabemos que $f^{]\nu-1[}$ se estende a uma função $\mathscr{C}^\rho$ - $f^{[\nu-1]}\colon X^{[\nu]} \to \mathbf{K}$. Por Lema 11.5, a função $f^{]\nu[}$ se estende a uma função $\mathscr{C}^\rho$ - $f^{<\nu>}$ em todos os $X^{[\nu]} - \triangle X^{[\nu]}$. Agora podemos estender $f^{]\nu[}$ a $X^{[\nu]}$ definindo

$$f^{[\nu]}(a_0,\ldots,a_\nu) = \begin{cases} f^{<\nu>}(a_0,\ldots,a_\nu), \text{ se } (a_0,\ldots,a_\nu) \in X^{[\nu]} - \triangle X^{[\nu]}, \\ \lim_{y\to\vec{a}} f^{]\nu[}(v), \text{ se } \vec{a} = (a,\ldots,a) \in \triangle X^{[\nu]}, \end{cases}$$

com y passando por $X^{]\nu[}$. Assim, se permitirmos que $A := X^{]\nu[}$ e $A \subseteq B := X^{[\nu]} \subseteq \overline{A}$, então em particular $f^{]\nu[}\colon A \to \mathbf{K}$ será uma função que é $\mathscr{C}^\rho$ em todo o B e, portanto, sua extensão contínua única $f^{[\nu]}\colon X^{[\nu]} \to \mathbf{K}$ é uma função $\mathscr{C}^\rho$ por **??**. □

Corolário 11.7. *Seja* $X \subseteq \mathbf{K}$ *um subconjunto não vazio sem pontos isolados e* $f \in \mathscr{C}^r(X,\mathbf{K})$. *Então as funções*

$$D_i f(a) := f^{[i]}(\vec{a}) \quad \textit{paraum} \in X$$

estão em $\mathscr{C}^{r-i}(X,\mathbf{K})$ *para* $i = 0,\ldots,\nu$.

Demonstração: Por [**?**, Lema 78.1], vale para todo $x = (x_0, \ldots, x_i) \in X^{]i[}$ que

$$(D_{\nu-i} f)^{]i[}(x) = \sum_{y \in S_{i,\nu}} f^{[\nu]}(y),$$

onde $S_\nu(x)$ é o conjunto de todas as tuplas $(x_{m_0}, \ldots, x_{m_\nu}) \in X^{[\nu]}$ para as quais $m_0 \leq \ldots \leq m_\nu$ e $\{m_0, \ldots, m_\nu\} = \{0, \ldots, i\}$. Porque cada tupla $y(x) \in S_\nu(x)$ como uma função $X^{]i[} \to X^{[\nu]}$ é apenas repetição de coordenadas, é localmente Lipschitzian, e **??**(i), (ii) nos diz que o lado direito da equação define uma função $\mathscr{C}^\rho$ - em $X^{]i[}$, produzindo $D_{\nu-i} f \in \mathscr{C}^{i+rho}(X, \mathbf{K})$. Em outras palavras, $D_i f \in \mathscr{C}^{r-i}(X, \mathbf{K})$ para $i = 0, \ldots, \nu$. □

Nota 11.8. (cf.[**?**, Teorema 29.5]) Seja $X \subseteq \mathbf{K}$ um subconjunto não vazio sem pontos isolados. Se $f \in \mathscr{C}^\nu(X, \mathbf{K})$ então f será ν vezes continuamente diferenciável no sentido de Arquimedes e nós temos $\nu! D_\nu f = f^{(\nu)}$, onde $f^{(\nu)}$ denota a derivada arquimedeana ν -fold de f.
No entanto, existem algumas sutilezas na característica $p > 0$: A função $f(x) = x^2$ é $\mathscr{C}^2$ sobre $\mathbf{K} = \mathbb{F}_2((t))$ que satisfaz $D_1 f \equiv 0$, mas $D_2 f \equiv 1$.

12. Séries de Potências

Estudemos séries de potências sobre um corpo não-arquimediano. Como se trata antes de tudo de séries de potências formais, a teoria segue primeiro analogamente à sobre um corpo arquimediano, isto é, sobre $\mathbb{R}$ ou $\mathbb{C}$, mas surgem novos fenómenos topológicos, por exemplo, vemos que devido ao valor p-ádico, o domínio de convergência da exponencial e do logaritmo são menores.

12.1. Convergência Uniforme

Seja X um conjunto e (Ω, ρ) um espaço métrico. Sejam $f_1, f_2, \ldots : X \to \Omega$ funções. A sequência (f_n) converge *uniformemente* a f, denotado por $f_n \to f$ uniformemente, se para todo $\epsilon > 0$ existe N tal que $\rho(f_n(x), f(x)) < \epsilon$ para todo x em X e $n \geq N$; isto é,

$$\sup\{\rho(f_n(x), f(x)) : x \in X\} < \epsilon \qquad \text{para todo } n \geq N.$$

Teorema 12.1. *Seja* X *um espaço métrico. Se todos os f_n são contínuas e $f_n \to f$ uniformemente, então f é contínua.*

Demonstração: Seja x_0 em X e $\epsilon > 0$. Mostremos que existe $\delta > 0$ tal que $d(x_0, x) < \delta$ implica $\rho(f(x_0), f(x)) < \epsilon$ para todo x em X.

Seja N tal que $\rho(f_n(x), f(x)) < \epsilon$ para todo x em X e $n \geq N$. Como f_N é contínua, em particular em x_0, existe $\delta > 0$ tal que $d(x_0, x) < \delta$ implica $\rho(f_N(x_0), f_N(x)) < \epsilon$ para todo x em X. Estimemos

$$\rho(f(x_0), f(x)) \leq \rho(f(x_0), f_N(x_0)) + \rho(f_N(x_0), f_N(x)) + \rho(f_N(x), f(x)) \leq 3\epsilon$$

Sejam $f_1, f_2, \ldots : X \to \Omega$ funções. A sequência (f_n) é uniformemente *Cauchy*, se para todo $\epsilon > 0$ existe N tal que $\rho(f_n(x), f_m(x)) < \epsilon$ para todo x em X e $n, m \geq N$.

Observação. Seja Ω um espaço vetorial normado. Seja B(X) o conjunto de todas as funções $f : X \to \Omega$ limitadas, isto é, tais que

$$\|f\|_\infty := \sup\{\|f(x)\| : x \in S\} < \infty$$

com a métrica $d_\infty(f, g) = \|f - g\|_\infty$. (Vide o espaço métrico B(S) de **??**.) Uma sequência (f_n) é convergente respectivamente Cauchy em B(X) se, e tão-somente se, ela é uniformemente convergente respectivamente uniformemente Cauchy.

Proposição 12.2. *Seja* X *um conjunto e* (Ω, ρ) *um espaço métrico. Seja* $f_1, f_2, \ldots : X \to \Omega$ *uma sequência de funções uniformemente Cauchy. Se* Ω *é completo, então* (f_n) *converge uniformemente a uma função* $f : X \to \Omega$.

Demonstração: Para todo x a sequência $f_n(x)$ é Cauchy, logo existe um limite. Defina $f : X \to \Omega$ por estes limites.

Seja $\epsilon > 0$ e x em X. Seja N tal que $\rho(f_n(x), f_m(x)) < \epsilon$ para todo x em X e $n, m \geq N$; Pela continuidade da métrica

$$\begin{aligned}\rho(f(x), f_n(x)) &= \rho(\lim_m f_m(x), f_n(x)) \\ &= \lim_m \rho(f_m(x), f_n(x)) \leq \max\{\rho(f_m(x), f_n(x)) : m \geq N\} < \epsilon;\end{aligned}$$

isto é, $f_n \to f$ uniformemente. □

Se f_1, f_2, ... são todas contínuas, então f é contínua por Teorema 12.1.

Definição. Seja Ω um espaço métrico que é um grupo abeliano; por exemplo, um espaço vetorial normado sobre $\mathbb{R}$ ou $\mathbb{C}$; em particular, $\Omega = \mathbb{R}$ ou $\mathbb{C}$.

Uma *série* $\sum u_n$ de funções u_1, u_2, ..., $: X \to \Omega$ *converge uniformemente* se a sequência (s_n) dos seus truncamentos finitos $s_n = u_1 + \cdots + u_n$ converge uniformemente.

Teorema 12.3 (Teste de Weierstrass). *Sejam* $u_n : X \to \mathbb{C}$ *funções e* $M_n > 0$ *constantes para* $n \in \mathbb{N}$. *Se* $\|u_n\| = \sup\{|u_n(x)| : x \in X\} \leq M_n$ *para todo* n *e* $M_1 + M_2 + \cdots < \infty$, *então a série* $\sum_{n\in\mathbb{N}} u_n$ *converge uniformemente.*

Demonstração: Seja (s_n) a sequência dos truncamentos finitos $s_n = u_1 + \cdots + u_n$ de $\sum u_n$. Como $\|u_n\| = \sup\{|u_n(x)| : x \in X\} \leq M_n$ para todo n e $M_1 + M_2 + \cdots < \infty$, a sequência s_n é uniformemente Cauchy. Logo, por Proposição 12.2, ela converge uniformemente. □

12.2. Séries de Potências

Seja **K** um corpo com um *valor absoluto.*

Definição 12.4. Uma *série* $\sum_n a_n = a_0 + a_1 + a_2 + \cdots$ com a_0, a_1, ... em **K** *converge* se a sequência (s_n) dos seus truncamentos $s_n = a_1 + \cdots + a_n$ converge (em **K**). Ela *converge absolutamente* se $\sum_n |a_n|$ converge (em $[0, \infty[$).

Proposição 12.5. *Seja* $\sum_n a_n$ *uma série. Se ela converge absolutamente e* **K** *é completo, então converge.*

Demonstração: Como **K** é completo, dada uma série $\sum_n u_n$, observemos q a sequência (s_n) dos seus truncamentos finitos $s_n = u_1 + \cdots + u_n$ converge se, e tão-somente se, $s_{n,\dots,m} = u_n + \cdots + u_m \to 0$ para $n \to \infty$; isto é, (s_n) converge se, e tão-somente se, para todo $\epsilon > 0$ existe N tal que $|s_{n,\dots,m}| < \epsilon$ para todo $n, m \geq \mathrm{N}$.

Pela desigualdade triangular, $|a_n + \cdots + a_m| \leq |a_n| + \cdots + |a_m|$. Logo, se $|a_n| + \cdots + |a_m| \to 0$, então $|a_n + \cdots + a_m| \to 0$. Logo, pela observação, a sequência dos truncamentos finitos $s_n = a_1 + \cdots + a_n$ converge; isto é, $\sum_n a_n$ converge. □

Se **K** é não-arquimediano, então, por Proposição 4.1.(iii), a série $\sum_n a_n$ converge se, e somente se, $a_n \to 0$.

Para (a_n) uma sequência em **K**, denote

$$\liminf a_n \coloneqq \lim_{n\to\infty} \inf\{a_n : n \in \mathbb{N}\} \quad \text{e} \quad \limsup a_n \coloneqq \lim_{n\to\infty} \sup\{a_n : n \in \mathbb{N}\};$$

ou, alternativamente,

$$\underline{\lim}\, a_n \coloneqq \liminf a_n \quad \text{e} \quad \overline{\lim}\, a_n \coloneqq \limsup a_n.$$

Os limites são monotonamente crescentes respectivamente decrescentes; por isso sempre existem (se incluímos a possibilidade dos limites $\pm\infty$).

Uma *série de potências à volta de a* em **K** é uma *série de potências formal* da forma

$$\sum_{n\in\mathbb{N}} a_n(\mathrm{Z} - a)^n;$$

isto é, uma sequência de polinómios $(s_\mathrm{N}(\mathrm{Z}) : \mathrm{N} \in \mathbb{N})$ com $s_\mathrm{N}(\mathrm{Z}) = \sum_{n=0,\dots,\mathrm{N}} a_n(\mathrm{Z}-a)^n$ em $\mathbf{K}[\mathrm{Z}]$.

Se substituirmos a incógnita Z por um elemento z em **K**, obtemos a série $\sum_{n\in\mathbb{N}} a_n(z-a)^n$ em **K**, isto é, a sequência de somas truncadas (s_N) com $s_\mathrm{N} = \sum_{n=0,\dots,\mathrm{N}} a_n(z-a)^n$ em **K**.

Se $\sum_{n\in\mathbb{N}} a_n(z-a)^n < \infty$, isto é, se a sequência (s_N) converge, para todo z em um subconjunto X em K, então obtemos uma função

$$\begin{aligned} \mathrm{D} &\mapsto \mathbf{K} \\ z &\mapsto \sum_{n\in\mathbb{N}} a_n(z-a)^n. \end{aligned}$$

Para distinguir entre estas três interpretações de uma série de potências, usaremos (que o leitor atento nos corrija em cada caso contrário!)

- uma letra *maiúscula* como Z para destacar que se trata de uma *série de potências formal*,
- uma letra *minúscula* como z para destacar que se trata de uma *série* em $\mathbf{K}$, e
- a notação $\sum_{n\in\mathbb{N}} a_n(\cdot - a)^n$ ou $z \mapsto \sum_{n\in\mathbb{N}} a_n(z-a)^n$ para destacar que se trata de uma *função* sobre um subconjunto de $\mathbf{K}$.

Definição. Seja X em $\mathbf{K}$ aberto. Uma função $f : \mathrm{X} \to \mathbf{K}$ é *analítica* se existe a em X e uma série de potências formal $\sum_{n\in\mathbb{N}} a_n(\mathrm{Z}-a)^n$ tal que $f(z) = \sum_{n\in\mathbb{N}} a_n(z-a)^n$.

Em um corpo não-arquimediano, todo ponto de uma bola é o seu centro. De fato, todo ponto pode ser o ponto de expansão de uma função analítica:

Teorema 12.6. *Seja* X *em* $\mathbf{K}$ *aberto e* $f : \mathrm{X} \to \mathbf{K}$ *uma função analítica. Para todo* b *em* X *existe uma série de potências formal* $\sum_{n\in\mathbb{N}} b_n(\mathrm{Z}-a)^n$ *tal que* $f(z) = \sum_{n\in\mathbb{N}} b_n(z-b)^n$.

Demonstração: A demonstração é dada em [Sch84, Theorem 25.1]. □

12.3. Raio de convergência

Toda série de potências $\sum_{n\in\mathbb{N}} a_n(\mathrm{Z}-a)^n$ é uma função (com valores em $\mathbf{K}$) definida sobre uma bola aberta em $\mathbf{K}$ à roda de a cujo raio calculemos agora:

Por exemplo, a *série geométrica* é a série de potências (em torno de 0)

$$\sum_{n\in\mathbb{N}} \mathrm{Z}^n.$$

Pela *soma telescópica*,

$$(1-\mathrm{Z})(1+\mathrm{Z}+\cdots+\mathrm{Z}^n) = 1-\mathrm{Z}^{n+1}$$

Logo, se $|z|<1$, então

$$\sum_n z^n = \frac{1}{1-z},$$

e se $|z|>1$, então $\sum_n z^n$ diverge.

Teorema 12.7. *Para uma série de potências* $\sum_n a_n(\mathrm{Z}-a)^n$ *define o seu* raio de convergência R *por*

$$\frac{1}{\mathrm{R}} \coloneqq \limsup |a_n|^{\frac{1}{n}} \qquad \textit{em } [0,\infty].$$

(i) *Se* $|z-a| < \mathrm{R}$, *então a série* $\sum_n a_n(z-a)^n$ *converge absolutamente.*

(ii) *Se* $|z-a| > \mathrm{R}$, *então a série* $\sum_n a_n(z-a)^n$ *diverge.*

(iii) *Se* $r < \mathrm{R}$, *então a série de funções* $\sum_n u_n$ *com* $u_n := a_n(\cdot - a)^n$ *definidas sobre a bola* $\{z : |z-a| \leq r\}$ *converge uniformemente.*

Demonstração: Pelos traslados mutuamente inversos $z \mapsto z-a$ e $z \mapsto z+a$, podemos supor $a = 0$.

(i) Temos

$$\limsup |a_n|^{\frac{1}{n}} \leq \frac{1}{\mathrm{R}} \quad \text{em } [0,\infty].$$

se, e tão-somente se, existe N tal que $|a_n|^{\frac{1}{n}} \leq \frac{1}{\mathrm{R}}$ para todo $n \geq \mathrm{N}$. Equivalentemente, $|a_n| \leq \frac{1}{\mathrm{R}^n}$ para todo $n \geq \mathrm{N}$. Logo, para $|z| = r < \mathrm{R}$ e $n \geq \mathrm{N}$,

$$|a_n z^n| = |a_n||z|^n \leq \left(\frac{r}{\mathrm{R}}\right)^n.$$

Como $\rho := \frac{r}{\mathrm{R}} < 1$, a série $\sum_{n\geq \mathrm{N}} |a_n z^n|$ é dominada pela série geométrica convergente em ρ; logo

- a série $\sum_n a_n z^n$ converge absolutamente sobre a bola $\{z : |z| \leq r\}$, e
- a série de funções $\sum_n a_n \cdot^n$ sobre $\{z : |z| \leq r\}$ converge uniformemente pelo Teste de Weierstrass Teorema 12.3.

Se $|z| > \mathrm{R}$, então existe $r > \mathrm{R}$ com $|z| \geq r$. Temos

$$\limsup |a_n|^{\frac{1}{n}} \geq \frac{1}{\mathrm{R}} \quad \text{em } [0,\infty].$$

se, e tão-somente se, existem n_0, n_1, ... em $\mathbb{N}$ tal que $|a_{n_0}|^{\frac{1}{n_0}}$, $|a_{n_1}|^{\frac{1}{n_1}} \ldots \geq \frac{1}{\mathrm{R}}$. Logo a série $\sum_n a_n z^n$ tem uma infinitude de parcelas com índices $n = n_0$, n_1, ... com

$$|a_n z^n| \geq \rho^n \quad \text{com } \rho := \frac{r}{\mathrm{R}} > 1.$$

Como $\rho^n \to \infty$, estas parcelas são ilimitadas; logo $\sum_n a_n z^n$ não converge (porque se convergisse, então as parcelas convergiriam a 0).

□

Observação 12.8.

(i) Se $|\mathbf{K}|$ é denso em $[0,\infty[$, então o *raio de convergência* R é o único número que satisfaz as propriedades (i) e (ii) em Teorema 12.7.

(ii) Se $|\mathbf{K}|$ é discreto, então todos os valores R entre $\mathrm{R}' = \max\{r \in |\mathbf{K}| : r \leq \mathrm{R}\}$ e $\mathrm{R}'' = \min\{r \in |\mathbf{K}| : r \geq \mathrm{R}\}$ satisfazem as propriedades (i) e (ii) em Teorema 12.7.

Demonstração: Ad (i): Sejam R e S dois tais números. Após uma eventual troca dos seus nomes $\mathrm{R} \leq \mathrm{S}$. Por Teorema 12.7.(ii) para R, obtemos que R é o maior número positivo T tal que Teorema 12.7.(i) é satisfeito para todo z em $\mathbf{K}$ com $|z| < \mathrm{T}$. Logo, como Teorema 12.7.(i) é satisfeito para S, obtemos $\mathrm{S} \leq \mathrm{R}$.

Ad (ii): Claro.

Ad (iii): □

Em outras palavras: se $|\mathbf{K}|$ é denso em $[0,\infty[$, então por Teorema 12.7, para toda série de potências $\sum_n a_n(\mathrm{Z}-a)^n$ existe um único número R em $[0,\infty]$, determinado por

$$\frac{1}{\mathrm{R}} := \limsup |a_n|^{\frac{1}{n}},$$

tal que, para todo z em $\mathbf{K}$ com $|z| \neq \mathrm{R}$, a série $\sum_n a_n(z-a)^n$ converge se, e tão-somente se, $|z| < \mathrm{R}$.

Se $|\mathbf{K}|$ é discreto, então isto não vale mais. Contudo, se R em $[0,\infty]$ satisfaz, para toda extensão completa $\mathbf{L}$ de $\mathbf{K}$ e todo z em $\mathbf{L}$ com $|z| \neq \mathrm{R}$, que a série $\sum_n a_n(z-a)^n$ converge se, e tão-somente se, $|z| < \mathrm{R}$, então R é o raio de convergência.

Proposição 12.9. *Seja* $\sum_n a_n(\mathrm{Z}-a)^n$ *uma série de potências cujo* raio de convergência R *é definido por*

$$\frac{1}{\mathrm{R}} := \limsup |a_n|^{\frac{1}{n}} \quad \textit{em } [0,\infty].$$

Se a sequência

$$\left(\left|\frac{a_n}{a_{n+1}}\right| : n \in \mathbb{N}\right)$$

converge, então

$$\left|\frac{a_n}{a_{n+1}}\right| \to \mathrm{R}.$$

Demonstração: Pelos traslados mutuamente inversos $z \mapsto z - a$ e $z \mapsto z + a$, podemos supor $a = 0$.

Exista $\alpha = \lim|\frac{a_n}{a_{n+1}}|$. Se $r < \alpha$, então existe N tal que para todo $n \geq \mathrm{N}$ temos $r < |\frac{a_n}{a_{n+1}}|$; equivalentemente, $\frac{1}{r} > |\frac{a_{n+1}}{a_n}|$.

Se $|z| = s < r$, então estimemos, para $n \geq \mathrm{N}$, usando o *produto telescópico*

$$\left|\frac{a_n}{a_\mathrm{N}}\right| = \left|\frac{a_n}{a_{n-1}} \cdots \frac{a_{\mathrm{N}+1}}{a_\mathrm{N}}\right| = \frac{|a_n|}{|a_{n-1}|} \cdots \frac{|a_{\mathrm{N}+1}|}{|a_\mathrm{N}|}$$

que

$$|a_n z^n| = |a_n||z^n| = \frac{|a_n|}{|a_{n-1}|} \cdots \frac{|a_{\mathrm{N}+1}|}{|a_\mathrm{N}|}|a_\mathrm{N}||z|^n < |a_\mathrm{N}|\xi^n$$

com $\rho := \frac{s}{r} < 1$. Logo, como $s < r < \alpha$ foram arbitrários, $\sum_n a_n z^n$ é dominada pela série geométrica convergente em ρ; logo converge para todo z com $|z| < \alpha$; isto é, $\mathrm{R} \geq \alpha$.

Se $r > \alpha$, então existe N tal que para todo $n \geq \mathrm{N}$ temos $|\frac{a_n}{a_{n+1}}| < r$; equivalentemente, $\frac{1}{r} < |\frac{a_{n+1}}{a_n}|$. Se $|z| = s > r$, então estimemos, para $n \geq \mathrm{N}$, usando o *produto telescópico*

$$|a_n z^n| = \frac{|a_n|}{|a_{n-1}|} \cdots \frac{|a_{\mathrm{N}+1}|}{|a_\mathrm{N}|}|a_\mathrm{N}||z|^n > |a_\mathrm{N}|\rho^n$$

com $\rho := \frac{s}{r} > 1$. Logo a série $\sum_n a_n z^n$ tem uma infinitude de parcelas com

$$|a_n z^n| \geq \rho^n$$

Como $\rho^n \to \infty$, estas parcelas são ilimitadas; logo $\sum_n a_n z^n$ não converge (porque se convergisse, então as parcelas convergiriam a 0). Isto é, $\mathrm{R} \leq \alpha$. □

Observação. Cautela, nada se diz sobre a convergência ou divergência no círculo $|z| = \mathrm{R}$. Neste caso depende

- da série específica, e
- da topologia do corpo

se ela converge ou não! Por exemplo,

- se $a_n = 1$, então $\sum_n a_n \mathrm{Z}^n$ tem raio de convergência $\mathrm{R} = 1$ e diverge no círculo $|z| = 1$;
- se $a_n = \frac{(-1)^n}{n}$ e $\mathbf{K} = \mathbb{R}$, então $\sum_n a_n \mathrm{Z}^n$ tem raio de convergência $\mathrm{R} = 1$ e sobre o círculo $|z| = 1$

– diverge para $z = -1$, e

– converge para $z = 1$ a $\log(1+z)$.

Observação. Se $\mathbf{K}$ é não-arquimediano, então

- ou a série de potências converge em todo ponto do círculo $|z| = R$,
- ou a série de potências diverge em todo ponto do círculo $|z| = R$.

Demonstração: Exercício. □

O raio de convergência depende dos valores absolutos $|a_n|$ dos coeficientes da série de potências $\sum_n a_n(Z-a)^n$:

Observação 12.10. Seja o *corpo primitivo* de $\mathbf{K}$ o menor corpo de $\mathbf{K}$. Se a característica de $\mathbf{K}$ é 0, então é $\mathbb{Q}$, e, caso contrário, então é $\mathbb{F}_p$ para um número p primo.

Pelo Teorema de Ostrowski, Teorema 3.1, se a característica é 0, então sobre o seu corpo primitivo:

- Ou $|\cdot|$ é o valor trivial,
- ou $|\cdot|$ é equivalente a um valor absoluto $|\cdot|_p$ para p primo, isto é, existe $\alpha > 0$ tal que $|\cdot| = |\cdot|_p^\alpha$,
- ou $|\cdot|$ é equivalente ao valor absoluto usual $|\cdot|_\infty$, isto é, existe $\alpha > 0$ tal que $|\cdot| = |\cdot|_\infty^\alpha$.

Se a característica é $p > 0$, então, pela multiplicatividade de $|\cdot|$, por $\mathbb{F}_p$ ser finito, $|\cdot|$ é equivalente ao valor absoluto trivial.

Exemplo 12.11. Logo, se a característica de $\mathbf{K}$ é 0, então, por exemplo, $\frac{1}{n!} \to 0$ em $\mathbb{R}$, mas cresce monotonamente em $\mathbb{Q}_p$ para todo p primo:

Em $\mathbf{K} = \mathbb{R}$ ou $\mathbb{C}$, por Proposição 12.9, a série de potências

$$\exp Z := \sum_n \frac{Z^n}{n!}$$

(também denotada por e^Z) converge para todo z em $\mathbf{K}$.

Ao contrário, em um corpo $\mathbf{K}$ que contém (como subcorpo normado) $\mathbb{Q}_p$ o raio de convergência de $\exp Z$ é $R = p^{-\frac{1}{p-1}}$.

Sobre um corpo de característica positiva, $\exp Z$ não é definida.

12.4. Derivadas de Séries de Potências

Calculemos explicitamente as derivadas de uma série de potência. Não há surpresas: a derivada é dada pelas derivadas (formais) dos polinómios truncados.

Lema 12.12. *Temos* $\limsup \sqrt[n]{|n|} = 1$.

Demonstração: Escreve $\sqrt[n]{n} = 1 + \delta_n$ para $\delta_n > 0$. Temos, pelo desenvolvimento binomial,

$$n = (1+\delta_n)^n = \sum_{k=0,\dots,n} \binom{n}{k} \delta_n^k > \binom{n}{2} \delta_n^2$$

se, e somente se,

$$\delta_n^2 < \frac{2}{n} \to 0 \quad \text{para } n \to \infty. \tag{$*$}$$

Para $|\cdot|$ equivalente à $|\cdot|_p$, temos $\frac{1}{n} \leq |n| \leq 1$ para todo n em $\mathbb{N}$. Logo, por $(*)$,

$$1 = \frac{1}{\lim \sqrt[n]{n}} = \lim \frac{1}{\sqrt[n]{n}} \leq \lim \sqrt[n]{|n|} \leq 1.$$

Lema 12.13. *Sejam* (a_n) *e* (b_n) *sequências reais. Se* $a_n, b_n \geq 0$ *para todo* n *e* $\lim a_n = a$ *e* $b = \limsup b_n$ *existem e* $a > 0$*, então*

$$\limsup a_n b_n = ab.$$

Demonstração: A estimativa $\leq$ é clara.

Quanto à estimativa $\geq$, se $b = 0$, então $b = \lim b_n$, logo $\limsup a_n b_n = \lim a_n b_n = \lim a_n \lim b_n = ab$.

Se $b > 0$, sejam $\alpha < a$, $\beta < b$ positivos e N em $\mathbb{N}$. Temos de encontrar $n \geq \mathrm{N}$ tal que $a_n b_n > \alpha\beta$: Como $a_n \to a$, existe $\mathrm{N}' \geq \mathrm{N}$ tal que $a_n > \alpha$ para todo $n \geq \mathrm{N}'$. Como $\sup\{b_m : m \geq n\} \to b$, existe $\mathrm{N}'' \geq \mathrm{N}$ tal que $\sup\{b_m : m \geq n\} > \beta$ para todo $n \geq \mathrm{N}''$. Logo, $a_n b_n > \alpha\beta$ para todo $n \geq \mathrm{N}', \mathrm{N}''$. □

Corolário 12.14. *Seja* (a_n) *uma sequência real. Se* $a_n \geq 0$ *para todo* n *e* $a = \limsup a_n$ *existe, então*

$$\limsup a_n a_n^{\frac{1}{n}} = a.$$

Demonstração: Ou $a = 0$, quando $\limsup a_n = \lim a_n$ e $0 \leq a_n^{\frac{1}{n}} \leq 1$ para todo n; logo Lema 12.13 se aplica às sequências (a_n) e $(a_n^{\frac{1}{n}})$.

Ou $a > 0$, quando $a_n^{\frac{1}{n}} \to 1$; logo Lema 12.13 se aplica às sequências $(a_n^{\frac{1}{n}})$ e (a_n). □

Proposição 12.15. *Seja* $f(\mathrm{Z}) = \sum_n a_n(\mathrm{Z}-a)^n$ *uma série de potências. Se* $f(\mathrm{Z})$ *tem raio de convergência* $\mathrm{R} > 0$*, então*

(i) *para todo* $k \geq 1$*, a série de potências*

$$\sum_{n\geq k} n(n-1)\cdots(n-k+1)a_n(\mathrm{Z}-a)^{n-k}$$

tem raio de convergência R*,*

(ii) *a função* $z \mapsto f(z)$ *sobre* $\mathrm{B}(a,\mathrm{R})$ *é infinitamente derivável e a sua* k*-ésima derivada* $f^{(k)}$ *é dada por*

$$f^{(k)}(z) = \sum_{n\geq k} n(n-1)\cdots(n-k+1)a_n(z-a)^{n-k},$$

(iii) *para todo* $n \geq 0$*,*

$$n!a_n = f^{(n)}(a)$$

Demonstração: Ad (i): Mostremos primeiro a conclusão para $k = 1$: Calculemos o raio de convergência R da série de potências $\sum_{n\geq 1} na_n(\mathrm{Z}-a)^{n-1}$ por

$$\frac{1}{\mathrm{R}} = \limsup|na_{n+1}|^{\frac{1}{n}} = \limsup|a_{n+1}|^{\frac{1}{n}} = \limsup|a_n|^{\frac{1}{n-1}} = \limsup|a_n|^{\frac{1}{n}},$$

onde se aplicaram, na segunda igualdade,

- Lema 12.12 para $\lim\sqrt[n]{|n|} = 1$, e
- Lema 12.13 às sequências $(n^{\frac{1}{n}})$ e $(|a_{n+1}|^{\frac{1}{n}})$;

na última igualdade,

$$\limsup|a_n|^{\frac{1}{n-1}} = \limsup\left(|a_n|^{\frac{1}{n}}\right)^{\frac{n}{n-1}} = \limsup\left(|a_n|^{\frac{1}{n}}\right)^{1+\frac{1}{n-1}} = \limsup b_n b_n^{\frac{1}{n-1}}$$

com $b_n = |a_n|^{\frac{1}{n}}$. Por Corolário 12.14, aplicado à sequência (b_n),

$$\limsup b_n b_n^{\frac{1}{n-1}} = \limsup \mathrm{b}_n = \limsup|a_n|^{\frac{1}{n}}.$$

Isto é, o raio de convergência de $\sum_{n\geq 1} na_n(\mathrm{Z}-a)^{n-1}$ é igual ao de $f(\mathrm{Z})$.

Para $k \geq 1$, pela hipótese de indução,

$$\sum_{n\geq 0} a_n^{(k)}(\mathrm{Z}-a)^n = \sum_{n\geq k} n(n-1)\cdots(n-k+1)a_n(\mathrm{Z}-a)^{n-k},$$

onde $a_n^{(k)} := (n+k)(n+k-1)\cdots(n+1)a_{n+k}$, tem raio de convergência R. Logo, pelo caso $k = 1$,

$$\sum_{n\geq k+1} n(n-1)\cdots(n-k+1)(n-k)a_n(\mathrm{Z}-a)^{n-(k+1)},$$

tem raio de convergência R.

Ad (ii): Pelos traslados mutuamente inversos $z \mapsto z-a$ e $z \mapsto z+a$, podemos supor $a = 0$. Ponhamos $g(z) = \sum_n na_nz^{n-1}$. Como todas as séries são absolutamente convergentes, logo os seus limites independentes da ordem do somatório,

$$\begin{aligned}
&\frac{f(z)-f(w)}{z-w} - g(w)\\
=&\frac{\sum_n a_nz^n - \sum_n a_nw^n}{z-w} - \sum_{n\geq 1} na_nw^{n-1}\\
=&\left[\frac{s_\mathrm{N}(z)-s_\mathrm{N}(w)}{z-w} - s'_\mathrm{N}(w)\right] + \frac{\mathrm{R_N}(z)-\mathrm{R_N}(z)}{z-w} + \left[s'_\mathrm{N}(w) - g(w)\right]
\end{aligned}$$

com

$$s_\mathrm{N}(z) = a_0 + a_1z + \cdots + a_\mathrm{N}z^\mathrm{N} \quad \text{e} \quad \mathrm{R_N}(z) = a_{\mathrm{N}+1}z^{\mathrm{N}+1} + a_{\mathrm{N}+2}z^{\mathrm{N}+2} + \cdots .$$

Precisamos de mostrar que esta soma converge a 0 quando w converge a z:

Como o polinomial $s_\mathrm{N}(z)$ é diferenciável, a primeira diferença posta em colchetes converge a 0.

Como a série de potências $g(w)$ converge por (i), a última diferença posta em colchetes converge a 0.

Quanto à fração no meio, pela convergência absoluta,

$$\frac{\mathrm{R_N}(z)-\mathrm{R_N}(z)}{z-w} = a_{\mathrm{N}+1}\frac{z^{\mathrm{N}+1}-w^{\mathrm{N}+1}}{z-w} + a_{\mathrm{N}+2}\frac{z^{\mathrm{N}+2}-w^{\mathrm{N}+2}}{z-w} + \cdots .$$

Pela soma telescópica

$$(z-w)(z^n + z^{n-1}w + \cdots + zw^{n-1} + w^n) = (z^{n+1} - w^{n+1}),$$

obtemos

$$\frac{\mathrm{R_N}(z)-\mathrm{R_N}(z)}{z-w} = \sum_{n>\mathrm{N}} a_n(z^{n-1} + z^{n-2}w + \cdots + zw^{n-2} + w^{n-1})$$

Como $|z|$ e $|w| < r < \mathrm{R}$, obtemos

$$|a_n(z^{n-1} + z^{n-2}w + \cdots + zw^{n-2} + w^{n-1})| \leq |a_n|n|r|^{n-1};$$

logo, a série com as parcelas $a_n(z^{n-1} + z^{n-2}w + \cdots + zw^{n-2} + w^{n-1})$ converge absolutamente por (i) (sobre $\mathbf{K} = \mathbb{R}$ ou $\mathbb{C}$). Em particular,

$$\frac{\mathrm{R_N}(z) - \mathrm{R_N}(z)}{z - w} \to 0 \quad \text{para } \mathrm{N} \to \infty.$$

Por indução sobre k, aplicando esta igualdade à série com os coeficientes $a_n^{(k)} = (n+k)(n+k-1)\cdots n a_{n+k}$, obtemos que f é k vezes diferenciável com a sua k-ésima derivada

$$f^{(k)} = \sum_n n(n-1)\cdots(n-k+1)a_n(\mathrm{Z} - a)^{n-k}.$$

Ad (iii): Por avaliação da série de potências que descreve $f^{(k)}$ obtida em (ii) no ponto a. □

Observação. Se $\lim_n \frac{|a_n|}{|a_{n+1}|}$ exista, então Proposição 12.15.(i) segue diretamente de Proposição 12.9 pois

$$\frac{|na_n|}{(n+1)|a_{n+1}|} = \frac{|a_n|}{|a_{n+1}|}\frac{n}{n+1} \to \lim_n \frac{|a_n|}{|a_{n+1}|}.$$

Corolário 12.16. *Seja $f(\mathrm{Z}) = \sum_n a_n(\mathrm{Z} - a)^n$ uma série de potências. Se $f(\mathrm{Z})$ tem raio de convergência $\mathrm{R} > 0$, então a função $z \mapsto \sum_n a_n(z-a)^n$ é bem-definida e suave em $\mathrm{B}(a, \mathrm{R})$.*

Corolário 12.17. *Para todo z em $\mathbf{K}$ com $|z| < 1$*

$$\sum_n nz^n = \frac{1}{(1-z)^2}.$$

Demonstração: Por Proposição 12.15 aplicado à série geométrica e usando a igualdade $\sum_n z^n = \frac{1}{1-z}$ para $|z| < 1$. □

12.5. Coeficientes de uma Função Analítica

Mostremos que os coeficientes (da série de potências) de uma função analítica são determinados pelos seus valores.

Proposição 12.18. *Seja* X *em* $\mathbf{K}$ *aberto e* $f : \mathrm{X} \to \mathbf{K}$ *uma função. Se* f *é analítica e dada por* $f(x+h) = \sum_n a_n x^n$ *para* $a_0, a_1, \ldots$ *em* $\mathbf{K}$*, então* $n! a_n = f^{(n)}(h)$.

Em particular, se a caraterística de $\mathbf{K}$ é 0, então a_n é determinado pelos valores de f. Se a caraterística de $\mathbf{K}$ é positiva, então a mesma conclusão vale; porém, a demonstração é (um pouco) diferente.

12.6. Composição de Funções Analíticas

Vejamos como criar a partir de funções analíticas uma nova função analítica.

Proposição 12.19. *Sejam* $\sum_n a_n$ *e* $\sum_n b_n$ *duas séries. Se convergem absolutamente, então a série*

$$\sum_n c_n \qquad \text{com } c_n := a_0 b_n + \cdots + a_n b_0$$

converge absolutamente com limite $\sum_n a_n \sum_n b_n$.

Demonstração: Exercício. □

Corolário 12.20. *Sejam* $\sum_n a_n (\mathrm{Z} - a)^n$ *e* $\sum_n b_n (\mathrm{Z} - a)^n$ *duas séries de potências. Se têm raios de convergência* $\geq r > 0$*, então*

$$\sum_n c_n (\mathrm{Z} - a)^n \qquad \text{com } c_n := a_0 b_n + \cdots + a_n b_0$$

e

$$\sum_n d_n (\mathrm{Z} - a)^n \qquad \text{com } c_n = a_n + b_n$$

tem raio de convergência $\geq r$ *com*

$$\sum_n c_n (z-a)^n = \sum_n a_n (z-a)^n \sum_n b_n (z-a)^n$$

e

$$\sum_n d_n (z-a)^n = \sum_n a_n (z-a)^n + \sum_n b_n (z-a)^n$$

para todo $z \in \mathbf{K}$ *com* $|z - a| < r$.

Na prática, Corolário 12.20 permite-nos adicionar e multiplicar valores de séries de potências sobre o seu domínio de convergência, graças a sua convergência absoluta, como se fossem somas finitas.

13. Exponencial e Logaritmo

Introduzimos a exponencial e o logaritmo sobre $\mathbb{Q}_p$ e as suas extensões, definidos pelos mesmos séries de potências como sobre $\mathbb{R}$; contudo, vemos que devido ao valor p-ádico, o seu domínio de convergência é menor.

13.1. Exponencial

Seja $\mathbf{K}$ um corpo não-arquimediano que inclui o corpo normado $\mathbb{Q}_p$. Por exemplo, $\mathbb{Q}_p$ ou uma extensão finita dele.

Definição. A *exponencial* sobre $\mathbf{K}$ é definido por

$$\exp x = 1 + x + x^2/2! + x^3/3! + \cdots$$

Lema 13.1. *Temos*

$$v_p(n!) = \frac{n - s(n)}{p-1} \qquad com \qquad s(\sum a_i p^i) := \sum a_i.$$

Demonstração: Existem $\lfloor n/p \rfloor$ números divisíveis por p em $\{1, 2, \ldots, n\}$; Existem $\lfloor n/p^2 \rfloor$ números divisíveis por p^2 em $\{1, 2, \ldots, n\}$, e assim por diante. Se $n = a_0 + a_1 p + \cdots + a_s p^s$, então

$$v_p(n!) = \lfloor n/p \rfloor + \lfloor n/p^2 \rfloor + \cdots + \lfloor n/p^s \rfloor.$$

Escrevendo

$$\begin{aligned} \lfloor n/p \rfloor &= a_1 + a_2 p + \cdots \\ \lfloor n/p^2 \rfloor &= a_2 + a_3 p + \cdots \\ \vdots &= \vdots \\ \lfloor n/p^s \rfloor &= a_s \end{aligned}$$

e

$$\begin{aligned} n &= a_0 + p\lfloor n/p \rfloor \\ \lfloor n/p \rfloor &= a_1 + \lfloor n/p \rfloor \\ \vdots &= \vdots, \end{aligned}$$

obtemos, adicionando as linhas e comparando,

$$n + v_p(n!) = n + \lfloor n/p \rfloor + \lfloor n/p^2 \rfloor + \cdots + \lfloor n/p^s \rfloor. = s(n) + p v_p(n!),$$

isto é, $(p-1)v_p(n!) = n - s(n)$. □

Corolário. *O domínio de convergência de* exp *sobre* **K** *é dado por*

$$\mathrm{B}(0, < p^{1/(1-p)}) := \{x \text{ em } \mathbf{K} \text{ com } |x| < p^{1/(1-p)}\}.$$

Demonstração: Por Lema 13.1, temos

$$\lim \lambda(n)/n = \frac{1}{p-1},$$

isto é,

$$\lim \sqrt[n]{|n!|} = p^{1/(p-1)}.$$

Pelo Teorema 12.7, falta verificar a divergência de exp para x em **K** com $|x| = \rho$: Se n é uma potência de p, então $s(n) = 1$ e $|x^n/n!| = p^{1/(1-p)}$. Isto é, a subsequência $x^{p^n}/(p^n)!$ não converge a 0. □

Demonstração: O raio de convergência de log é $\rho = 1$. Como, por exemplo, $|1/p^n| = p^{-n}$, a série não converge sobre a borda $\{x \text{ em } \mathbf{K} \text{ com } |x| = 1\}$. □

Denote $\mathrm{E} := \mathrm{B}(0, < p^{1/(1-p)})$ este domínio de convergência de exp.

Teorema. *A exponencial satisfaz* $\exp(x+y) = \exp(x)\exp(y)$ *e* $\exp' = \exp$.

Demonstração: Usa que se $a_0 + a_1 + \cdots$ e $b_0 + b_1 + \cdots$ convergem, então $c_0 + c_1 + \cdots$ onde $c_n = a_0 b_n + \cdots + a_n b_0$ converge, e

$$(a_0 + a_1 + \cdots)(b_0 + b_1 + \cdots) = c_0 + c_1 + \cdots .$$

□

Teorema 13.2. *A exponencial é uma* isometria $\exp\colon \mathrm{E} \to 1 + \mathrm{E}$*, isto é*

$$|\exp \cdot| = |\cdot|.$$

Demonstração: Segue da definição de exp como série de potências. □

Em particular, a aplicação exp é injetora.

13.2. Logaritmo

Definamos o inverso da exponencial por duas vias:

O Logaritmo como Série de Potências. A definição mais comum é a como série de potências:

Definição. O *logaritmo* sobre **K** é definido por

$$\log 1 + x = x - x^2/2 + x^3/3 + \cdots$$

Corolário. *O domínio de convergência de* log *sobre* **K** *é dado por*

$$\mathrm{B}(1, < 1) = 1 + \mathrm{B}(0, < 1) := \{x \text{ em } \mathbf{K} \text{ com } |x - 1| < 1\}.$$

Denote $\mathrm{L} := \mathrm{B}(1, < 1)$ este domínio de convergência de log.

Observação. Temos $1 + \mathrm{E} \subset \mathrm{L}$.

Teorema. *O logaritmo é diferenciável com a derivada* $\log' x = 1/x$.

Demonstração: Por Proposição 12.15. □

Corolário. *O logaritmo é o inverso da exponencial, isto é,* $\exp \circ \log = \mathrm{id} = \log \circ \exp$.

Demonstração: Pela regra da cadeia,

$$(\exp \circ \log)'(x) = \exp(\log x) \cdot \frac{1}{x} = 1$$

e

$$(\log \circ \exp)'(x) = \frac{1}{\exp x} \cdot \exp x = 1.$$

Logo, $\exp \circ \log = x + \mathrm{C}$. Como $\exp \circ \log(1) = 1$, obtemos $\mathrm{C} = 0$. □

Corolário. *O logaritmo satisfaz* $\log(x) + \log(y) = \log(xy)$.

Demonstração: Como $\exp(x+y) = \exp(x)\exp(y)$ e $\exp \circ \log = \mathrm{id}$, temos $\log(x) + \log(y) = \log(xy)$ pela injetividade de exp. □

Corolário 13.3. *O logaritmo é uma* isometria $\log\colon 1 + \mathrm{E} \to \mathrm{E}$, *isto é*

$$|\log \cdot| = |\cdot|.$$

Demonstração: Como log é inverso a exp, e exp é uma isometria. □

Em particular, log é injetor sobre $1 + \mathrm{E}$ (mas, cautela, não sobre L).

Lema 13.4. *Para* x *em* $1 + \mathrm{B}(0, < 1)$, *vale* $x^{p^n} \to 1$ *para* $n \to \infty$.

Demonstração: Mostramos que se $|x-1| \leq \epsilon$, então $|x^p - 1| \leq \max\{|p|, \epsilon\}$.

A esta fim, Seja $x = 1 + a$ com $|a| \leq \epsilon$. Tem-se $x^p - 1 = \binom{p}{1}a + \binom{p}{2}a^2 + \cdots + \binom{p}{p}$. Como $|\binom{p}{1}|, |\binom{p}{2}|, \ldots \leq |p|$ e $|a|, |a^2|, \ldots \leq \epsilon$, segue a proposta desigualdade.

Por indução, segue $x^{p^n} \to 0$. □

Teorema 13.5. *Seja x em $1 + \mathrm{B}(0, <1)$. Vale $\log x = 0$ se, e somente se, existe n em $\mathbb{N}$ tal que $x^{p^n} = 1$.*

Demonstração: Seja x em $1+\mathrm{B}(0,<1)$. Pelo Lema 13.4, vale $x^{p^n} \to 1$ para $n \to \infty$. Em particular, para n suficientemente grande, x^{p^n} em $1 + \mathrm{E}$. Se $x^{p^n} = 1$, então $\log x^{p^n} = p^n \log x = 0$; isto é $\log x = 0$. Se $\log x^{p^n} = p^n \log x = 0$, então, pela injetividade de log sobre $1 + \mathrm{E}$, vale $x^{p^n} = 1$. □

O Logaritmo como Limite de p-Potências. Recordemo-nos da definição da exponencial sobre $\mathbb{R}$ pelos juros compostos

$$\exp(x) = \lim \left(1 + \frac{1}{n}\right)^n,$$

a qual leva à definição da função inversa

$$\log(x) = \lim n(x^{\frac{1}{n}} - 1) = \lim \frac{1}{\epsilon}(x^\epsilon - 1)$$

onde $\epsilon = \frac{1}{n} \to 0$.

Ora, em $\mathbb{Z}_p$ temos $p^n \to 0$. Logo, o bom análogo sobre U_1 é

$$\log(x) = \lim \frac{1}{p^{e-1}}(x^{p^{e-1}} - 1).$$

Com efeito, sobre U_1

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots$$

é um valor bem-definido em $p\mathbb{Z}/p^e\mathbb{Z}$, porque se p divide x, então nenhum denominador cortado é divisível por p e todos os números indivisíveis por p são invertíveis em $\mathbb{Z}/p^e\mathbb{Z}$. Da mesma maneira, sobre $p\mathbb{Z}/p^e\mathbb{Z}$,

$$\exp(x) = \sum_{n \geq 0} \frac{x^n}{n!}$$

é um valor bem definido em $1 + p\mathbb{Z}/p^e\mathbb{Z}$. Pois se p divide x, então nenhum denominador da fração cortada é divisível por p e todos os números indivisíveis por p são invertíveis em $\mathbb{Z}/p^e\mathbb{Z}$.

De interesse particular é a base e^p do logaritmo natural em $1 + p\mathbb{Z}/p^e\mathbb{Z}$, isto é, o argumento y tal que $\log y = 1$. Por exemplo, para $p = 7$ e $e = 4$, calculamos

$$\exp(p) = \sum_{n\geq 0} \frac{p^n}{n!} = 1 + p + \frac{p^2}{2} + \frac{p^3}{3!} = 1 + 7 \cdot 127 = 1 + 1 \cdot 7 + 4 \cdot 7^2 + 2 \cdot 7^3.$$

Logaritmo Modular. Na criptografia, usa-se o logaritmo modular para cifrar e assinar mensagens. Veremos porque o módulo usado é sempre um número primo pela redução de um módulo composto aos seus fatores primários:

Produto de Primos Diferentes. Se o módulo $m = pq$ é produto de dois fatores p e q sem fator comum, então o logaritmo modular

$$\log_g \bmod m$$

pode ser computado, pelo Teorema Chinês dos Restos, pelos logaritmos

$$\log_g \bmod p \quad \text{e} \quad \log_g \bmod q$$

Mais exatamente, existem inteiros a e b, computados (em tempo linear no número dos bits de p e q) pelo Algoritmo de Euclides (estendido), tais que que $ap + bq = 1$ e

$$\log_g \bmod m = a(\log_g \bmod p) + b(\log_g \bmod q).$$

Potência de um Primo. Se o módulo $m = p^e$ é uma potência de um primo p, então o logaritmo modular módulo m para uma base g

$$\log_g \colon \mathbb{Z}/m\mathbb{Z}^* \to \mathbb{Z}/\phi(m)\mathbb{Z}$$

pode ser computado em tempo polinomial pelo módulo p. Exponhamos os passos para um número primo $p > 2$:

1. Recordemo-nos de que $\mathbb{Z}/p^e\mathbb{Z}^*$ é cíclico de ordem $(p-1)p^{e-1}$. Logo, existe uma aplicação multiplicativa

$$\mathbb{Z}/p^e\mathbb{Z}^* \to \mu_{p-1} \times \mathrm{U}_1$$

dada por

$$x \mapsto x^{p^{e-1}}, x/x^{p^{e-1}} \quad (*)$$

onde

$$\mu_{p-1} = \{\zeta \in \mathbb{Z}/p^e\mathbb{Z}^* : \zeta^{p-1} = 1\}$$

denote o grupo das $(p-1)$-ésimas raízes da unidade e

$$U_1 = 1 + p\mathbb{Z}/p^e\mathbb{Z}$$

o das unidades unitárias.

2. Temos o isomorfismo

$$\mu_{p-1} \to \mathbb{F}_p^*$$

dado por $x \mapsto x \bmod p$ e o seu inverso por $X \mapsto X^{p^{e-1}}$ para qualquer X em $\mathbb{Z}/p^e\mathbb{Z}$ tal que $X \equiv x \bmod p$. (Observa que a restrição do homomorfismo

$$\mathbb{Z}/p^e\mathbb{Z}^* \to U_1$$

dado por $x^{1-p^{e-1}}$ a U_1 é a identidade porque a ordem de U_1 é p^{e-1}.)

3. Temos o logaritmo para a base g

$$\log_g \colon \mathbb{F}_p^* \to \mathbb{Z}/(p-1)\mathbb{Z}$$

e temos o logaritmo natural

$$\log \colon U_1 \to p(\mathbb{Z}/p^e\mathbb{Z})$$

que é calculado em tempo polinomial pela fórmula

$$eq : logaritmo - p - adicou \mapsto [x^{p^e} - 1]/p^e;$$

e o qual fornece o logaritmo $\log_g \colon U_1 \to p(\mathbb{Z}/p^e\mathbb{Z})$ para a base g pelo escalamento

$$\log_g = \log \cdot / \log g.$$

4. Pelo Teorema Chinês dos Restos, temos o isomorfismo

$$\mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/p^{e-1}\mathbb{Z} \to \mathbb{Z}/(p-1)p^{e-1}\mathbb{Z}$$

dada pelo produto e o seu inverso dado por $y \mapsto (ay \bmod p, by \bmod p^{e-1})$ onde a e b satisfazem $a(p-1)+b(p^{e-1}) = 1$ e foram obtidos pelo Algoritmo de Euclides (estendido).

Concluímos que, dado

- o número y em $\mathbb{Z}/p^e\mathbb{Z}$ e
- o seu valor $\log_g(y)$ sob $\log_g\colon \mathbb{F}_p^* \to \mathbb{Z}/(p-1)\mathbb{Z}$,

o valor $\log_g(y)$ de $\log_g\colon (\mathbb{Z}/p^e\mathbb{Z})^* \to \mathbb{Z}/(p-1)p^{e-1}\mathbb{Z}$ é computado em tempo polinomial.

Observação. Para facilitar a computação, ao invés da projeção

$$\mathbb{Z}/p^e\mathbb{Z}^* \to \mathrm{U}_1$$

dada em $(*)$ por $x \mapsto x^{1-p^{e-1}}$, é mais rápido usar a dada por $\pi\colon x \mapsto x^{p-1}$. Porém, a sua restrição a U_1 não é a identidade. Logo, é preciso usar ao invés de

$$\log g\colon \mathrm{U}_1 \to p\mathbb{Z}/p^e\mathbb{Z}$$

o logaritmo escalado

$$(p-1)^{-1}\log_g$$

para obter

$$\log_g = (p-1)^{-1}\log_g \circ\pi = (\log(g)p-1)^{-1}\log\circ\pi\colon \mathrm{U}_1 \to p\mathbb{Z}/p^e\mathbb{Z}.$$

13.3. Transcendência de e

Lembremo-nos de que por definição

$$\mathrm{e}^x = x + x^2/2! + x^3/3! + \cdots$$

Por Lema 13.1, vemos que

$$\mathrm{e}^p = p + p^2/2! + p^3/3! + \cdots$$

converge em $\mathbb{Q}_p$, isto é e^p existe em $\mathbb{Q}_p$.

Teorema 13.6 (Hensel). *Para todo número primo p, o polinômio*

$$\mathrm{X}^p - \mathrm{e}^p \quad \textit{em } \mathbb{Q}_p[\mathrm{X}]$$

é irredutível.

Corolário 13.7 ((errado!)). *O número* e *é transcendente.*

Demonstração: A falácia é que a série de e^p converge em $\mathbb{Q}_p$ a um número diferente da de e^p em $\mathbb{R}$. □

14. A base de Mahler

Seja **K** um corpo não-arquimediano que contenha $\mathbb{Q}_p$.

Definição. Seja V um espaço de Banach sobre **K**. Uma sequência (b_n) é

- uma *base ortogonal* se

$$\|\lambda_1 b_1 + \cdots + \lambda_n b_n\| = \max\{\|\lambda_1 b_1\|, \ldots, \|\lambda_n b_n\|\}$$

  para todo n em $\mathbb{N}$ e $\lambda_1, \ldots, \lambda_n$ em **K**, e

- uma *base ortonormal* se ela é uma base ortogonal e $\|b_1\| = \|b_2\| = \ldots = 1$, isto é,

$$\|\lambda_1 b_1 + \cdots + \lambda_n b_n\| = \max\{|\lambda_1|, \ldots, |\lambda_n|\}$$

  para todo n em $\mathbb{N}$ e $\lambda_1, \ldots, \lambda_n$ em **K**. Equivalentemente,

  – denote $c_0(\mathbb{N})$ o espaço de Banach das sequências com entradas em **K** que convergem a zero equipado com a norma $\|(a_n)\| := \sup\{|a_n| : n \in \mathbb{N}\}$, e

  – denote e_n para n em $\mathbb{N}$ a sequência cuja única entrada diferente de zero é 1 na posição n.

  A sequência (b_n) em V é uma base ortonormal, se

$$\begin{aligned} c_0(\mathbb{N}) &\xrightarrow{\sim} \mathscr{C}^0(\mathbb{Z}_p) \\ e_n &\mapsto \binom{x}{n}, \end{aligned}$$

  é um isomorfismo isométrico (isto é, uma bijeção linear que preserva a norma) entre espaços de Banach.

Observação. Um corpo não-arquimediano não tem uma ordem, como $\mathbb{R}$. Por isso, é inviável definir um produto escalar (isto é, uma aplicação bilinear simétrica $\langle , \rangle$ que satisfaz $\langle x, x\rangle > 0$) sobre um espaço vetorial sobre um corpo não-arquimediano. Logo, é inviável definir a ortogonalidade entre dois vetores por um produto escalar. Sobre espaços vetoriais normados sobre um corpo não-arquimediano, temos de contentar com a norma para definir ortogonalidade.

Veja [Scho3, Section 2.1]

Nesta seção, exibiremos uma base ortonormal *distinguida* das funções contínuas sobre $\mathbb{Z}_p$, os *polinômios de Mahler*, os coeficientes binomiais $\binom{x}{0}, \binom{x}{1}, \ldots$ como funções sobre $\mathbb{Z}_p$. Equivalentemente,

- denote $\mathscr{C}_0(\mathbb{Z}_p)$ o espaço de Banach das funções contínuas $f : \mathbb{Z}_p \to \mathbf{K}$, equipado com a norma $\|f\| := \sup\{|f(x)| : x \in \mathbb{Z}_p\}$, e
- denote $\binom{\cdot}{n}$ para n em $\mathbb{N}$ o coeficiente binomial $x \mapsto \binom{x}{n} = x(x-1)\cdots(x-n)/n!$ como função sobre $\mathbb{Z}_p$.

Teorema (Isomorfismo de Mahler, Teorema 14.2'). *A aplicação*

$$\begin{aligned} c_0(\mathbb{N}) &\xrightarrow{\sim} \mathscr{C}^0(\mathbb{Z}_p) \\ e_n &\mapsto \binom{x}{n}, \end{aligned}$$

é um isomorfismo isométrico (isto é, uma bijeção linear que preserva a norma) entre espaços de Banach.

O que distingue a *base de Mahler* de todas as outras bases ortonormais sobre as funções contínuas sobre $\mathbb{Z}_p$ (que existem em abundância), é que a avaliação de uma integral sobre ela dá um isomorfismo entre *álgebras* (e não meramente de espaços vetoriais). Em mais detalhes: Seja $\mathscr{D}^0(\mathbb{Z}_p)$ as integrais sobre $\mathbb{Z}_p$, isto é, o dual

$$\mathscr{D}^0(\mathbb{Z}_p) = \mathscr{C}^0(\mathbb{Z}_p)^* = \{\mu : \mathscr{C}^0(\mathbb{Z}_p) \to \mathbf{K} : \text{ linear e contínua }\},$$

dos funcionais sobre as funções contínuas sobre $\mathbb{Z}_p$ equipado com o produto $*$ de *convolução*

$$\mu * \nu := \mu[x \mapsto \nu f(x + \cdot)].$$

Teorema (Isomorfismo de Iwasawa). *A aplicação*

$$\begin{aligned} \mathscr{D}^0(\mathbb{Z}_p) &\to \{ \textit{ séries de potências } a_0 + a_1 X + \cdots \textit{ limitados } \} \\ \mu &\mapsto a_0 + a_1 X + \ldots \qquad \textit{com } a_0 = \mu(e_0), a_1 = \mu(e_1), \ldots \end{aligned} \tag{$*$}$$

é um isomorfismo de álgebras topológicas sobre $\mathbf{K}$

- *entre as integrais sobre* $\mathbb{Z}_p$ *(com o produto de* convolução*), e*
- *as séries de potências que convergem sobre a bola de unidade* fechada *em* $\mathbb{C}_p$ *(com o seu produto natural),*

Para convencermo-nos que o isomorfismo de Iwasawa é o dual do isomorfismo de Mahler, observamos

- que, como espaço vetorial normado, as séries de todas as potências limitadas são todas as sequências limitadas, e
- que por Proposição 15.8 o dual de $c^0(\mathbb{N})$, das sequências que convergem a zero, é $c^{\mathrm{b}}(\mathbb{N})$, as sequências limitadas,

$$c^0(\mathbb{N})^* = c^{\mathrm{b}}(\mathbb{N}).$$

Por substituição da variável, uma série de potências é uma função sobre todos os pontos em que converge. Observamos que os coeficientes de uma série de potências são limitados se, e somente se, converge sobre a bola de unidade fechada (por exemplo, em $\mathbb{C}_p$). O *isomorfismo de Amice* é o isomorfismo análogo ao isomorfismo de Iwasawa, obtido pela mesma avaliação $(*)$ da integral sobre a base de Mahler, mas onde

- o domínio consiste nas integrais sobre as funções *localmente analíticas* (ao invés das funções contínuas), isto é, em cada ponto a função é definida por uma série de potências convergente, e
- o codomínio nas séries de potências que convergem sobre a bola de unidade *aberta* em $\mathbb{C}_p$ (ao invés da bola fechada).

Com efeito, veremos que o isomorfismo de Iwasawa é o dual do isomorfismo de Mahler, e pela *Dualidade de Schikhof*, o primeiro isomorfismo implica o segundo.

14.1. Isomorfismo de Iwasawa

Continue $\mathbf{K}$ a ser um corpo completo não-arquimediano que contenha $\mathbb{Q}_p$. Seja $\mathbf{K}^\circ = \{x \in \mathbf{K} \mid |x| \leq 1\}$ o seu anel de inteiros e π um uniformizador. Denotem

$$\mathscr{C}^0(\mathbb{Z}_p)^\circ := \{ \text{ todas as funções contínuas } f\colon \mathbb{Z}_p \to \mathbf{K}^\circ\},$$

e

$$\mathscr{D}^0(\mathbb{Z}_p)^\circ := \{ \text{ todas as formas lineares contínuas } \mu\colon \mathscr{C}^0(\mathbb{Z}_p)^\circ \to \mathbf{K}^\circ\}.$$

(Para facilitar o argumento, restringimo-nos inicialmente aos valores em $\mathbf{K}^\circ$ ao invés de $\mathbf{K}$). Ambos os conjuntos são, com efeito,

- módulos sobre $\mathbf{K}^\circ$, e

- normados pelas normas de supremo

$$\|f\| := \sup\{|f(x)| : x \text{ em } \mathbb{Z}_p\} \quad \text{e} \quad \|\mu\| := \sup\{|\mu(f)| : f \text{ em } \mathscr{C}^0(\mathbb{Z}_p)^\circ\}.$$

Seja $\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]$ a álgebra do grupo $\mathbb{Z}/p^n\mathbb{Z}$ sobre $\mathbf{K}^\circ$, isto é

- o $\mathbf{K}^\circ$-módulo libre com base $\{g : g \text{ em } \mathbb{Z}/p^n\mathbb{Z}\}$,
- com a multiplicação $*$ definida sobre esta base por $g * h := g \cdot h$; isto é, o produto $*$ do anel entre os vetores da base é dado pelo produto $\cdot$ do grupo.

Como conjunto,

$$\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}] = \{f \colon \mathbb{Z}/p^n\mathbb{Z} \to \mathbf{K}^\circ\}$$

Identificamos pelo homomorfismo de quociente $\pi \colon \mathbb{Z}_p \twoheadrightarrow \mathbb{Z}_p/p^n\mathbb{Z}_p \xrightarrow{\sim} \mathbb{Z}/p^n\mathbb{Z}$ (o isomorfismo por Proposição 2.8),

$$\begin{aligned} \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}] &\xrightarrow{\sim} \mathscr{C}^{n-\mathrm{cst}}(\mathbb{Z}_p)^\circ \\ f &\mapsto f \circ \pi \end{aligned}$$

onde

$$\mathscr{C}^{n-\mathrm{cst}}(\mathbb{Z}_p)^\circ := \{f \colon \mathbb{Z}_p \to \mathbf{K}^\circ \text{ tais que } f(x) = f(y) \text{ se } x \equiv y \bmod p^n\mathbb{Z}_p\}.$$

Em particular,

$$\bigcup_{n\in\mathbb{N}} \mathbf{K}[\mathbb{Z}/p^n\mathbb{Z}] = \bigcup_{n\in\mathbb{N}} \mathscr{C}^{n-\mathrm{cst}}(\mathbb{Z}_p)^\circ = \mathscr{C}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ$$

onde

$$\mathscr{C}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ := \{ \text{ todas as funções } f \colon \mathbb{Z}_p \to \mathbf{K}^\circ \text{ localmente constantes } \}.$$

Proposição 14.1. *A inclusão*

$$\mathscr{C}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ \subseteq \mathscr{C}^0(\mathbb{Z}_p)^\circ$$

é densa.

Demonstração: Seja $\epsilon = |\pi^m| > 0$. Como $\mathbb{Z}_p$ é por Corolário 4.3 compacto, f é uniformemente contínua. Logo existe $\delta = |p^n| > 0$ tal que $x \equiv y \bmod p^n$ implica que $f(x) = f(y) \bmod \pi^m$ Fixa para cada a em $\mathbb{Z}/p^n\mathbb{Z}$ um $\hat{a}$ em $a + p^n\mathbb{Z}_p$; define $\mathrm{F}(x) := f(\hat{a})$ para x em $a + p^n\mathbb{Z}_p$. Então F é localmente constante e $|\mathrm{F}(x) - f(x)| \leq \epsilon$ para todo x em $\mathbb{Z}_p$. □

Corolário. *A restrição*

$$\begin{aligned}\mathscr{D}^0(\mathbb{Z}_p)^\circ &\xrightarrow{\sim} \mathscr{D}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ\\ \mu &\mapsto \mu_{\restriction \mathscr{C}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ}\end{aligned}$$

é uma bijeção.

Demonstração: Ela é

- injetora pela densidade e continuidade do funcional, e
- sobrejetora, ambos sendo espaços de Banach, pelo Teorema de Hahn-Banach Proposição 15.7. (Para poder aplicá-lo, observamos que $\mathscr{C}^0(\mathbb{Z}_p)$ é de tipo finito é justamente testemunhado por $\mathscr{C}^{\mathrm{lc}}(\mathbb{Z}_p)$.) □

Denote

$$\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]^* = \{ \text{ todas } \mu\colon \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}] \to \mathbf{K}^\circ \text{ lineares e contínuas } \}$$

Como $\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]$ é de dimensão finita, vale, por escolha de uma base,

$$\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}] \xrightarrow{\sim} \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]^*.$$

Obtemos

- que $\mathscr{D}^0(\mathbb{Z}_p)^\circ \xrightarrow{\sim} \mathscr{D}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ$,
- que $\mathscr{C}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ = \bigcup_{n\in\mathbb{N}} \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]$, e
- que $\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]^* = \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]$,

logo

$$\mathscr{D}^0(\mathbb{Z}_p)^\circ = \mathscr{D}^{\mathrm{lc}}(\mathbb{Z}_p)^\circ = \varprojlim \mathscr{D}^{n-\mathrm{cst}}(\mathbb{Z}_p)^\circ = \varprojlim \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]^* = \varprojlim \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}],$$

onde as aplicações de transição

$$\ldots \leftarrow \mathscr{D}^{n-\mathrm{cst}}(\mathbb{Z}_p)^\circ \leftarrow \mathscr{D}^{n+1-\mathrm{cst}}(\mathbb{Z}_p)^\circ \leftarrow \ldots$$

do limite inverso são dadas pela restrição de $\mathscr{C}^{n+1-\mathrm{cst}}(\mathbb{Z}_p)^\circ$ a $\mathscr{C}^{n-\mathrm{cst}}(\mathbb{Z}_p)^\circ$. Se denote

$$\mathbf{K}^\circ[[\mathbb{Z}_p]] := \varprojlim \mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}],$$

então concluímos

$$\mathscr{D}^0(\mathbb{Z}_p)^\circ \xrightarrow{\sim} \mathbf{K}^\circ[[\mathbb{Z}_p]]. \tag{*}$$

Ao produto de $\mathbf{K}^\circ[\mathbb{Z}/p^n\mathbb{Z}]$ corresponde o produto de *convolução* em $\mathscr{D}^0(\mathbb{Z}_p)^\circ$ dada por

$$\mu * \nu \colon f \mapsto \mu[x \mapsto \nu(f(\cdot + x))],$$

em particular, $\delta_1^n = \delta_1 * \cdots * \delta_1 = \delta_n$.

O grupo aditivo $\mathbb{Z}_p$ é topologicamente cíclico, isto é, gerado por um elemento, por exemplo, pelo elemento $\mathbf{1} = (\bar{1}, \bar{1}, \ldots)$ (visto que $\mathbb{N}$ é denso em $\mathbb{Z}_p$). Esta observação torna plausível:

Teorema (Isomorfismo de Iwasawa). *A aplicação dada por*

$$\begin{aligned} \mathbf{K}^\circ[[\mathbb{Z}_p]] &\xrightarrow{\sim} \mathbf{K}^\circ[[\mathrm{X}]] \\ \mathbf{1} &\mapsto 1 + \mathrm{X}. \end{aligned} \tag{**}$$

é um isomorfismo entre álgebras topológicas (sobre $\mathbf{K}^\circ$) onde

- *o lado esquerdo tem a topologia do limite inverso dos espaços vetoriais normados, e*
- *o lado direito tem a topologia da convergência ponto-por-ponto.*

Demonstração: Denote

$$\Gamma_n = \mathbb{Z}/p^n\mathbb{Z}$$

O grupo Γ_n é cíclico de ordem p^n. Seja $\gamma = \bar{1}$ o gerador. Tem-se

$$\mathbb{Z}_p[\Gamma_{n+1}] = \mathbb{Z}_p[\mathrm{T}]/\mathrm{T}^{p^{n+1}}$$

por $\gamma \mapsto \mathrm{T}$. Por $\mathrm{T} = 1 + \mathrm{X}$, obtemos

$$\mathbb{Z}_p[\mathrm{T}]/(\mathrm{T}^{p^n} - 1) \xrightarrow{\sim} \mathbb{Z}_p[\mathrm{X}]/((1 + \mathrm{X})^{p^n} - 1)$$

Seja

$$h_n(\mathrm{X}) := (1 + \mathrm{X})^{p^n} - 1.$$

Então

$$h_n(\mathrm{X}) = \mathrm{X}^{p^n} + \ldots$$

é um *polinômio distinguido*, isto é, onde todos os coeficientes exceto o primeiro são divisíveis por p.

Para um polinômio distinguido h, o Algoritmo de Euclides, Teorema A.1, mostra que $\mathbb{Z}_p[X]/(h)$ e $\mathbb{Z}_p[[X]]/(h)$ têm o mesmo posto $\deg h$ sobre $\mathbb{Z}_p$, e que

$$\mathbb{Z}_p[X]/(h) \to \mathbb{Z}_p[[X]]/(h).$$

é sobrejetor; logo, é um isomorfismo.

Obtemos então para cada n um homomorfismo

$$\mathbb{Z}_p[[X]] \to \mathbb{Z}_p[\Gamma_n] \xrightarrow{\sim} \mathbb{Z}_p[X]/(h_n),$$

então um homomorfismo

$$\epsilon : \mathbb{Z}_p[[X]] \to \varprojlim_n \mathbb{Z}_p[X]/(h_n).$$

Por indução, mostra-se que

$$h_n(X) = (1+X)^{p^n} - 1 \in (p, X)^{n+1}$$

onde (p, X) denota o ideal máximo de $\mathbb{Z}_p[X]$ gerado por p e X. Logo a interseção dos núcleos (h_n) é 0, isto é, o homomorfismo é injetor. Como o homomorfismo é evidentemente sobrejetor, ele é um isomorfismo. □

Quanto aos análogos de $\mathscr{C}^0(\mathbb{Z}_p)^\circ$ e $\mathscr{D}^0(\mathbb{Z}_p)^\circ$ para valores em $\mathbf{K}$ ao invés de $\mathbf{K}^\circ$,

$$\mathscr{C}^0(\mathbb{Z}_p) := \{ \text{ todas as funções contínuas } f : \mathbb{Z}_p \to \mathbf{K} \}$$

e

$$\mathscr{D}^0(\mathbb{Z}_p) := \{ \text{ todas as formas lineares contínuas } \mu : \mathscr{C}^0(\mathbb{Z}_p) \to \mathbf{K} \},$$

recordemo-nos de que por Lema 16.4 uma forma linear $\mu : \mathscr{C}^0(\mathbb{Z}_p) \to \mathbf{K}$ é contínua se, e somente se, ela é limitada. Isto é,

$$\mathscr{D}^0(\mathbb{Z}_p) = \mathbf{K} \otimes_{\mathbf{K}^\circ} \mathscr{D}^0(\mathbb{Z}_p)^\circ.$$

Juntando $(*)$ e $(**)$, concluímos:

Teorema 14.2. *O homomorfismo entre álgebras topológicas (sobre* $\mathbf{K}$*)*

$$\begin{aligned} \phi^* : \mathscr{D}^0_{\mathbf{K}}(\mathbb{Z}_p) &\xrightarrow{\sim} \mathbf{K} \otimes \mathbf{K}^\circ[[X]] \\ \delta_1 &\mapsto 1 + X \end{aligned} \tag{$\dagger$}$$

é um isomorfismo, onde

- *o lado esquerdo tem a topologia da convergência ponto-por-ponto, e*
- *o lado direito tem a topologia da da convergência ponto-por-ponto.*

Demonstração: Como $\delta_1 \mapsto \mathbf{1}$ sob o isomorfismo $\mathbf{K}^\circ[[\mathbb{Z}_p]] \xrightarrow{\sim} \mathscr{D}^0(\mathbb{Z}_p)^\circ$. □

14.2. Isomorfismo de Amice

Uma descrição semelhante pela base de Mahler vale também para o dual das funções *localmente analíticas*:

Definição 14.3. Uma função $f\colon \mathbb{Z}_p \to \mathbf{K}$ é *localmente analítica* se para todo x_0 em $\mathbb{Z}_p$ existe uma bola $\mathrm{B} = x_0 + p^n\mathbb{Z}_p$ em volta de x_0 tal que $f_{\restriction \mathrm{B}}(x_0 + \cdot)$ é uma função *analítica*, isto é, existem coeficientes $a_0, a_1, \ldots$ em $\mathbf{K}$ tal que

$$f_{\restriction \mathrm{B}}(x_0 + x) = a_0 + a_1 x + a_2 x^2 + \cdots$$

Como $\mathbb{Z}_p = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$ é totalmente desconexo e compacto, f é localmente analítica se, e somente se, $f = f_1 + \cdots + f_m$ para funções analíticas $f_1, \ldots, f_m$ sobre bolas disjuntas $\mathrm{B}_1 = x_1 + p^{n_1}\mathbb{Z}_p, \ldots, \mathrm{B}_n = x_n + p^{n_m}\mathbb{Z}_p$.

Denotem

$$\mathscr{C}^{\mathrm{la}}(\mathbb{Z}_p) = \{ \text{ todas as funções localmente analíticas } f\colon \mathbb{Z}_p \to \mathbf{K}\},$$

e

$$\mathscr{D}^{\mathrm{la}}(\mathbb{Z}_p) = \{ \text{ todas as aplicações lineares e contínuas } \mu\colon \mathscr{C}^{\mathrm{la}}(\mathbb{Z}_p) \to \mathbf{K}\}.$$

O *isomorfismo de Amice* é o análogo do *isomorfismo de Iwasawa*,

$$\mathscr{D}^0(\mathbb{Z}_p) \xrightarrow{\sim} \mathbf{K} \otimes \mathbf{K}^\circ[[\mathrm{X}]]$$

para $\mathscr{D}^{\mathrm{la}}(\mathbb{Z}_p)$ (ao invés de $\mathscr{D}^0(\mathbb{Z}_p)$):

Teorema 14.4 (Isomorfismo de Amice). *Seja*

$$\mathrm{B} = \{x \textit{ em } \mathbb{C}_p \textit{ com } |x| < 1\},$$

a bola de unidade aberta em $\mathbb{C}_p$*, e*

$$\mathscr{O}(\mathrm{B}) = \{\textit{todas as } \mathrm{F} \textit{ em } \mathbf{K}[[\mathrm{X}]] \textit{ tais que } \mathrm{F}(x) \textit{ converge para todo } x \textit{ em } \mathrm{B}\},$$

sejam as séries de potências que convergem sobre B*. O homomorfismo entre álgebras topológicas (sobre* $\mathbf{K}$*)*

$$\begin{aligned} \mathscr{D}^{\mathrm{la}}(\mathbb{Z}_p) &\xrightarrow{\sim} \mathscr{O}(\mathrm{B}) \\ \mu &\mapsto a_0 + a_1\mathrm{X} + a_2\mathrm{X}^2 + \cdots, \end{aligned}$$

onde os coeficientes $a_0 = \mu(\binom{\cdot}{0})$*,* $a_1 = \mu(\binom{\cdot}{1})$*,* $a_2 = \mu(\binom{\cdot}{2})$*, ... são dados pela avaliação na base de Mahler, é um isomorfismo.*

Demonstração: Veja [Sch99, Seção 2]. □

Por Proposição 14.1 as funções localmente constantes são densas nas funções contínuas; em particular, as funções localmente analíticas são densas nas funções contínuas. Logo, pela continuidade do funcional, a restrição de um funcional sobre $\mathscr{C}^0(\mathbb{Z}_p)$ a $\mathscr{C}^{\mathrm{la}}(\mathbb{Z}_p)$ é injetora,

$$\mathscr{D}^0(\mathbb{Z}_p) \hookrightarrow \mathscr{D}^{\mathrm{la}}(\mathbb{Z}_p).$$

Sob os isomorfismos de Iwasawa e Amice, vemos que esta inclusão corresponde a inclusão das séries de potências que convergem sobre a bola de unidade *fechada* nas séries de potências que convergem sobre a bola de unidade *aberta*.

14.3. Dualidade de Schikhof

Recordemos uma formulação da dualidade de Schikhof ([Sch95]) apta para nós (dada por [ST02, Theorem 1.2]):

Teorema 14.5 (Dualidade de Schikhof). *As categorias*

- *de espaços de* $\mathbf{K}$*-Banach* V *com aplicações lineares e contínuas (tal que* $\|\mathrm{V}\|$ *seja contido em* $|\mathbf{K}|$*), e*
- *de módulos topológicos* compactos *livres de torção (= produtos arbitrários de* $\mathcal{O}$ *consigo) com aplicações lineares e contínuas,*

são anti-equivalentes pelos funtores quase-inversos dados

- *por* $\mathrm{V} \mapsto \mathrm{V}^d = \{$ *todas as aplicações lineares e* uniformemente *contínuas* $f : \mathrm{V}^\circ \to \mathcal{O} \}$ *munido com a topologia da convergência* ponto-por-ponto *(onde* V° *denote a bola de unidade em* V*), e*
- *por* $\mathrm{M} \mapsto \mathrm{M}' = \{$ *todas as aplicações lineares e contínuas* ponto-por-ponto $f : \mathrm{M} \to \mathbf{K} \}$ *munido com a topologia da convergência* uniforme.

Demonstração: Veja [ST02, Theorem 1.2]. □

14.4. Teorema de Mahler

Denotemos

- por $c^0(\mathbb{N})$ as sequências com entradas em $\mathbf{K}$ que convergem a zero,

- por e_n para n em $\mathbb{N}$ a sequência cuja única entrada diferente de zero é 1 na posição n, e
- por $\binom{\cdot}{n}$ para n em $\mathbb{N}$ o coeficiente binomial $x \mapsto \binom{x}{n} = x(x-1)\cdots(x-n)/n!$ como função sobre $\mathbb{Z}_p$.

Teorema (Isomorfismo de Mahler, Teorema 14.2'). *O homomorfismo entre espaços de Banach (isto é, uma aplicação linear e contínua)*

$$\begin{aligned} \phi\colon\ c_0(\mathbb{N}) &\xrightarrow{\sim} \mathscr{C}^0(\mathbb{Z}_p) \\ e_n &\mapsto \binom{x}{n}, \end{aligned} \tag{‡}$$

é um isomorfismo (isto é, um homomorfismo bijetor).

Demonstração: Observamos

- que, como espaço vetorial normado, $\mathbf{K} \otimes \mathbf{K}^\circ[[\mathrm{X}]] = c^\flat(\mathbb{N})$, e
- que por Proposição 15.8 o dual de $c^0(\mathbb{N})$, das sequências que convergem a zero, é $c^\flat(\mathbb{N})$, as sequências limitadas,

$$c^\flat(\mathbb{N}). = c^0(\mathbb{N})^*$$

Tem-se $\delta_1^n = \delta_1 * \cdots * \delta_1 = \delta_n$. Logo, sob o isomorfismo (†),

$$\delta_n \mapsto (1+\mathrm{X})^n = \sum_{i=0,\dots,n} \binom{n}{i} \mathrm{X}^i$$

Como $\delta_n(\binom{\cdot}{i}) = \binom{n}{i}$, a aplicação (†) é o dual de (‡), isto é, ϕ^* é dado por

$$\begin{aligned} \mathscr{C}^0(\mathbb{Z}_p)^* &\to c^0(\mathbb{N})^* \\ \mu &\mapsto \mu \circ \phi. \end{aligned}$$

Como ϕ^* é um isomorfismo, ϕ é pela dualidade de Schikhof Teorema 14.5 também um isomorfismo. □

A imagem de $\mathscr{C}^r(\mathbb{Z}_p) \subseteq \mathscr{C}^0(\mathbb{Z}_p)$ das funções r-vezes diferenciáveis permite sob este isomorfismo a seguinte descrição concisa:

Teorema 14.6. *Temos o isomorfismo*

$$c_r(\mathbb{N}) = \{(a_n)\colon |a_n| n^r \to 0\} \xrightarrow{\sim} \mathscr{C}^r(\mathbb{Z}_p).$$

Isto é, uma função $f\colon \mathbb{Z}_p \to \mathbf{K}$ é r-vezes diferenciável se, e somente se, $f(x) = \sum a_n \binom{x}{n}$ com $|a_n| n^r \to 0$.

Demonstração: Veja [Nag12] □

15. Teorema de Hahn-Banach

Veremos que o Teorema de Hahn-Banach não-arquimediano vale se, e somente se, o corpo de coeficientes é esfericamente completo.

15.1. Corpos Esfericamente Completos

Recordemo-nos de que, pelo Critério de Cantor, um corpo $\mathbf{K}$ é *completo* se, e somente se, para cada sequência de bolas fechadas $B_1 \supseteq B_2 \supseteq \dots$ em $\mathbf{K}$ *cujos raios convergem a* 0, vale $\bigcap_n B_n \neq \emptyset$.

Definição. Um corpo $\mathbf{K}$ é *esfericamente completo* se, e somente se, para cada sequência de bolas fechadas $B_1 \supseteq B_2 \supseteq \dots$, vale $\bigcap_n B_n \neq \emptyset$.

Isto é, a condição é mais forte porque não se restringe mais a sequências de bolas cujos raios convergem a 0.

Observação. Se o corpo é não-arquimediano, então não importa se as bolas sejam fechadas ou abertas.

Proposição 15.1. *Todo corpo localmente compacto* $\mathbf{K}$ *é esfericamente completo.*

Demonstração: Por Proposição C.10, um conjunto K é localmente compacto se, e tão-somente se, para toda coleção $\mathcal{F}$ de subconjuntos fechados em K cujas interseções finitas são todas não-vazias, a sua interseção total $\bigcap \mathcal{F}$ é não-vazia. Se $\mathbf{K}$ é localmente compacto, então toda bola fechada $\bar{B}(x, \epsilon)$ é compacta. As interseções finitas da coleção $\{B_1, B_2, \dots\}$ são em particular todas não-vazias em $K = B_1$; logo, por K ser compacto, $\bigcap_n B_n \neq \emptyset$. □

Exemplo 15.2.

- Por Corolário 4.3, todos os corpos locais, tais como os corpos $\mathbb{Q}_p$ e $\mathbb{F}_p((t))$ e as suas extensões finitas, são localmente compactos. (Por Teorema 4.4, estes corpos são todos os corpos localmente compactos.)
- Pelo Teorema de Heine-Borel, todo subconjunto limitado e fechado em $\mathbb{R}$ (e em $\mathbb{C}$) é compacto; em particular, toda bola fechada é compacta. Logo $\mathbb{R}$ e $\mathbb{C}$ são localmente compactos.

Recordemo-nos que, por Fato 9.1, o fecho algébrico $\overline{\mathbb{Q}}_p$ de $\mathbb{Q}_p$ é incompleto e que $\mathbb{C}_p$ denota o seu completamento. Seção 9 alistou umas das suas propriedades principais: Em particular, o seu grupo de valores é denso, $|\mathbb{C}_p^*| = p^{-\mathbb{Q}}$.

Teorema 15.3. *Seja* **K** *um corpo não-arquimediano. Se* **K** *é*

- *separável, isto é, contenha um subconjunto enumerável denso, e*
- *a sua valoração é densa, isto é, cuja imagem é densa em* $\mathbb{R}$*,*

então **K** *não é esfericamente completo.*

Demonstração: Pela valoração densa, existem raios $r_1 > r_2 > \dots$ tal que $r := \lim r_n > 0$. Pela separabilidade, existe um subconjunto enumerável $\{s_1, s_2, \dots\}$ denso em **K**. Seja B_1 uma bola fechada de raio r_1 que não contenha s_1. Pela valoração densa, a bola B_1 de raio r_1 contem uma infinitude de bolas de raio r_2; seja B_2 uma tal bola "fechada" embutido em B_1 que não contenha s_2. A sequência $\mathrm{B}_1 \supseteq \mathrm{B}_2 \supseteq \dots$ assim construida tem intersecção vazia: Caso contrário, seria uma bola "fechada" de raio r, em particular, aberta, isto é, conteria um dos elementos s_1, s_2, ... em contradição às escolhas de B_1, B_2, ... □

Corolário 15.4. *O corpo* $\mathbb{C}_p$ *não é esfericamente completo.*

Demonstração: O corpo $\mathbb{C}_p$ é

- separável, isto é, contém o subconjunto enumerável denso dos números algébricos $\overline{\mathbb{Q}}$ sobre $\mathbb{Q}$, e
- a sua valoração é densa, isto é, a sua imagem é densa em $\mathbb{R}$, por exemplo, porque contém $\{p, \sqrt[2]{p}, \sqrt[3]{p}, \dots\}$.

Logo, por Teorema 15.3, é esfericamente incompleto. □

15.2. Hahn-Banach

Seja **K** um corpo não-arquimediano. Um *funcional* sobre um espaço vetorial normado V sobre um corpo **K** é uma aplicação linear e contínua $f : \mathrm{V} \to \mathbf{K}$.

Lema 15.5. *Uma aplicação linear entre espaços vetoriais normados é contínua se, e somente se, é* limitada*, isto é, tal que a sua imagem da bola unitária é contida em uma bola.*

Demonstração: Se é limitada, em particular é (uniformemente) contínua.

Seja $f : \mathrm{V} \to \mathrm{W}$ ilimitada, isto é, existe uma sequência x_1, x_2, ... (cujas entradas são) em V tal que $\|x_n\| \leq 1$ e $\|f(x_1)\|$, $\|f(x_2)\|$, ... é ilimitada. Sejam λ_1, λ_2, ... em **K** tal que

$$\{\|f(\lambda_1 x_1)\|, \|f(\lambda_2 x_2)\|, \dots\}$$

é limitada e o seu máximo > 1. Como $x_1, x_2, \ldots$ é limitada, $\lambda_1 x_1, \lambda_2 x_2, \ldots$ converge a 0. Como $\|f(\lambda_1 x_1)\|, \|f(\lambda_2 x_2)\|, \ldots > 1$, $f(x_1), f(x_2), \ldots$ não converge a $f(0) = 0$. Em particular, f não é contínua (em 0). □

A norma $\|f\|$ de um funcional f é definida por

$$\|f\| := \sup\{|f(v)| : v \text{ em V com } \|v\| \leq 1\}.$$

O Teorema de Hahn-Banach demonstra-se para um espaço vetorial sobre um corpo esfericamente completo como para um espaço vetorial sobre os números reais: Trocam-se os intervalos reais pelas bolas não-arquimedianos; ora, a sua interseção não é vazia porque o corpo é esfericamente completo.

Teorema 15.6 (Hahn-Banach não-arquimediano). *Seja* $\mathbf{K}$ *um corpo normado. Seja* X *um espaço vetorial normado sobre* $\mathbf{K}$ *e seja* M *um sob-espaço em* X. *Se* $\mathbf{K}$ *é esfericamente completo, então todo funcional* $f : \mathrm{M} \to \mathbf{K}$ *estende-se a um funcional* $\mathrm{F}: \mathrm{X} \to \mathbf{K}$ *com* $\|\mathrm{F}\| = \|f\|$.

Demonstração: Se $f = 0$, então $\mathrm{F} = 0$ estende f como requerido. Logo, podemos supor f não-nulo.

Se $|\mathbf{K}| = \|\mathrm{V}\|$, então substitui f por λf com $|\lambda| = \|f\|^{-1}$ para obter $\|f\| = 1$. Caso contrário, o argumento permanece o mesmo, mas um fator adicional aparece em várias equações. Permitamo-nos supor que $\|f\| = 1$.

Primeiro, consideremos o caso $\mathrm{X} = \mathrm{M} \oplus \mathbf{K} \cdot m_0$ para m_0 em X: Precisamos de determinar α em $\mathbf{K}$ tal que

$$\mathrm{F}(m + \lambda m_0) := f(m) + \lambda\alpha$$

tem norma 1, isto é,

$$|f(m) + \lambda\alpha| \leq \|m + \lambda m_0\|$$

para todo m em M e λ em $\mathbf{K}$. Em particular, para $\tilde{m} = -\lambda m$,

$$|f(\tilde{m}) + \lambda\alpha| = |\lambda||f(m) - \alpha| \quad \text{e} \quad \|\tilde{m} + \lambda m_0\| = |\lambda|\|m - m_0\|.$$

Por isto, basta demonstrar que exista α tal que

$$|f(m) - \alpha| \leq \|m - m_0\|$$

para todo m em M. Esta desigualdade vale se, e somente se, α em

$$\bigcap_{m \text{ em } \mathrm{M}} \mathrm{B}_{\leq \|m - m_0\|}(f(m)).$$

Como **K** é esfericamente completo, para esta interseção infinita não ser vazia, basta toda interseção sua finita não ser vazia; isto é, para m' e m'' em M,

$$\mathrm{B}_{\leq\|m'-m_0\|}(f(m')) \cap \mathrm{B}_{\leq\|m''-m_0\|}(f(m'')) \neq \emptyset.$$

Isto vale pois

$$|f(m') - f(m'')| = |f(m' - m'')| \leq \|m' - m''\| \leq \|m' - m_0\| + \|m'' - m_0\|.$$

Ora, consideremos o caso que X seja qualquer espaço vetorial normado que contém M. As extensões F de f com $\|\mathrm{F}\| = \|f\|$ são parcialmente ordenados por $\mathrm{F}' \leq \mathrm{F}''$ se $\mathrm{F}''_{\restriction \mathrm{N}'} = \mathrm{F}'$ e toda cadeia $(\mathrm{F}_i : i \in \mathrm{I})$ tem a cota superior dada pelo gráfico $\bigcup\{$ gráfico de $\mathrm{F}_i : i \in \mathrm{I}\}$. Pelo Lema de Zorn, Teorema D.2, existe uma extensão máxima $\mathrm{F}^* \colon \mathrm{M}^* \to \mathbf{K}$ com $\|\mathrm{F}\| = \|f\|$.

Se existisse m_0 em $\mathrm{X} - \mathrm{M}^*$, então existe uma extensão $\mathrm{F} \colon \mathrm{M}^* \oplus \mathbf{K}m_0 \to \mathbf{K}$ com $\|\mathrm{F}\| = |\mathrm{F}^*| = |f|$ pelo caso que acabamos de mostrar; em particular, F^* não seria máxima. Logo, $\mathrm{M}^* = \mathrm{X}$. □

Com efeito, o Teorema de Hahn-Banach vale para todos os espaços vetoriais sobre um corpo **K** se, e somente se, **K** é esfericamente completo! Contudo, para certos espaços vetoriais normados, o Teorema de Hahn-Banach vale independentemente do copo ser esfericamente completo ou não:

Proposição 15.7 ([PGS10, Theorem 4.2.4]). *Seja* X *um espaço vetorial sobre um corpo não-arquimediano* **K**. *Se* X *é de* tipo finito*, isto é, tem um subespaço de* dimensão numerável denso*, então a conclusão do Teorema de Hahn-Banach vale (sem a condição que* **K** *seja esfericamente completo).*

Por exemplo, $c^0(\mathbb{N})$ é de tipo finito.

15.3. Contra-Exemplo ao Teorema de Hahn-Banach

Construamos um contra-exemplo ao Teorema de Hahn-Banach se **K** esfericamente incompleto: Denotamos

- por $c^0(\mathbb{N})$ as sequências com entradas em **K** que convergem a zero, e
- por $c^{\mathrm{b}}(\mathbb{N})$ as sequências com entradas em **K** que são limitadas.

Ambos os conjuntos são espaços de Banach pela norma

$$\|(a_n)\| := \sup\{|a_n| : n \text{ em } \mathbb{N}\}.$$

Para um espaço vetorial normado V, seja

$$\mathrm{V}^* := \{ \text{ todas as aplicações } f : \mathrm{V} \to \mathbf{K} \text{ contínuas e lineares } \}$$

o seu dual contínuo dos funcionais sobre V, isto é, o espaço vetorial com a norma

$$\|f\| := \sup\{|f(v)| : v \text{ em } \mathrm{V}\}.$$

Proposição 15.8. *Temos*

$$c^0(\mathbb{N})^* = c^{\mathrm{b}}(\mathbb{N})$$

Demonstração: Por Lema 16.4 uma aplicação linear entre espaços vetoriais normados é contínua se, e somente se, é limitada. □

Dado que $\mathbf{K}$ não é esfericamente completo, explicitemos o espaço vetorial sem qualquer funcional diferente de zero: Denotemos

- por $\mathrm{V} = c^{\mathrm{b}}(\mathbb{N})$ as sequências com entradas em $\mathbf{K}$ que são limitadas,
- por $\mathrm{W} := c^0(\mathbb{N})$ as sequências com entradas em $\mathbf{K}$ que convergem a zero, e
- por $\mathrm{X} := \mathrm{V}/\mathrm{W}$ o seu quociente.

Os espaços vetoriais V e W são equipados com a sua norma de supremo natural, e o espaço vetorial $\mathrm{X} = \mathrm{V}/\mathrm{W}$ é equipado com a norma de quociente dada por

$$\|\bar{x}\| := \inf\{\|x\| \text{ para todos os } x \text{ em } \bar{x} = x + \mathrm{V}\}.$$

Com efeito, vale $\|\bar{x}\| = \limsup|x_n|$ se (x_n) em V é um representante de $\bar{x}$ em X.

Definição. Um espaço vetorial normado V é *esfericamente completo* se, e somente se, para cada sequência de bolas $\mathrm{B}_1 \supseteq \mathrm{B}_2 \supseteq \ldots$ em V vale $\bigcap_n \mathrm{B}_n \neq \emptyset$.

Proposição. *O espaço* X *é esfericamente completo.*

Demonstração: Seja $\mathrm{B}_1 = \mathrm{B}_{\leq r_1}(\bar{x}_1) \supseteq \mathrm{B}_{\leq r_2}(\bar{x}_2) \supseteq \ldots$ uma sequência de bolas em X com $r_1 > r_2 > \ldots$ Sejam $x_1, x_2, \ldots$ em V representantes de $\bar{x}_1, \bar{x}_2, \ldots$ em X indutivamente escolhidos tal que $\|x_n - x_{n+1}\| \leq r_{n-1}$ para todo $n > 1$. Seja $x := (x_{n,n})$ a sequência diagonal. Observamos que x em V, e que para qualquer $n > 1$,

$$\|\bar{x} - \bar{x}_n\| = \|\overline{x - x_n}\| = \limsup_i |x_{i,i} - x_{n,i}| \leq \limsup_i \|x_i - x_n\| \leq r_{n-1}$$

Por isso,

$$\bar{x} \in \bigcap \mathrm{B}_{\leq r_{n-1}}(\bar{x}_n) = \bigcap \mathrm{B}_{\leq r_n}(\bar{x}_{n+1}) = \bigcap \mathrm{B}_{\leq r_n}(\bar{x}_n).$$

Lema 15.9. *Se um espaço vetorial normado é esfericamente completo, então o seu quociente por todo subespaço fechado é esfericamente completo.*

Demonstração: Seja V esfericamente completo e W um subespaço fechado. Seja $X = V/W$. Seja $B_{\leq r_1}(\bar{x}_1) \supseteq B_{\leq r_2}(\bar{x}_2) \supseteq \dots$ uma sequência de bolas em X com $r_1 > r_2 > \dots$ Sejam $x_1, x_2, \dots$ em V representantes de $\bar{x}_1, \bar{x}_2, \dots$ em X indutivamente escolhidos tal que $\|x_n - x_{n+1}\| \leq r_{n-1}$ para todo $n > 1$. Como V é esfericamente completo, existe x em V tal que $\|x - x_n\| \leq r_{n-1}$ para todo $n > 1$. Em particular, $\|\bar{x} - \bar{x}_n\| \leq r_{n-1}$ para todo $n > 1$. Por isso,

$$\bar{x} \in \bigcap B_{\leq r_{n-1}}(\bar{x}_n) = \bigcap B_{\leq r_n}(\bar{x}_n).$$

Corolário. *Se* $\mathbf{K}$ *não é esfericamente completo, então o único funcional* $f : X \to \mathbf{K}$ *é* $f = 0$.

Demonstração: Como f é contínuo, $\ker f = f^{-1}\{0\}$ é fechado. Se $f \neq 0$, então $V/\ker f = \mathbf{K}$ seria esfericamente completo por Lema 15.9. □

Por exemplo, por Corolário 15.4, $X^* = 0$ sobre $\mathbf{K} = \mathbb{C}_p$.

16. Teorema de Alaoglu

Seja $\mathbf{K}$ um corpo normado. Um espaço vetorial é *topológico* se é um espaço vetorial com uma topologia tal que a adição + e a multiplicação escalar · são contínuas. Um *funcional* sobre um espaço vetorial topológico V sobre $\mathbf{K}$ é uma aplicação linear e contínua $f : \mathrm{V} \to \mathbf{K}$. Denote

$$\mathrm{V}^* := \{ \text{ todos os funcionais } f : \mathrm{V} \to \mathbf{K}\}.$$

Exemplo 16.1. Por exemplo, se $\mathbf{K}$ é não-arquimediano e V consiste das sequências de nulo,

$$\mathrm{V} = c_0(\mathbb{N}) := \{(a_n : n \in \mathbb{N}) : a_n \to n\}$$

então V^* consiste das sequências limitadas,

$$\mathrm{V}^* = c^b(\mathbb{N}) := \{(a_n : n \in \mathbb{N}) : \sup\{|a_n|\} < \infty\}.$$

16.1. Topologia Fraca e Fraca*

A topologia *fraca* de V é a topologia inicial dos funcionais sobre V; isto é, a menor topologia tal que todo funcional $f : \mathrm{V} \to \mathbf{K}$ seja contínua; isto é, a topologia gerada pelas pré-imagens $f^{-1}\mathrm{U}$ para os funcionais $f : \mathrm{V} \to \mathbf{K}$ e conjuntos abertos U em $\mathbf{K}$.

Exemplo 16.2. Por exemplo, se $\mathrm{V} = c_0(\mathbb{N})$, então (v_n) em V converge para a topologia fraca se, e tão-somente se, (v_n) é limitada e converge em cada coordenada.

A topologia *fraca** de V^* é a topologia inicial das avaliações $f \mapsto f(v)$ para todos os v em V; isto é, a menor topologia tal que toda avaliação $e_v : \mathrm{V}^* \to \mathbf{K}$ definida por $e_v : f \mapsto f(v)$ para v em V seja contínuas; isto é, a topologia gerada pelas pré-imagens $e_v^{-1}\mathrm{U}$ para todas as avaliações $e_v : \mathrm{V}^* \to \mathbf{K}$ e conjuntos abertos U em $\mathbf{K}$.

Se V é separável com subconjunto denso enumerável $\{v_n\}$, então a topologia fraca* é induzida pela métrica

$$\mathrm{d}(x, y) := \sum_{n \in \mathbb{N}} 2^{-n} \frac{|[x - y](v_n)|}{1 + |[x - y](v_n)|}.$$

Logo, formulado de forma mais palpável pela convergência das sequências (que definem os conjuntos fechados, logos os abertos como os seus complementos): Uma sequência (f_n) em V^* converge se $(f_n(x))$ converge para todo x em V.

Exemplo 16.3. Por exemplo, se $\mathrm{V} = c_0(\mathbb{N})$ e $\mathrm{V}^* = c^b(\mathbb{N})$, então (f_n) converge para a topologia fraca* se, e tão-somente se, (f_n) é limitada e converge em cada coordenada.

16.2. Topologia Forte (ou do Operador)

Lema 16.4. *Uma aplicação linear entre espaços vetoriais normados é contínua se, e somente se, é limitada.*

Demonstração: Se é limitada, em particular é (uniformemente) contínua.

Seja $f\colon \mathrm{V} \to \mathrm{W}$ ilimitada, isto é, existe uma sequência $x_1, x_2, \ldots$ (cujas entradas são) em V tal que $\|x_n\| \leq 1$ e $\|f(x_1)\|, \|f(x_2)\|, \ldots$ é ilimitada. Sejam $\lambda_1, \lambda_2, \ldots$ em $\mathbf{K}$ tal que $\|f(\lambda_1 x_1)\|, \|f(\lambda_2 x_2)\|, \ldots$ é limitada e maior do que 1. Como $x_1, x_2, \ldots$ é limitada, $\lambda_1 x_1, \lambda_2 x_2, \ldots$ converge a 0. Como $\|f(\lambda_1 x_1)\|, \|f(\lambda_2 x_2)\|, \ldots > 1$, $f(x_1), f(x_2), \ldots$ não converge a $f(0) = 0$. Em particular, f não é contínua (em 0). □

A *norma de operador* $\|f\|$ de um funcional f é definida por

$$\|f\| := \sup\{|f(v)| : v \text{ em } \mathrm{V}\};$$

que dá a V^* uma métrica d definida por $\mathrm{d}(f,g) = |f - g|$. que induz a topologia gerada pela base das bolas abertas $\mathrm{B}(x,r) := \{g \in \mathrm{V}^* : \mathrm{d}(g,f) < \epsilon\}$.

16.3. Reflexividade

Define o homomorfismo *canónico* $\mathrm{V} \to \mathrm{V}^{**}$ entre V o seu dual duplo $\mathrm{V}^{**} := (\mathrm{V}^*)^*$ por $v \mapsto [f \mapsto f(v)]$.

Definição 16.5. Um espaço vetorial topológico V é *reflexivo*, se a aplicação $\mathrm{V} \hookrightarrow \mathrm{V}^{**}$ é sobrejetora.

Exemplo 16.6.

- Se $\mathbf{K}$ é não-arquimediano e esfericamente completo, então um espaço de Banach V é reflexivo se, e tão-somente se, $\dim \mathrm{V} < \infty$.
- Se $\mathbf{K} = \mathbb{R}$ ou $\mathbb{C}$, então todo espaço vetorial de dimensão finita é reflexivo e todo espaço de Hilbert V é reflexivo, isto é, um espaço vetorial com um *produto escalar*, isto é, uma aplicação bilinear (ou, equivalentemente, linear em um dos dois argumentos) $\langle\cdot,\cdot\rangle\colon \mathrm{V} \times \mathrm{V} \to \mathbf{K}$ tal que

$$\overline{\langle x,y\rangle} = \langle y,x\rangle$$

onde $\bar{\cdot}$ é a conjugação complexa, e

$$\langle x, x\rangle \geq 0 \quad \text{e} \quad \langle x, x\rangle = 0 \text{ somente se } x = 0.$$

tal que ele é completo para a (métrica induzida pela) norma $\|\cdot\|$ definida por $\|x\| = \langle x, x\rangle$.

- Se $\mathbf{K}$ é uma extensão completa de $\mathbb{Q}_p$, então o espaço vetorial topológico das funções *localmente analíticas* (isto é, dada em uma vizinhança de cada ponto por uma série de potências) $\mathbb{Z}_p \to \mathbf{K}$ é reflexivo para a sua topologia natural (definida pelas normas das séries de potências locais, vide [Sch02, Exemplo após Proposition 16.10]).

Lema 16.7. *Seja* V *um espaço vetorial topológico sobre um corpo normado* $\mathbf{K}$. *Se* V *é normado e* $\mathbf{K}$ *é esfericamente completo, então o homomorfismo* $\mathrm{V} \to \mathrm{V}^{**}$ *é injetor.*

Demonstração: Se $v \neq 0$, então existe pelo Teorema de Hahn-Banach f em V^* com $f(v) = 1$. □

Corolário 16.8. *Seja* V *um espaço vetorial topológico sobre um corpo normado* $\mathbf{K}$. *Se* V *é normado e* $\mathbf{K}$ *é esfericamente completo, então* V *é reflexivo se, e tão-somente se,* $\mathrm{V} \xrightarrow{\sim} \mathrm{V}^{**}$ *como homomorfismo isométrico.*

Demonstração: Por Lema 16.7. □

Se $\mathbf{K}$ é completo, então V^* é completo como espaço normado (pela norma de operador). Logo, se V é normado e reflexivo, então $\mathrm{V} = \mathrm{V}^{**}$ é completo. Isto é, V é um espaço de Banach.

Corolário 16.9. *Seja* V *um espaço vetorial topológico. Se* V *é normado e reflexivo, então a topologia* fraca* *de* V^* *é igual à topologia* fraca *de* V^*.

Demonstração: Por Corolário 16.8, o homomorfismo $\mathrm{V} \to \mathrm{V}^{**}$ dado por $v \mapsto [f \mapsto f(v)]$ é bijetor. Logo, pela sua definição, as topologias fraca e fraca* são iguais. □

Fato 16.10. *Um espaço de Banach* V *sobre* $\mathbf{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *é reflexivo se, e tão-somente se, a sua bola unitária* $\{v \in \mathrm{V} \colon \|v\| \leq 1\}$ *é compacta para a topologia fraca.*

16.4. Teorema de Alaoglu

Definição 16.11. Um corpo $\mathbf{K}$ é *localmente compacto* se para todo x em $\mathbf{K}$ existe uma vizinhança $\mathrm{V} \ni x$ compacta.

Equivalentemente, um corpo $\mathbf{K}$ é *localmente compacto* se para todo x em $\mathbf{K}$ existe uma vizinhança aberta $\mathrm{V} \ni x$ tal que o seu fecho é compacto.

Se a bola unitária de $\mathbf{K}$ é compacta, então $\mathbf{K}$ é localmente compacto.

Por exemplo, $\mathbf{K} = \mathbb{R}$, $\mathbb{C}$ e as extensões finitas de $\mathbb{Q}_p$ e $\mathbb{F}_p((t))$ são localmente compactos (e de fato constituem todos tais corpos). Já vimos que $\mathbb{Z}_p$, a bola unitária de $\mathbb{Q}_p$, é compacto; logo $\mathbb{Q}_p$ é localmente compacto (e semelhantemente $\mathbb{F}_p((t))$). Para ver que $\mathbb{R}$ (e logo $\mathbb{C}$) é localmente compacto, basta por Teorema C.13 mostrar que todo subconjunto limitado é sequencialmente compacto, isto é, que toda sequência limitada tem uma subsequência convergente.

A bola *unitária* B de V^* é a bola fechada $\mathrm{B}(0,1) := \{f \in \mathrm{V}^* : \|f\| \leq 1\}$

Teorema 16.12 (Teorema de Banach-Alaoglu). *Seja $\mathbf{K}$ um corpo topológico e* V *um espaço normado sobre $\mathbf{K}$. Se a bola unitária em $\mathbf{K}$ é compacta, então a bola unitária fechada de* V^* *é compacta para a topologia fraca*.*

Demonstração: Seja B a bola unitária fechada de V e seja B^* a bola unitária fechada de V^*. Seja b a bola unitária fechada de $\mathbf{K}$. Logo

$$\mathrm{B}^* \hookrightarrow b^{\mathrm{B}}$$

por $f \mapsto f_{\restriction \mathrm{B}}$. Como, por definição,

- a topologia do produto é a menor topologia que torna todas as projeções contínuas,
- a topologia fraca* de V^* é a menor topologia que torna todas as avaliações $f \mapsto f(v)$ para v em V contínuas,

a inclusão é um mergulho, isto é, um homeomorfismo à sua imagem. Logo, B^* é compacta se, e tão-somente se, a sua imagem em b^{B} é compacta. Pelo Teorema de Tychonoff, o produto b^{B} dos compactos b é compacto para a topologia do produto. Logo, basta verificar que B^* é fechada em b^{B}. Dado $f : \mathrm{B} \to b$, vale f em B^* se, e tão-somente se, é restrição de uma aplicação linear, isto é,

- $f(v+w) = f(v) + f(w)$ para todo v e w em B tal que $v+w$ em B, e
- $f(\lambda w) = \lambda f(v)$ para todo λ em $\mathbf{K}$ e v em B tal que λv em V.

Isto é,

$$\mathrm{B}^* = \bigcap \{\mathrm{T}_{v,w}^{-1}\{0\} : v, w \in \mathrm{B} \text{ com } v + w \in \mathrm{B}\}$$
$$\cap \bigcap \{\mathrm{T}_{\lambda,w}^{-1}\{0\} : \lambda \in \mathbf{K}, w \in \mathrm{B} \text{ com } \lambda v \in \mathrm{B}\}$$

onde as avaliações lineares com domínio B e contra-domínio b são dadas por

$$\mathrm{T}_{v,w} \colon f \mapsto f(v+w) - (f(v) + f(w)) \quad \text{e} \quad \mathrm{T}_{\lambda,w} \colon f \mapsto f(\lambda v) - (\lambda f(v)).$$

Como todas as $\mathrm{T}_{v,w}$ e $\mathrm{T}_{\lambda,w}$ são contínuas, o conjunto unitário $\{0\}$ é fechado e a interseção de conjuntos fechados é fechada, concluímos que B^* é fechada. □

Lema 16.13. *Seja* V *um espaço vetorial topológico. Se* V *é separável, então existe uma métrica cujas bolas geram a topologia fraca* de* V^*.

Demonstração: Para um subconjunto $\{x_n\}$ denso em V, define a função distância de X^* por

$$\mathrm{d}(x,y) := \sum_{n\in\mathbb{N}} 2^{-n} \frac{|[x-y](x_n)|}{1+|[x-y](x_n)|}.$$

Corolário 16.14. *Seja* V *um espaço vetorial normado. Se* V *é separável, então a bola unitária fechada de* V^* *é sequencialmente compacta para a topologia fraca*, isto é, toda sequência* (f_n) *em* V^* *tem uma subsequência* (f_{n_k}) *tal que existe* f *em* V^* *com* $f_{n_k}(v) \to f(v)$ *para todo* v *em* V.

Demonstração: Primeiro Teorema 16.12, depois Lema 16.13 permite aplicar Teorema C.13. □

Corolário 16.15. *Seja* V *um espaço vetorial normado. Se* V *é reflexivo, então a bola unitária fechada de* V *é compacta para a topologia fraca. Se* V *é reflexivo e* V^* *é separável, então a bola unitária fechada de* V *é sequencialmente compacta para a topologia fraca, isto é, toda sequência* (v_n) *em* V *tem uma subsequência* (v_{n_k}) *tal que existe* v *em* V *com* $f(v_{n_k}) \to f(v)$ *para todo* f *em* V^*.

Demonstração: Se V é reflexivo, então a topologia fraca sobre V^* é igual à topologia fraca* sobre V^*. Logo, a topologia fraca* sobre V^{**} é igual à fraca sobre $\mathrm{V}^{**} = \mathrm{V}$. Logo, por Teorema 16.12 aplicado a $\mathrm{W} = \mathrm{V}^*$, a bola unitária fechada de $\mathrm{V} = \mathrm{W}^*$ é compacta para a topologia fraca.

Se V^* é um espaço normado e separável, então V é separável (sem demonstração). Logo, se V^* é separável e reflexivo, então V^{**} é separável. Então por Corolário 16.14, a bola unitária fechada de V^{**} é sequencialmente compacta para a topologia fraca*. Como V é reflexivo, a topologia fraca* sobre V^{**} é igual à fraca, e $\mathrm{V}^{**} = \mathrm{V}$. □

A. Divisão com Resto e o Algoritmo de Euclides

Sejam a e b números inteiros. O número b **divide** a ou, em fórmulas, $b \mid a$, se

$$\text{existe um número inteiro } c \text{ tal que } a = bc.$$

Por exemplo, um número inteiro, é *par* ou *ímpar* se é divisível por 2 ou não.

Definição. Um número $p > 1$ é **primo** se somente 1 e p dividem p.

$$\{ \text{números primos } \} = \{2,3,5,7,11,13,17,23,\ldots\}.$$

A.1. Divisão com Resto

Sejam a e b números inteiros positivos. Que a **dividido por** b tem **quociente** q e **resto** r significa

$$a = b \cdot q + r \quad \text{com } 0 \leq r < b.$$

Por exemplo, para $a = 230$ e $b = 17$, calculamos

$$230 = 17 \cdot 13 + 9.$$

Isto é, 230 *divido por* 17 tem *quociente* 13 e *resto* 9.

Para dois números inteiros a e b,

$$\textbf{divisor comum} = \textbf{número inteiro positivo que divide } a \textbf{ e } b.$$

O **maior** divisor comum de a e b é o **maior** número inteiro positivo que divide a e b. Denote

$$\text{mdc}(a,b) := \text{ o maior divisor comum de } a \text{ e } b \ .$$

Por exemplo,

$$\{ \text{divisores comuns de 24 e 26} \} = \{2,3,4,6,\mathbf{12}\}$$

e

$$\text{mdc} = \text{ o maior divisor comum } = 12.$$

A.2. Computar o Maior Divisor Comum pelo Algoritmo de Euclides

Sejam a e b números inteiros positivos tal que $a \geq b$ e com

$$a = b \cdot q + r \quad \text{com } 0 \leq r < b. \qquad (\dagger)$$

Equação $(\dagger) \implies d \mid a, b$ se, e somente se, $d \mid b, r, \implies$

$$\{\text{divisores comuns de } a \text{ e } b\} = \{\text{divisores comuns de } b \text{ e } r\},$$

em particular

$$\text{mdc}(a, b) = \text{mdc}(b, r).$$

Reaplicando a divisão com resto aos números menores b e r,

$$b = r \cdot q' + r' \quad \text{com } 0 \leq r' < r$$

e por isto

$$\text{mdc}(b, r) = \text{mdc}(r, r').$$

Iterando, chegamos a $s := r'^{\cdots'}$ e $r'^{\cdots''}$ com $r'^{\cdots''} = 0$, isto é

$$\text{mdc}(a, b) = \ldots = \text{mdc}(s, 0) = \mathbf{s}.$$

Por exemplo, calculamos o maior divisor comum de $a = 748$ e $b = 528$ pela iterada divisão com resto:

$$\begin{aligned} 748 &= 528 \cdot 1 + 220 \\ 528 &= 220 \cdot 2 + 88 \\ 220 &= 88 \cdot 2 + \mathbf{44} \\ 88 &= 44 \cdot 2 + 0, \end{aligned}$$

logo $\text{mdc}(528, 220) = 44$. Isto é,

$$\text{o maior divisor comum} \;=\; \text{o penúltimo resto}\,.$$

Teorema (Algoritmo de Euclides). Sejam a e b números inteiros positivos com $a \geq b$. O seguinte algoritmo calcula $\text{mdc}(a, d)$ em um número finito de passos:

*(*inicialização*) Põe* $r_0 = a$ *e* $r_1 = b$*, e* $i = 1$.
*(*divisão com resto*) Divide* r_{i-1} *por* r_i *com resto para obter*

$$r_{i-1} = r_i q_i + r_{i+1} \quad \textit{com } 0 \leq r_{i+1} \leq r_i.$$

*(*iteração*) Distingue entre:*

- *ou* $r_{i+1} = 0$*, então* $r_i = \text{mdc}(a, b)$ *e o algoritmo termina,*
- *ou* $r_{i+1} > 0$*, então ponha* $i := i + 1$ *e continue no passo (*divisão com resto*).*

A.3. Algoritmo de Euclides Estendido = Computar o Maior Divisor Comum como Combinação Linear

Para números inteiros $v_1, \ldots, v_d$, uma **combinação linear** de $v_1, \ldots, v_d$ é uma soma s de múltiplos inteiros deles,

$$s = \lambda_1 v_1 + \cdots + \lambda_d v_d \quad \text{com inteiros } \lambda_1, \ldots, \lambda_d.$$

O *Algoritmo de Euclides Estendido* mostra iterativamente que

$$\textbf{maior divisor comum} \text{ de } a \text{ e } b \;=\; \textbf{combinação linear} \text{ de } a \text{ e } b\ .$$

Teorema A.1 (O algoritmo de Euclides Estendido). *Para quaisquer números inteiros positivos a e b, o seu maior divisor comum* $\mathrm{mdc}(a,b)$ *é uma combinação linear de a e b; isto é, existem números inteiros u e v tais que*

$$\mathrm{mdc}(a,b) = au + bv.$$

Demonstração: Inicialmente, com $r_0 := a$, $r_1 := b$,

$$r_0 = r_1 q_i + r_2 \quad \text{com } 0 \le r_2 < r_1.$$

Isto é, $r_2 = r_0 - r_1 q_1$, $\Longrightarrow$

$$r_0, r_1, \text{ e } r_2 = \text{ combinações lineares de } a \text{ e } b\ .$$

Por indução,

$$r_{i-1} = r_i q_i + r_{i+1} \quad \text{com } 0 \le r_{i+1} \le r_i.$$

Como r_{i-1} e r_i são combinações lineares de a e b,

$$r_{i+1} = r_{i-1} - r_i q_i = \text{ uma combinação linear de } a \text{ e } b\ .$$

Em particular, quando finalmente $r_{I+1} = 0$,

$$r_I = \mathrm{mdc}(r_I, r_{I+1}) = \mathrm{mdc}(a,b) = \text{ combinação linear de } a \text{ e } b\ .$$

Já calculamos o maior divisor comum de $a = 748$ e $b = 528$,

$$\begin{aligned} 748 &= 528{\cdot}1{+}220 \\ 528 &= 220{\cdot}2{+}88 \\ 220 &= 88\ {\cdot}2{+}\mathbf{44} \\ 88 &= 44\ {\cdot}2{+}0, \end{aligned}$$

Logo,

$$\begin{aligned} 220 &= 748 - 528 \cdot 1 = a - b \\ 88 &= 528 - 220 \cdot 2 = b - (a-b)2 = 3b - 2a \\ \mathbf{44} &= 220 - 88 \cdot 2 = (a-b) - (3b-2a)2 = 5a - 7b, \end{aligned}$$

e, com efeito,

$$44 = 5 \cdot 748 - 7 \cdot 528.$$

B. Aritmética Modular

Denotem

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} \quad \text{e} \quad \mathbb{R}$$

os *anéis* dos números inteiros e dos números reais (= a reta).

Um **anel** (comutativo com 1) é

- um conjunto que possui
  - **0** (= o elemento neutro da adição), e
  - **1** (= o elemento neutro da multiplicação);
- sobre o qual operam
  - a **adição** + (e o seu *inverso* −), e
  - a **multiplicação** ·,

  que satisfazem a lei *associativa*, *comutativa* e *distributiva*.

B.1. Aritmética Modular no dia-a-dia

Damos exemplos da aritmética modular do nosso dia-a-dia (como o relógio, os dias da semana, ou até o alfabeto). A propriedade comum entre estes exemplos é a sua circularidade (de onde a nomenclatura "anel").

Relógio. O exemplo protótipo de aritmética modular é a aritmética do relógio em que o ponteiro volta após 12 horas no início; isto é, vale a equação

$$12 = 0,$$

que implica, entre outros, as equações

$$9 + 4 = 1 \quad \text{e} \quad 1 - 2 = 11; \quad (*)$$

Quer dizer, 4 horas depois das 9 horas é 1 hora, e 2 horas antes da 1 hora são 11 horas. Podemos ir mais longe:

$$9 + 24 = 9, \quad (**)$$

quer dizer se agora são 9 horas, então 24 horas (= um dia) mais tarde também.

Figura B.1: Relógio como Anel dos números $1, 2, \ldots, 11, 12 = 0$

Dias da Semana. Além das horas, outro exemplo do aritmética modular no dia-a-dia são os dias da semana: tendo passado 7 dias, os dias da semana recomeçam. Se numeramos `sábado`, `domingo`, `segunda`, `terça`, `quarta`, `quinta` e `sexta-feira` por 0, 1, 2, 3, 4, 5, 6 então vale a equação

$$7 = 0,$$

e que implica, entre outros, as equações

$$4 + 4 = 1 \quad \text{e} \quad 1 - 2 = 5;$$

Quer dizer, 4 dias depois da quarta-feira é domingo, e 2 dias antes do domingo é sexta-feira. Podemos ir mais longe: $5 + 14 = 5$, quer dizer se agora é `quinta-feira`, então daqui em 14 dias (= duas semanas) também.

A cifração de César. Na cifração de César, trasladamos cada letra do alfabeto por uma distancia t fixa; por exemplo, para $t = 3$, obtemos

$$\mathrm{A} \mapsto \mathrm{D}, \mathrm{B} \mapsto \mathrm{E}, \ldots, \mathrm{W} \mapsto \mathrm{Z}.$$

Para trasladarmos as últimas $t = 3$ letras X, Y e Z do alfabeto, consideramos o alfabeto como anel, isto é:

Figura B.2: A circularidade semanal de tomar comprimidos

Assim,

$$X \mapsto A, \ldots, Z \mapsto C.$$

Se identificamos cada letra do alfabeto romano com a sua posição,

$$A \mapsto 0, B \mapsto 1, \ldots, X \mapsto 23, Y \mapsto 24, Z \mapsto 25.$$

então vale $23 + 3 \mapsto 0$, $24 + 3 = 1$ e $25 + 3 = 2$; isto é, $26 = 0$.

Formalização. Formalmente, derivamos as equações em $(*)$ e $(**)$ das igualdades

$$9 + 4 = 13 = 12 + 1 = 0 + 1 = 1 \quad \text{e} \quad 1 - 2 = -1 = -1 + 0 = -1 + 12 = 11.$$

e

$$9 + 24 = 9 + 2 \cdot 12 = 9 + 2 \cdot 0 = 9.$$

Em geral, para quaisquer a e x em $\mathbb{Z}$,

$$a + 12 \cdot x = a$$

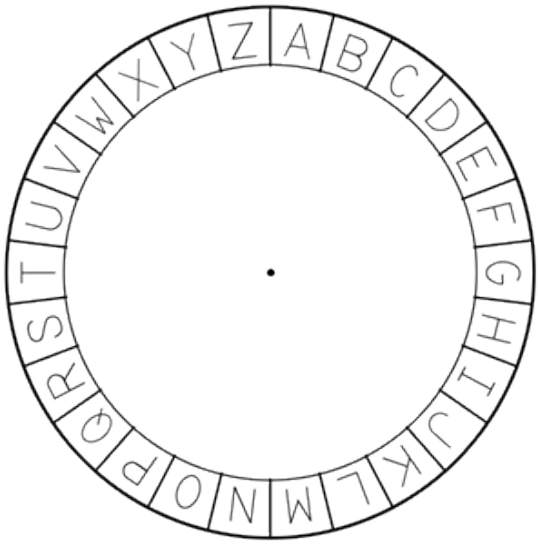


Figura B.3: Roda das Letras do Alfabeto Latin

ou, equivalentemente, para quaisquer a e b em $\mathbb{Z}$,

$$a = b \quad \text{se } 12 \mid a - b.$$

Em palavras, a e b deixam o mesmo resto divido por 12.

As mesmas observações valem para os dias da semana.

Não há nada de especial sobre o número $m = 12$ (horas até meio-dia), $m = 7$ (dias da semana) ou $m = 26$ (letras do alfabeto latim). Por exemplo, as mesmas observações valeriam se o relógio indicasse $m = 15$ horas (como no planeta Netuno em que o dia, a rotação completa em torno do próprio eixo, dura 16 horas):

Definição. Seja $m \geq 1$ um inteiro. Os números inteiros a e b são **congruentes módulo** m ou, em fórmulas,

$$a \equiv b \bmod m.$$

se $m \mid a - b$, isto é se a sua diferença $a - b$ é divisível por m. Em outras palavras, se a e b deixam o mesmo resto divido por m. O número m é o **módulo**.

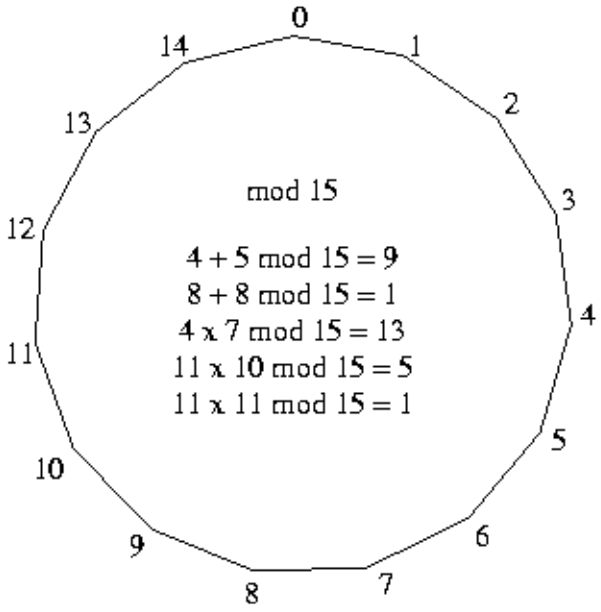


Figura B.4: O relógio com 15 horas

B.2. O Anel Quociente

Dado um número inteiro m, construiremos de duas vias, primeiro pela teórica e depois pela prática,

- o **menor** anel (= conjunto que possui 0 e 1 e sobre o qual + e · operam), denotado por $\mathbb{Z}/m\mathbb{Z}$,
- com **uma aplicação** $\bar{\cdot}\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ denotada por

  $$x \mapsto \bar{x}$$

  que **satisfaz**

  $$\overline{x+y} = \bar{x} + \bar{y}$$

  (e portanto $\overline{x \cdot y} = \bar{x} \cdot \bar{y}$, em particular, $\bar{0} = 0$ e $\bar{1} = 1$), e
- tal que

  $$x \mapsto 0 \iff m \mid x \qquad (\dagger)$$

  ou, equivalentemente,

  $$x \equiv y \bmod m \iff \bar{x} = \bar{y} \text{ em } \mathbb{Z}/m\mathbb{Z}.$$

Explicamos o que significa o *menor* anel que satisfaz as propriedades precedentes. Com efeito, menor não tem significado matemático. Em verdade, falamos matematicamente de uma propriedade *universal:* Se A é outro anel com uma aplicação $\pi\colon \mathbb{Z} \to A$ que satisfaz (†), isto é, tal que $\pi(x) = 0$ se, e somente se, $m \mid x$, então existe uma (com efeito, única) aplicação $\phi\colon \mathbb{Z}/m\mathbb{Z} \to A$ tal que $\pi = \phi \circ \bar{\cdot}$. (Diz-se que a aplicação π *fatora* através de $\bar{\cdot}$.)

Construção *Teórica.* Esta construção do anel residual é a ordinariamente dada, como conjunto de classes residuais: Pela Equação (†) exatamente o conjunto $m\mathbb{Z} \mapsto 0$, e consequentemente para x em $\mathbb{Z}$ qualquer, exatamente os seus traslados

$$x + m\mathbb{Z} \mapsto \bar{x}.$$

Isto nos leva a definir

- como **conjunto**

  $$\mathbb{Z}/m\mathbb{Z} := \{x + m\mathbb{Z} : x \text{ em } \mathbb{Z}\}$$

  e como **aplicação** $\bar{\cdot}\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$

  $$x \mapsto x + m\mathbb{Z};$$

- como **elementos neutros** da adição e multiplicação

  $$0 = 0 + m\mathbb{Z} = m\mathbb{Z} \quad \text{e} \quad 1 = 1 + m\mathbb{Z};$$

- como **operações** + e ·

  $$\bar{x} + \bar{y} := \overline{x + y} \quad \text{e} \quad \bar{x} \cdot \bar{y} := \overline{x \cdot y}.$$

Os elementos (ou números) em $\mathbb{Z}/m\mathbb{Z}$ são subconjunto de $\mathbb{Z}$, com efeito *classes residuais.*

Construção *Prática.* Esta definição reflete melhor o nosso ponto de vista intuitivo da aritmética modular, por exemplo, sobre o relógio, isto é, $\mathbb{Z}/12\mathbb{Z}$: Pela divisão com resto,

$$\{x + m\mathbb{Z} : x \text{ em } \mathbb{Z}\} = \{0 + m\mathbb{Z}, \ldots, m - 1 + m\mathbb{Z}\},$$

isto é, $0, \ldots, m - 1$ representam o conjunto $\mathbb{Z}/m\mathbb{Z}$. Por isso definamos

- como **conjunto**

$$\mathbb{Z}/m\mathbb{Z} := \{0, \ldots, m-1\};$$

  e como **aplicação** $\bar{\cdot}\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$

$$x \mapsto r \quad \text{onde } x = qm + r \quad \text{com } r \text{ em } \{0, \ldots, m-1\};$$

- como **elementos neutros** da adição e multiplicação 0 e 1,
- como **operações** + e ·

$$x + y := r \quad \text{onde } x + y = qm + r \quad \text{com } r \text{ em } \{0, \ldots, m-1\}$$

  e

$$x \cdot y = r \quad \text{onde } x \cdot y = qm + r \quad \text{com } r \text{ em } \{0, \ldots, m-1\}.$$

Por exemplo, as tabelas da adição e multiplicação para $m = 4$ são:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

e

*	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Exercício B.1. Mostra que um número inteiro é divisível por 3 (respectivamente 9) se, e tão-somente se, a soma dos seus algarismos decimais é divisível por 3 (respectivamente 9).

B.3. Números Invertíveis

Em $\mathbb{Z}$, os únicos números que dividem todos são ±1. Em $\mathbb{Z}/m\mathbb{Z}$, depende do módulo m se todos os seus elementos podem ser divididos

- só pelas imagens dos números invertíveis ±1 em $\mathbb{Z}$, ou

- por todos os números (exceto 0).

Um número que divide todos os outros é uma *unidade:*

Definição B.2 (Unidades). O elemento $\bar{x}$ em $\mathbb{Z}/m\mathbb{Z}$ é uma **unidade** se existe $\bar{y}$ em $\mathbb{Z}/m\mathbb{Z}$ tal que

$$\bar{y}\bar{x} = 1.$$

*O elemento $\bar{y}$ é o **inverso** de $\bar{x}$ e denotado por $\bar{x}^{-1}$.* Denotamos

$$(\mathbb{Z}/m\mathbb{Z})^* := \{ \text{ as unidades em } \mathbb{Z}/m\mathbb{Z}\}.$$

Por exemplo, a tabela de multiplicação de $\mathbb{Z}/4\mathbb{Z}$ mostra que

$$(\mathbb{Z}/4\mathbb{Z})^* = \{1,3\}.$$

pois $1 \cdot 1 = 1$ e $3 \cdot 3 = 1$; ao contrário, 2 não é uma unidade.

Definição. Os números inteiros a e b são **relativamente primos** se

$$\text{mdc}(a,b) = 1,$$

isto é, se nenhum número inteiro (fora 1) divide a e b.

Proposição B.3 (Caracterização de Unidades). *Dado um número inteiro x, sua imagem*

$$\bar{x} \text{ é } \textbf{\textit{unidade}} \text{ em } \mathbb{Z}/m\mathbb{Z} \iff x \text{ e } m \text{ são } \textbf{\textit{relativamente primos}}.$$

Por exemplo, $\text{mdc}(1,4) = \text{mdc}(3,4) = 1$, mas $\text{mdc}(2,4) = 2$. Pela proposição, concluímos que em $\mathbb{Z}/4\mathbb{Z}$ são unidades 1 e 3, mas 2 não é.

Demonstração (da Caracterização de Unidades):: Pelo **Algoritmo de Euclides Estendido**, existe u e v em $\mathbb{Z}$ tais que

$$ux + vm = \text{mdc}(x,m).$$

Então $\text{mdc}(x,m) = 1$ se, e somente se, existe u em $\mathbb{Z}$ tal que

$$ux \equiv 1 \bmod m$$

ou, equivalentemente,

$$\bar{u}\bar{x} = 1 \quad \text{em } \mathbb{Z}/m\mathbb{Z}.$$

Isto é, $\bar{x}$ é uma **unidade** em $\mathbb{Z}/m\mathbb{Z}$. □

Corolário B.4. Se p é um número primo, então

$$(\mathbb{Z}/p\mathbb{Z})^* = \{1, \ldots, p-1\}.$$

Isto é, todos os elementos, exceto 0*, são unidades.*

Demonstração: Se p é primo, então $\text{mdc}(x, p) = 1$ para $x = 1, \ldots, p-1$. □

Recordemo-nos de que um **corpo** é um anel cujos números são todos unidades ($\neq 0$). Isto é, um anel em que se pode dividir por qualquer número ($\neq 0$). Para um número primo p, pelo Corolário, $\mathbb{Z}/p\mathbb{Z}$ é um corpo, o **corpo com p elementos**, denotado por

$$\mathbf{F}_p := \mathbb{Z}/p\mathbb{Z}.$$

B.4. Teorema Chinês dos Restos

Teorema B.5. *Se m e n são coprimos, isto é* $\text{mdc}(m, n) = 1$*, então*

$$\mathbb{Z}/mn\mathbb{Z} = \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

Demonstração: Olha a aplicação natural

$$\pi\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

e observa que $\pi(x) = 0$ se, e somente se, $x \equiv 0 \bmod m$ e $x \equiv 0 \bmod n$, isto é, se, e tão-somente se, $\text{mmc}(n, m) = nm$ divide x. Logo o seu núcleo é $nm\mathbb{Z}$, e a aplicação induzido

$$\pi\colon \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

é injetora.

Como m e n são relativamente primos existem pelo **Algoritmo de Euclides Estendido** i e j em $\mathbb{Z}$ tal que $im + jn = 1$. Como a imagem é um ideal sobre $\mathbb{Z}$, para mostrar que a aplicação é sobrejetora, é suficiente mostrar que os seus geradores $(1,0)$ e $(0,1)$ são valores. Calculemos

$$jn \equiv im + jn = 1 \bmod m \quad \text{e} \quad jn \equiv 0 \bmod n$$

e

$$im \equiv 0 \bmod m \quad \text{e} \quad im \equiv im + jn = 1 \bmod n.$$

C. Topologia

Seja X um conjunto. Uma *topologia* sobre X é uma família τ de subconjuntos de X tal que

- os conjuntos $\emptyset$ e X são em τ,
- se U e V em τ, então $U \cap V$ em τ, e
- se $\{U_i : i \in I\}$ (para qualquer conjunto I) é em τ, então $\bigcup_{i\in I} U_i$ é em τ.

Exemplo.

- A topologia *discreta* sobre X é τ = { todos os subconjuntos de X}, e
- a topologia *mínima*, a menor topologia possível, é $\tau = \{\emptyset, X\}$.
- Seja X um conjunto. Uma aplicação $d\colon X \times X \to [0,\infty[$ é uma função *distância* ou *métrica* se
  - $d(x,y) = 0$ se, e somente se, $x = y$, (Hausdorff)
  - $d(x,y) = d(y,x)$, e (simetria)
  - $d(x,z) \leq d(x,y) + d(y,z)$. (desigualdade triangular)

  Um *espaço métrico* é um par (X,d) de um conjunto X e uma função distância d como acima.

  Para x em X e $r > 0$, defina a *bola aberta* respectivamente *fechada* com *raio* r e *centro* x (ou *à volta de* x) por

$$B(x;r) := \{y \in X : d(x,y) < r\} \quad \text{e} \quad \bar{B}(x;r) := \{y \in X : d(x,y) \leq r\}.$$

Uma *vizinhança à volta de um ponto* x em um espaço topológico X é um conjunto em X que contém um conjunto aberto U que contém x. Por exemplo, se X é um espaço métrico, então uma bola é uma vizinhança à volta do seu centro.

Uma *base* em um ponto x em X é uma família $\{U_i : i \in I\}$ de conjuntos abertos que contêm x tal que

$$\{\bigcup_{j\in J} U_j : J \subseteq I\} = \{U \text{ em } \tau : U \ni x\}.$$

Por exemplo, as bolas à volta de um ponto x em um espaço métrico X constituem uma base.

Um grupo é *topológico* se as suas operações $\cdot$ e $\cdot^{-1}$ são contínuas. Se $X = G$ é um grupo topológico, então basta por esta continuidade definir uma base $\{U_i : i \in I\}$ em volta de 1, o elemento neutro de G para obter τ inteira: A família τ_1 das uniões arbitrárias de $\{U_i : i \in I\}$ é à família de todos os conjuntos abertos que contêm 1. Seja g em G um ponto arbitrário. Como o traslado $g\cdot$ é contínuo e invertível, $g\mathcal{B}_1 = \tau_g$ onde τ_g é a família de todos os conjuntos que contêm g. Como $\tau = \bigcup_{g\in G} \tau_g$, concluímos que τ foi inteiramente restituída por $\mathcal{B}$.

Uma topologia é *totalmente desconexa* se todo subconjunto é *desconexo*, isto é, é união disjunta de dois subconjunto abertos. Por exemplo, $\mathbb{Z}_p$ é totalmente desconexo.

C.1. Sequências e Completude

Definição C.1. Uma sequência $(x_n : n \in \mathbb{N})$ em X *converge ao limite* x se para toda vizinhança U à volta de x existe n_0 tal que $x_n \in U$ para todo $n \geq n_0$. Isto é, *quase todos* os membros (= todos exceto um número finito) de (x_n) estão em U. Denote $\lim x_n = x$ ou $x_n \to x$ que (x_n) converge a x.

Observe que $x_n \to x$ se, e tão-somente se, $d(x_n, x) \to 0$.

Se X é *Hausdorff*, isto é, dois pontos diferentes são sempre contidos em duas vizinhanças disjuntas, então o limite é único.

Exemplo C.2 (Juros compostos). Ponhamos o nosso dinheiro em uma conta de poupança de um banco generoso, 100 reais por 100 dias com uma taxa de juros de 100%. Logo, teremos após 100 dias 200 reais na conta.

Agora proponhamos ao banco, no lugar de 100% uma única vez, pague 10% dez vezes, isto é após cada décimo dia um décimo dos juros. Logo, teremos

- após 10 dias $110 = 100 \cdot (1{,}1)$ reais na conta,
- após 20 dias $110 \cdot (1{,}10) = 121 = 100 \cdot (1{,}1)^2$ reais na conta, ... e
- finalmente, após 100 dias $100 \cdot (1{,}1)^{10} = 100 \cdot 2{,}5937424601 \approx 259$ no lugar de 200 reais na conta!

Tornamo-nos arbitrariamente ricos, ao diminuirmos os intervalos de pagamento mais e mais? Não! Se o banco pagasse, por exemplo, os juros (compostos) em 1.000.000 de intervalos, isto é, quase cada segundo, teríamos

$$100 \cdot (1{,}0000001^{1.000.000}) \approx 100 \cdot 2{,}71828 = 271{,}828$$

reais na conta. Isto é, observamos que a sequência $(1 + 1/n)^n$ converge ao *número de Euler* $e = 2{,}718\ldots$!

Um ponto x em X é um *ponto de acumulação* (ou *ponto de limite*) se é o limite de uma sequência (x_n)

- cujos membros x_n são todos diferentes, ou
- equivalentemente, tal que $\{x_n\}$ é infinito, ou
- equivalentemente, tal que $x \notin \{x_n\}$.

Um ponto x não é um ponto de acumulação, é um *ponto isolado*, se o conjunto unitário $\{x\}$ é aberto.

C.2. Funções Contínuas

Sejam X e Ω espaços topológicos e $f : X \to \Omega$. Seja a em X e ω em Ω.

Definição C.3. A função f *tem limite* ω em a, denotado por $\lim_{x\to a} f(x) = \omega$, se para toda vizinhança V à volta de ω em Ω existe uma vizinhança U à volta de a em X tal que $f^{-1}(V) \supseteq U$.
Ela é *contínua em a* se $\lim_{x\to a} f(x) = f(a)$.
Ela é *contínua* se é contínua em todos os pontos.

Observemos que a primeira condição $\lim_{x\to a} f(x) = \omega$ não é vazia se, e tão-somente se, a é ponto de acumulação.

Exemplo C.4.

- Todas as operações aritméticas + e ·, − e $\cdot^{-1}$ são contínuas.
- Toda função linear sobre um espaço vetorial finito é contínua.
- Todo polinómio é contínuo.
- Todas as funções especiais como exp, log, sin, cos, tan, arctan ...

Proposição C.5. *Sejam* X *e* Ω *espaços topológicos e* $f : X \to \Omega$. *Seja* a *em* X *e* $\alpha = f(a)$ *em* Ω. *São equivalentes as seguintes condições:*

(i) *A função* f *é contínua em* a.

(ii) *Para toda sequência* (x_n), *se* $x_n \to a$, *então* $f(x_n) \to \alpha$.

Demonstração: (i) $\implies$ (ii): Seja O em Ω uma vizinhança em torno de α. Pela hipótese, existe uma vizinhança B em torno de x contida em $f^{-1}\mathrm{O}$. Como $x_n \to x$, esta vizinhança B contém quase todos (isto é, todos exceto um número finito) membros de (x_n). Logo O contém quase todos os membros de $(f(x_n))$; isto é, $f(x_n) \to f(x) = \alpha$.

(ii) $\implies$ (i): Por contraposição: Seja a em X e, para todo $n \in \mathbb{N}$, seja V_n uma vizinhança de $f(a)$ tal que, para toda vizinhança U de a, existe x_n em X tal que x em U e $f(x) \notin \mathrm{V}_n$. Logo $x_n \to a$, mas $f(x_n) \not\to f(a)$. □

Proposição C.6. *Sejam* X *e* Ω *espaços topológicos e* $f : \mathrm{X} \to \Omega$. *São equivalentes as seguintes condições:*

(i) *A função f é contínua.*

(ii) *A pré-imagem sob f de todo subconjunto aberto é aberta.*

(iii) *A pré-imagem sob f de todo subconjunto fechado é fechada.*

Demonstração: Por definição, um conjunto G é aberto, se, e tão-somente se, para todo g em G existe uma bola em torno de g. Por Proposição C.5, uma função f é contínua se, e tão-somente se, a pré-imagem de toda vizinhança O em Ω contém uma vizinhança B em X à volta de x.

(i) $\implies$ (ii): Se f é contínua e Δ é um subconjunto aberto em Ω , então, para todo x em $\mathrm{D} = f^{-1}\Delta$, existe uma vizinhança B em X à volta de x; isto é, D é aberto.

(ii) $\implies$ (i): Se O uma vizinhança em Ω, então $f^{-1}\mathrm{O}$ é aberto, em particular, contém uma vizinhança B em X à volta de x.

(ii) $\iff$ (iii): A pré-imagem de um complemento é o complemento da pré-imagem. □

C.3. Compacidade

Definição. Seja K um subconjunto de X. Uma *cobertura* de K é uma família de subconjuntos abertos

$$\mathcal{U} = \{\mathrm{U}_i : i \in \mathrm{I}\}$$

de X tal que a sua união contenha K, isto é

$$\bigcup \mathcal{U} \supseteq \mathrm{K}.$$

O espaço K é *compacto* se toda cobertura $\mathcal{U}$ de K tem um *refinamento finito*; isto é, se contém uma subcobertura finita de K; isto é, existem $U_1, \ldots, U_n$ em $\mathcal{U}$ tal que $U_1 \cup \ldots \cup U_n \supseteq K$.

O fecho $\bar{S}$ de um conjunto S em X é o menor conjunto fechado que contém S, isto é,

$$\bar{S} = \bigcap \{ \text{todos os conjuntos fechados F que contêm S} \}.$$

O espaço X é *localmente compacto* se todo ponto x tem uma vizinhança U tal que o seu fecho $\bar{U}$ é compacto.

Exemplo C.7.

- Todos os conjuntos finitos são compactos
- a bola unitária aberta $B(0,1)$ não é compacto em $\mathbb{R}$: a cobertura $B(0,1-\frac{1}{n})$ contém nenhuma subcobertura finita.

Proposição C.8. *Seja* K *um subconjunto de* X.

(i) *Se* K *é compacto, então é fechado.*

(ii) *Se* K *é compacto, então é* limitado*, isto é, contido em uma bola.*

(iii) *Se* F *é fechado e* $F \subseteq K$*, então* F *é compacto.*

Demonstração: Ad (i): Mostremos que $K = K^-$ por contraposição: Seja $K \neq K^-$, isto é, exista $x \in K^- - K$, e mostremos que K não é compacto, isto é, existe uma cobertura $\mathcal{C}$ de K tal que $K \not\subseteq U_1 \cup \ldots \cup U_n$ para quaisquer $U_1, \ldots, U_n \in \mathcal{C}$. Seja

$$\mathcal{C} = \{C_n : n \in \mathbb{N}\} \quad \text{com } C_n := X - \bar{B}(x, \frac{1}{n})$$

uma cobertura de K.

Como $x \in K^-$, por **??**.(vi), toda bola em torno de x intersecta K. Em particular, toda bola da forma $\bar{B}(x, \frac{1}{n})$ intersecta K; equivalentemente, nenhum conjunto $C_n = X - \bar{B}(x, \frac{1}{n})$ em $\mathcal{C}$ contém K.

Como $C_1 \subseteq C_2 \subseteq \ldots$, a união de toda coleção finita $\{C_{i_1}, \ldots, C_{i_n}\}$ de $\mathcal{C}$ é contida em C_I para $i = \max\{i_1, \ldots, i_n\}$. Logo, nenhuma coleção finita de $\mathcal{C}$ cobre K.

Ad (ii): Por contraposição: Se K é ilimitado, isto é, contido em nenhuma bola $B(0,n)$ para n em $\mathbb{N}$, então $\mathcal{C} = \{B(0,n) : n \in \mathbb{N}\}$ é uma cobertura de K sem refinamento finito.

Ad (iii): Seja K compacto e F em K um subconjunto fechado. Seja $\mathscr{C}$ uma cobertura de F. Logo $\mathscr{D} = \mathscr{C} \cup \{X - F\}$ é uma cobertura de K; Como K é compacto, existe uma subcobertura finita de $\mathscr{D}$; logo uma subcobertura de $\mathscr{D}$ finita de F. □

Observação C.9. Todo subconjunto compacto é fechado e limitado; porém, a implicação inversa não vale em geral, isto é, existe um espaço métrico com um subconjunto limitado e fechado que não é compacto!

Por exemplo, seja $V = \mathbb{R}^{(\mathbb{N})}$ o espaço vetorial das sequências cujas entradas reais são quase todas nulas (= todas, exceto um número finito) e a sua norma dada por $\|(a_n)\| := \max\{|a_n| : n \in \mathbb{N}\}$; em particular, $\|\cdot\|$ induz a métrica $d(a,b) = \|a - b\|$. Seja $B = \{e_n : n \in \mathbb{N}\}$ a base canónica de V dada por e_n = a sequência cuja única entrada não-nula é a n-ésima com valor 1. O conjunto B é limitada e, como $d(e_n, e_m) = 1$ se (e tão-somente se) $n \neq m$, ele tem nenhum ponto de acumulação, logo é fechado. Porém, a cobertura de B dada pela coleção

$$\{B(e_n, 1) : n \in \mathbb{N}\}$$

não tem refinamento finito porque $B(e_n, 1) \cap B = \{e_n\}$.

Contudo, o Teorema de Heine-Borel abaixo mostrará que para $X = \mathbb{R}$ (e os seus produtos finitos), a implicação inversa vale: todo subconjunto limitado e fechado em $\mathbb{R}$ (e nos seus produtos finitos) é compacto.

Formulado por conjuntos fechados ao invés de abertos, um espaço X é compacto se, e tão-somente se, para toda família $\mathscr{F}$ de subconjuntos fechados, se todas as suas interseções finitas (ou equivalentes, todas as interseções entre dois subconjuntos em $\mathscr{F}$) são não-vazias, então a interseção total $\bigcap \mathscr{F}$ é não-vazia.

Uma coleção $\mathscr{C}$ de subconjuntos de um espaço métrico X tem a *Propriedade de Interseções Finitas* ou a *PIF*, se para todos $C_1, \ldots, C_n$ em $\mathscr{C}$ temos $C_1 \cap \ldots \cap C_n \neq \emptyset$. Por exemplo, a coleção dos conjuntos $X - B(0, \frac{1}{n})$ para n em $\mathbb{N}$ tem a PIF.

Proposição C.10. *Um subconjunto* K *em* X *é compacto se, e tão-somente se, para toda coleção $\mathscr{C}$ de subconjuntos fechados em* K*, se $\mathscr{C}$ tem a PIF, então $\bigcap \mathscr{C} \neq \emptyset$.*

Isto é, se a interseção entre todos os conjuntos de uma coleção $\mathscr{C}$ de subconjuntos fechados em K *é vazia, então já a interseção entre um número finito de conjuntos em $\mathscr{C}$ é vazia.*

Demonstração: Seja $\mathscr{F}$ uma coleção de subconjuntos em X tal que: Se $\mathscr{F}$ tem a PIF, então $\bigcap \mathscr{F} \neq \emptyset$; ou equivalentemente, por contraposição: Se $\bigcap \mathscr{F} = \emptyset$,

então há um número finito de conjuntos em $\mathcal{F}$ cuja interseção é vazia; ou equivalentemente, para a coleção dos complementos $\mathcal{C} = \{X - F : F \in \mathcal{F}\}$: Se $\bigcup \mathcal{C} = X$, então há um número finito de conjuntos em $\mathcal{C}$ cuja união é X.

Concluímos em particular que toda coleção de subconjuntos fechados em X tem a PIF se, e tão-somente se, X é compacto.

Todo subconjunto K em X é compacto se, e tão-somente se, é relativamente compacto, isto é, toda cobertura de conjuntos *abertos em* K, isto é, de conjuntos da forma $U \cap K$ para U um subconjunto aberto em X, tem um refinamento finito. Equivalentemente, se, e tão-somente se, toda coleção $\mathcal{F}$ de conjuntos *fechados em* K, isto é, de conjuntos da forma $K \cap F$ para F fechado em X, tem a PIF.

Se K é compacto, então é fechado por Proposição C.8. Logo um conjunto é relativamente fechado, isto é, fechado em K se, e tão-somente se, é fechado em X. Concluímos que toda coleção de subconjuntos fechados em K tem a PIF se, e tão-somente se, K é compacto. □

Teorema C.11 (Tychonoff). *Seja* $\{X_\alpha : \alpha \in A\}$ *uma família (de cardinalidade arbitrária) de espaços topológicos. Se todos os* X_α *são compactos, então o seu produto* $X := \prod_{\alpha \in A} X_\alpha$ *é compacto.*

Demonstração: Usemos a caracterização da compacidade pelos subconjuntos fechados, isto é, dada uma família $\mathcal{F}$ de subconjuntos fechados em X, se as suas interseções finitas são não-vazias, então a interseção total $\bigcap \mathcal{F}$ é não-vazia:

Seja $\mathcal{F}$ uma tal família de subconjuntos fechados em X que tem a PIF (= Propriedade das Interseções Finitas), isto é, cujas interseções finitas todas são não-vazias. O conjunto das famílias de subconjuntos (não necessariamente fechados) em X que têm a PIF, ordenado por inclusão, satisfaz a condição do Lema de Zorn; logo existe uma família máxima $\mathcal{F}^* \supseteq \mathcal{F}$.

Observação: Por $\mathcal{F}^*$ ser máxima, se G é um subconjunto em X tal que $G \cap F \neq \emptyset$ para todo F em $\mathcal{F}^*$, então G em $\mathcal{F}^*$. Em particular, dados $F_1, \ldots, F_n$ em $\mathcal{F}^*$, não apenas $F := F_1 \cap \ldots \cap F_n \neq \emptyset$, mas F em $\mathcal{F}^*$. (Isto é, $\mathcal{F}^*$ é fechada sob interseções finitas.)

Para uma coordenada $\alpha \in A$ denote $p_\alpha : X \to X_\alpha$ a projeção. Como $\mathcal{F}^*$ tem a PIF, também a família $\mathcal{F}^*_\alpha := \{p_\alpha(F) : F \in \mathcal{F}^*\}$ das suas projeções tem a PIF. Por X_α ser compacto, a interseção total $I_\alpha = \bigcap\{\overline{F_\alpha} : F_\alpha \in \mathcal{F}^*_\alpha\}$ dos seus fechamentos é não-vazia. Seja, pelo axioma de escolha, $x = (x_\alpha)_{\alpha \in A} \in \prod_{\alpha \in A} I_\alpha$.

Temos x em $\prod_{\alpha \in A} \bigcap \mathcal{F}^*_\alpha \supseteq \bigcap \mathcal{F}^*$; para completar a demonstração, precisa de demonstrar que x já é em cada F em $\mathcal{F}$. Basta mostrar que toda vizinhança à volta de x contém um conjunto que é contida em $\mathcal{F}^*$: Pela PIF todo F em $\mathcal{F}^*$

intersecta todo conjunto em $\mathcal{F}^*$; em particular, toda vizinhança de x, logo o seu fecho $\overline{\mathrm{F}}$ contém x. Em particular, todo F em $\mathcal{F}$, sendo fechado, contém x.

Seja $\mathrm{V} \ni x$ aberto. Por definição da topologia do produto, existe um número finito de subconjuntos abertos $\mathrm{U}_{\alpha_1}, \ldots, \mathrm{U}_{\alpha_n}$ de $\mathrm{X}_{\alpha_1}, \ldots, \mathrm{X}_{\alpha_n}$ tais que que

$$\mathrm{U} := \mathrm{U}_{\alpha_1} \times \ldots \times \mathrm{U}_{\alpha_n} \times \prod_{\alpha \neq \alpha_1, \ldots, \alpha_n} \mathrm{X}_\alpha$$

contém x e é contido em V. Como $\mathrm{U} \subseteq \mathrm{V}$, basta demonstrar que U em $\mathcal{F}^*$. Como

$$\mathrm{U} = p_{\alpha_1}^{-1}\mathrm{U}_{\alpha_1} \cap \ldots \cap p_{\alpha_n}^{-1}\mathrm{U}_{\alpha_n}$$

basta, pela observação acima que $\mathcal{F}^*$ é fechada sob interseções finitas, demonstrar que $p_\alpha^{-1}(\mathrm{U}_\alpha)$ em $\mathcal{F}^*$ para todo conjunto aberto U_α e α em A. Pela observação acima, basta demonstrar que $p_\alpha^{-1}(\mathrm{U}_\alpha) \cap \mathrm{F} \neq \emptyset$ para todo F em $\mathcal{F}^*$. Como x_α em $\mathrm{I}_\alpha \subseteq \overline{p_\alpha(\mathrm{F})}$ e em U_α, em particular, $\mathrm{F} \cap p_\alpha^{-1}(\mathrm{U}_\alpha) \neq \emptyset$. □

O Teorema de Tychonoff, em geral, equivale ao Axioma da Escolha. Se os fatores são *Hausdorff*, isto é, cada sequência que converge tem um único limite, por exemplo, espaços métricos, então não precisa dele.

C.4. Compacidade Sequencial

Um subconjunto de um espaço topológico é *denso* se intersecta todo subconjunto aberto. Um espaço topológico é *separável* se tem um subconjunto enumerável denso. Um espaço topológico é *sequencialmente compacto* se toda sequência tem uma subsequência convergente.

Exemplo C.12. Nem todo espaço (topológico) compacto é sequencialmente compacto: Por exemplo, a bola unitária do dual $c^b(\mathbb{N})^*$ é compacta para a topologia fraca* pelo Teorema de Alaoglu. Porém, a sequência $([f \mapsto f(n)] : n \in \mathbb{N})$ tem nenhuma subsequência convergente.

Contudo, um espaço métrico é compacto se, e tão-somente se, é topologicamente compacto:

Teorema C.13. *Se* X *é um espaço métrico, então* X *é sequencialmente compacto se, e tão-somente se,* X *é compacto.*

Demonstração: Mostremos a implicação $\Longleftarrow$ por contraposição, isto é, se não é sequencialmente compacto, então não é compacto: Seja (x_n) tal que

nenhuma subsequência converge, isto é, para todo x em X existe $\epsilon(x) > 0$ tal que $\mathrm{B}(x, \epsilon(x)) \cap \{x_n\}$ é finita. Se a cobertura $\{\mathrm{B}(x, \epsilon(x)) : x \in \mathrm{X}\}$ tivesse um refinamento finito, então (x_n) seria finita, em particular, convergente; em contradição ao que nenhuma subsequência converge. Logo X não é compacto.

Mostremos a implicação $\Longrightarrow$ em três passos:

(i) *Lema do Número da Cobertura de Lebesgue*: Para toda cobertura $\{\mathrm{U}_i\}$ existe $\epsilon > 0$ tal que, para todo x em X, existe i tal que $\mathrm{B}(x, \epsilon) \subseteq \mathrm{U}_i$.

(ii) *Cota Total*: Para todo $\epsilon > 0$ existe uma cobertura finita de X por bolas $\mathrm{B}(x, \epsilon)$.

(iii) O espaço topológico X é compacto.

Ad (i): Por contraposição: Se existe uma cobertura $\{\mathrm{U}_i\}$ tal que, para todo n, existe x_n em X tal que $\mathrm{B}(x_n, 1/n)$ é contida em nenhum U_i, então (x_n) tem nenhuma subsequência convergente: Se (x_n) tivesse uma subsequência (x_{n_k}) tal que $x_{n_k} \to x$, então

- existe $\epsilon > 0$ e i_0 tal que $\mathrm{B}(x, \epsilon) \subseteq \mathrm{U}_{i_0}$, e
- existe N tal que x_{n_k} em $\mathrm{B}(x, \epsilon)$ para todo $n_k \geq \mathrm{N}$.

Logo, existe k suficientemente grande tal que $\mathrm{B}(x_{n_k}, \frac{1}{n_k}) \subseteq \mathrm{B}(x, \epsilon) \subseteq \mathrm{U}_{i_0}$; em contradição à escolha de x_{n_k}!

Ad (ii): Por contraposição: Se existe $\epsilon > 0$ tal que para todo n e $a_1, \ldots, a_n$ em X existe

$$a \notin \mathrm{B}(a_1, \epsilon) \cup \ldots \cup \mathrm{B}(a_n, \epsilon),$$

então define a sequência (x_n) por uma escolha arbitrária de x_1 e a escolha de

$$x_{n+1} \notin \mathrm{B}(x_1, \epsilon) \cup \ldots \cup \mathrm{B}(x_n, \epsilon).$$

Se (x_n) tivesse uma subsequência (x_{n_k}) tal que $x_{n_k} \to x$, então existiria para $\frac{\epsilon}{2} > 0$ um N tal que x_{n_k} em $\mathrm{B}(x, \frac{\epsilon}{2})$ para todo $n_k \geq \mathrm{N}$. Logo, $\mathrm{d}(x_{n_k}, x_{\mathrm{N}}) \leq \mathrm{d}(x_{n_k}, x) + \mathrm{d}(x, x_{\mathrm{N}}) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, isto é,

$$x_{n_k} \in \mathrm{B}(x_{\mathrm{N}}, \epsilon)$$

para todo $n_k > \mathrm{N}$; em contradição à escolha de x_{n_k} para $n_k > \mathrm{N}$!

Ad (iii): Seja $\{\mathrm{U}_i\}$ uma cobertura de X. Seja $\epsilon > 0$ dado por (i) e $x_1, \ldots, x_n$ por (ii). Logo o refinamento dado pelos $\mathrm{U}_{i_1}, \ldots, \mathrm{U}_{i_n}$ que contêm respetivamente $x_1, \ldots, x_n$ cobre X. □

Lema C.14. *Toda sequência infinita em $\mathbb{R}$ tem uma subsequência monótona infinita.*

Demonstração: Um índice n é um *pico* se $x_n > x_{n+1}, x_{n+2}, \ldots$. Se tem uma infinitude de picos $n_1, n_2, \ldots$, então $x_{n_1}, x_{n_2}, \ldots$, é uma subsequência monotonamente decrescente. Caso contrário, seja n_0 o último pico. Isto é, para todo $n > n_0$ existe $m > n$ com $x_m \geq x_n$. Seja $n_1 = n_0 + 1$; logo, existe $n_2 > n_1$ tal que $x_{n_2} \geq x_{n_1}$; semelhantemente para n_2 no lugar de n_1, e assim por diante se constrói uma subsequência (x_{n_k}) monotonamente crescente. □

Lema C.15. *Toda sequência real monótona e limitada converge; caso crescente ao seu supremo, caso decrescente ao seu ínfimo.*

Demonstração: Seja (x_n) monotonamente crescente e limitada. Seja $s = \sup\{x_n\}$ e $\epsilon > 0$. Logo existe N tal que $|s - x_N| < \epsilon$. Pela monotonia, a fortiori $|s - x_n| < \epsilon$ para todo $n > N$. Isto é, $x_n \to s$.

Analogamente para uma sequência monotonamente decrescente. □

Corolário C.16. *Seja (x_n) é uma sequência em $\mathbb{R}$. Se (x_n) é limitado, então tem uma subsequência convergente.*

Demonstração: Por Lema C.14 e Lema C.15. □

Corolário C.17 (Teorema de Heine-Borel)**.** *Um subconjunto* K *de $\mathbb{R}$ é compacto se, e tão-somente se, é limitado e fechado.*

Demonstração: Como espaço métrico $\mathbb{R}$ é compacto se, e tão-somente se, é sequencialmente compacto pelo Teorema C.13.

$\Longleftarrow$: Pelo **??**, toda sequência tem uma subsequência convergente. Como K é fechado, este limite é pela **??** em K. Logo K é sequencialmente compacto.

$\Longrightarrow$: Pela Proposição C.8, todo subconjunto compacto é fechado e limitado. □

D. O Lema de Zorn

Uma *ordenação parcial* é uma relação $\leq$ sobre um conjunto X que é

- *reflexiva*, isto é, $x \leq x$,
- *anti-simétrica*, isto é, se $x \leq y$ e $y \leq x$, então $x = y$, e
- *transitiva*, isto é, se $x \leq y$ e $y \leq z$, então $x \leq z$.

Se $x \leq y$, digamos que x é *menor* que y. Denote $x < y$ que $x \leq y$ e $x \neq y$. Denote $y \geq x$ que $x \leq y$, digamos que y é *maior* que x, e denote $y > x$ que $x < y$.

Uma *ordenação total* ou *cadeia* é uma ordenação parcial que satisfaz, além disso, que ou $x \leq y$, ou $y \leq x$.

Um elemento x_0 em X é

- *mínimo* se não existe $x < x_0$ em X, e
- *máximo* se não existe $x > x_0$ em X.

Um elemento x_0 em X é

- o *menor* elemento se $x_0 \leq x$ para todo x em X, e
- o *maior* elemento se $x_0 \geq x$ para todo x em X.

Se X é totalmente ordenado, então todo elemento mínimo é o menor elemento e todo elemento máximo é o maior elemento.

Uma *boa ordenação* é uma ordenação total tal que todo subconjunto não-vazio tem um menor elemento.

Seja Y um subconjunto de X. Um elemento x em X é

- uma cota *inferior* se $x \leq y$ para todo y em Y, e
- uma cota *superior* se $x \geq y$ para todo y em Y.

Se X é bem-ordenado e existe uma cota superior de Y que não pertence a Y, então existe um *supremo* s de Y, uma menor cota superior entre todas as cotas superiores de Y que não pertence a Y, e $Y = \{x \in X : x < s\}$.

Um subconjunto Y é um *segmento inicial* (ou *fechado*) em um conjunto parcialmente ordenado X, se, para todo y em Y, se $x \leq y$, então x em Y. Denote $X \leq Y$ que X é um segmento inicial em Y (e $Y \geq X$ que $X \leq Y$), e $X < Y$ que $X \leq Y$ e $X \neq Y$ (e $Y > X$ que $X < Y$).

D.1. Demonstração

A união de uma cadeia de conjuntos bem-ordenados é bem-ordenada:

Lema D.1. *Seja* X *parcialmente ordenado e* $\mathcal{F}$ *uma coleção de subconjuntos de* X *bem-ordenados. Se para todo* C *e* D *em* $\mathcal{F}$*, ou* $C \leq D$*, ou* $C \geq D$*, então* $E = \bigcup \mathcal{F}$ *é bem-ordenado e* $E \geq C$ *para todo* C *em* $\mathcal{F}$*.*

O Lema de Zorn frequentemente é formulado com uma condição mais restritiva, isto é: Todo subconjunto totalmente ordenado, ao invés de apenas todo subconjunto bem-ordenado (= subconjunto totalmente ordenado cujos subconjuntos não-vazios todos têm um elemento menor), tem uma cota superior. Porém, na prática, esta restrição revela-se irrelevante. Demonstremos a formulação mais geral:

Teorema D.2 (Lema de Zorn). *Dado um conjunto* X *não-vazio e parcialmente ordenado. Se todo subconjunto bem-ordenado tem uma cota superior, então* X *tem um elemento máximo.*

Demonstração (segundo H. Kneser): Suponhamos o contrário, isto é, para cada elemento x em X exista $y > x$. Logo, pela hipótese, para cada subconjunto bem-ordenado $C \subseteq X$ existe um supremo de C que não pertence a C, denotado por $g(C)$. Define um *g-conjunto* como um subconjunto bem-ordenado $C \subseteq X$ tal que todo $c \in C$ é supremo de $\{c' \in C : c' < c\}$, isto é, $c = g(\{c' \in C : c' < c\})$.

Afirmação: Se C e D são g-conjuntos, então, ou $C \leq D$, ou $C \geq D$.

Prova: Seja $W = \bigcup\{B \subseteq X : B \leq C \text{ e } B \leq D\}$. Como todos os subconjuntos B são segmentos iniciais, também W é um segmento inicial; logo $W \leq C$ e $W \leq D$, e W é o maior subconjunto com esta propriedade. Se $W = C$ ou $W = D$, então a demonstração é finalizada. Suponhamos o contrário, isto é, $W < C$ e $W < D$, e sejam c em C e d em D os supremos de W em C respectivamente D, isto é,

$$\{c' \in C : c' < c\} = W = \{d' \in D : d' < d\}.$$

Como C e D são g-conjuntos, $c = g(W) = d$. Põe $W' := W \cup \{g(W)\}$; é um g-conjunto $> W$ que satisfaz $W' \leq C$ e $W' \leq D$: contradição a W ser máximo com esta propriedade!

Põe $W := \bigcup\{$ todos os g-conjuntos$\}$. A união de uma cadeia de g-conjuntos é um g-conjunto: Pela Afirmação as condições de Lema D.1 são satisfeitas, logo W é bem-ordenado. Além disto, como união de g-conjuntos, W é um g-conjunto; é o maior g-conjunto em X. Porém, $W' = W \cup \{g(W)\}$ é um g-conjunto $> W$: contradição a W ser máximo com esta propriedade! □

O Lema de Zorn equivale ao Axioma da Escolha. O nome faz referência ao matemático Max Zorn, mas sua primeira formulação se deve ao matemático polonês Kazimierz Kuratowski.

D.2. Uma Aplicação

Como aplicação do *Lema de Zorn*, provemos que *todo espaço vetorial* V *possui uma base*, um subconjunto de vetores linearmente independentes que gera V. O Lema de Zorn fornecerá um subconjunto B de vetores linearmente independentes máximo; provemos que tal B é uma base, isto é, gera V:

Lema D.3. *Seja* V *um espaço vetorial e* B *um subconjunto de vetores linearmente independentes em* V. *Se* B *é máximo, então* B *gera* V.

Demonstração: Seja $x \in V - B$ não-nulo. Como B é máximo, o superconjunto próprio de B definido por $B \cup \{x\}$ contém vetores linearmente dependentes; isto é, existe uma combinação linear $\alpha_1 v_1 + \cdots + \alpha_n v_n + \beta x = 0$ com alguns coeficientes não-nulos. Logo $\beta \neq 0$, caso contrário já B teria tido vetores linearmente independentes. Portanto, $x = \frac{-\alpha_1}{\beta} v_1 + \cdots + \frac{-\alpha_n}{\beta} v_n$. Isto é, B gera V. □

Teorema D.4 (Todo espaço vetorial tem uma base). *Seja* V *um espaço vetorial. Se* L *é um conjunto de vetores linearmente independentes em* V*, então existe um conjunto de vetores linearmente independentes máximo em* V *que contém* L.

Demonstração: Para aplicar o Lema de Zorn, construamos um conjunto e definir uma relação de ordem parcial: Como desejamos aumentar um conjunto de vetores linearmente independentes, definamos

$$X = \{x \in P(V) : x \supseteq L \text{ e } x \text{ consiste de vetores linearmente independentes } \}$$

e equipamos X com a ordem parcial natural dada por $x \leq y$ se $x \subseteq y$. O conjunto X não é vazio, porque $L \in X$.

Provemos que todo subconjunto totalmente ordenado de X tem uma cota superior: Seja $T \subseteq X$ não-vazio e totalmente ordenado.

Põe $Q = \bigcup\{x \in T\}$. Pela sua definição, Q é uma *cota superior*, isto é, $x \subseteq Q$ para todo $x \in T$. Para concluir, falta verificar que Q em X, isto é, $L \subseteq Q$ e que todos os vetores em Q são linearmente independentes:

Como L é um subconjunto de todo elemento de X, então L é um subconjunto de todo elemento de T. Logo, $L \subseteq Q$

Verifiquemos que Q é linearmente independente: Seja $\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n = 0$ uma combinação linear de elementos distintos de Q. Como Q é uma união de conjuntos, existe para todo $i = 1, \ldots, n$ um x_i em T tal que $v_i \in x_i$. Como T é totalmente ordenado, dentre os x_i existe um deles x_{i_0} que é superconjunto de todos os outros. Logo $v_i \in x_{i_0}$ para todo $i = 1, \ldots, n$; portanto $\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n = 0$ é uma combinação linear de vetores de x_{i_0}. Como x_{i_0} é um conjunto de vetores linearmente independentes, obtemos $\alpha_1, \ldots, \alpha_n = 0$. Isto é, todos os vetores em Q são linearmente independentes.

Ora se aplica o *Lema de Zorn* ao conjunto X para obter um elemento máximo B. □

D.3. História

Kazimierz Kuratowski provou em 1922 [1] uma versão menos genérica do Lema de Zorn (usando conjuntos parcialmente ordenados pela inclusão e fechados relativamente à união arbitrária de cadeias bem-ordenadas). O Lema na sua forma atual (usando qualquer relação de ordem, e usando qualquer cadeia totalmente ordenada) foi proposto independentemente por Max Zorn em 1935. [2] Zorn propôs esta formulação como um novo axioma da teoria dos conjuntos, como um substituto do teorema da bem-ordenação, exibiu algumas das suas aplicações na álgebra. Ele também prometeu mostrar a equivalência entre o seu lema e o *axioma da escolha* em outro artigo, mas nunca o escreveu.

[1] Casimir Kuratowski, *Une méthode d'élimination des nombres transfinis des raisonnements mathématiques*, Fundamenta Mathematicae 3 (1922), pp. 76–108. icm

[2] Max Zorn, *A remark on method in transfinite algebra*, Bulletin of the American Mathematical Society 41 (1935), no. 10, pp. 667–670.

E. Teorema de Hasse-Minkowski

Recordemos:

Definição. Uma *forma quadrática* é um polinômio da forma $\sum_{i,j=1,\dots,n} a_{i,j} X_i X_j$.

Dizemos que um polinômio $P(X_1,\dots,X_n)$ com coeficientes em $\mathbb{Q}$ *tem um zero (não-trivial)* (ou *representa zero*) sobre $\mathbb{Q}$ (ou $\mathbb{Q}_p$ ou $\mathbb{R}$) se existe $(x_1,\dots,x_n)$ diferente de 0 com $P(x) = 0$ cujas entradas são em $\mathbb{Q}$ (ou $\mathbb{Q}_p$ ou $\mathbb{R}$).

Para demonstrar o Teorema de Hasse-Minkowski, seguimos [Ser70, Chapitre IV] e introduzimos à Teoria das Formas Quadráticas: Ao invés de considerar uma forma quadrática como polinômio, definiremo-la como aplicação entre espaços vetoriais. Em seguida, simplificaremos o problema de uma forma quadrática geral a uma forma quadrática *não-degenerada com base ortogonal.*

E.1. Formas Quadráticas Gerais

Seja $\mathbf{K}$ um corpo e V um espaço vetorial de dimensão finita sobre $\mathbf{K}$. Uma *forma quadrática* é uma aplicação $q : V \to \mathbf{K}$ tal que

- vale $q(\alpha v) = \alpha^2 q(v)$ para todo α em $\mathbf{K}$ e v em V, e
- a função $v, w \mapsto q(v+w) - q(v) - q(w)$ é bilinear.

Em um corpo da característica $\neq 2$, formas quadráticas e formas bilineares simétricas correspondem um-a-um por

$$q \mapsto [v, w \mapsto v \cdot w := 1/2(q(v+w) - q(v) - q(w))].$$

Chamamos um espaço vetorial sobre que uma forma quadrática é definida um *módulo quadrático.* Se q' é uma forma quadrática sobre V′ e q'' é uma forma quadrática sobre V″, uma aplicação $f : V' \to V''$ é um *homomorfismo (métrico)* se f é um homomorfismo entre espaços vetoriais e

$$q'' \circ f = q';$$

então $f(v) \cdot f(w) = v \cdot w$.

Por escolha de uma base $\boldsymbol{e} = (e_1,\dots,e_n)$ em V, uma forma quadrática q corresponde à matriz simétrica $A = (a_{i,j})$ definida por $a_{i,j} = e_i \cdot e_j$. Se $x = x_1 e_1 + \cdots + x_n e_n$ em V, então

$$q(x) = \sum_{i,j=1,\dots,n} a_{i,j} x_i x_j.$$

Seja A' a matriz de q com respeito a base e'. Se a base e' é trocada pela base e'', então a matriz A'' de q com respeito a e'' é dada por

$$A'' = XA'X^t$$

onde X designa a matriz que transforma e' em e'' e X^t a sua matriz *transposta* (cujas colunas são as linhas de X, ou, equivalentemente, X^t é a matriz obtida por reflexão de X ao longo da diagonal).

Ortogonalidade. A *ortogonalidade* permite decompor uma forma quadrática em várias formas quadráticas sobre um espaço vetorial de dimensões menores.

Dois vetores v e w são *ortogonais* se $v \cdot w = 0$. O *complemento ortogonal* de um subespaço vetorial U em V é definido por $U^0 := \{v \in V : v \cdot u = 0 \text{ para cada } u \text{ em } U\}$. Dois subespaços vetoriais U e W são *ortogonais* se $U \subseteq W^0$, isto é, $u \cdot w = 0$ para todo u em U e w em W.

Se $V^0 = 0$, a forma quadrática chama-se *não-degenerada*. (Equivalentemente, $\det A \neq 0$ para uma (ou toda) matriz A que define q.)

Escrevemos $V = V_1 \hat{\oplus} \cdots \hat{\oplus} V_n$ para subespaços vetoriais $V_1, \ldots, V_n$ em V se

- $V = V_1 \oplus \cdots \oplus V_n$, isto é, V é a soma direta de $V_1, \ldots, V_n$, e
- cada par de subespaços em $V_1, \ldots, V_n$ é ortogonal.

Vetores isótropos. Um vetor v em V é *isótropo* se $v \cdot v = 0$. Um subespaço é *isótropo* se todos os seus vetores são isótropos. Quer dizer, U é isótropo se, e somente se, $U \subseteq U^0$, ou, se, e somente se, $q_{\restriction U} = 0$.

Definição. Um *plano hiperbólico* é um espaço vetorial com uma forma quadrática que tem uma base de vetores x e y que

- são isótropos, e
- satisfazem $x \cdot y \neq 0$

Após multiplicação por $1/x \cdot y$, a matriz que define um plano hiperbólico é dada por

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Proposição E.1. *Seja q uma forma quadrática. Se ela é não-degenerada, então existe para cada vetor isótropo $\neq 0$ um plano hiperbólico que o contém.*

Demonstração: Seja V o espaço vetorial sobre que q é definido e x um vetor isótropo. Como q é não-degenerada, existe z em V tal que $x \cdot z = 1$. O vetor $y = 2z - (z \cdot z)x$ é isótropo e $x \cdot y = 2$. O plano hiperbólico é gerado por x e y. □

Bases Ortogonais. Uma base $e_1, \ldots, e_n$ de um módulo quadrático (q, V) é *ortogonal* se $\mathrm{V} = \mathbf{K}e_1 \hat{\oplus} \cdots \hat{\oplus} \mathbf{K}e_n$. A matriz de q com respeito a esta base tem a figura

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

Isto é, para $x = x_1e_1 + \cdots + x_ne_n$ em V, vale $q(x) = a_1x_1^2 + \cdots + a_nx_n^2$.

Teorema E.2. *Todo módulo quadrático q não-degenerado tem uma base ortogonal.*

Demonstração: Por indução sobre o posto de V. Se todo vetor x fosse isótropo, então $q = 0$ e a forma bilinear é zero; em particular, ela é degenerada. Logo, como q é suposta não-degenerada, existe um vetor x anisótropo. Vale $\mathrm{V} = \mathbf{K} \cdot x \hat{\oplus} \mathrm{U}$ onde U é o complemento ortogonal de x em V. Para aplicar a hipótese de indução a U para concluir, falta verificar que o módulo quadrático U seja não-degenerado. Seja u_0 em U. Se u_0 fosse ortogonal a U, isto é, $u_0 \cdot u = 0$ para todo u em U, então

$$u_0 \cdot v = u_0 \cdot (\lambda x + u) = u_0 \cdot \lambda x + u_0 \cdot u = 0$$

para todo $v = \lambda x + u$ em V; isto é, u_0 é ortogonal a V, isto é, V é degenerado. Logo, como V é suposto não-degenerado, U não é degenerado. □

Traslações. Duas formas quadráticas q' e q'' são *isomorfas*, denotado por $q' \sim q''$, se os seus módulos quadráticos são isomorfos, isto é, se existe um isomorfismo métrico entre eles. Por exemplo, um módulo quadrático (V, q) é um plano hiperbólico se, e somente se,

$$q \sim \mathrm{XY} \sim \mathrm{X}^2 - \mathrm{Y}^2.$$

Em particular, q tem um zero se, e somente se, existe um vetor isótropo.

Dados dois módulos quadráticos (V', q') e (V'', q''), denota

$$q = q' + q''$$

a forma quadrática sobre $\mathrm{V}' \oplus \mathrm{V}''$ que se restringe a q' sobre V′ e a q'' sobre V″. Apontamos as seguintes reformulações de Proposição E.1 e Teorema E.2:

Proposição (E.1'). *Se f tem um zero e é não-degenerada, então $f \sim f' + g$ com f' uma forma sobre um plano hiperbólico.*

Teorema (E.2'). *Seja f uma forma quadrática não-degenerada em n variáveis. Existem $a_1, \ldots, a_n$ tal que $f \sim a_1 X_1^2 + \cdots + a_n X_n^2$.*

E.2. Normas

Seja $\mathbf{K}$ ou o corpo $\mathbb{R}$, ou o corpo $\mathbb{Q}_p$ para um número primo p e $|\cdot|$ o seu valor absoluto. Seja b em $\mathbf{K}^*$ e $\mathbf{K}(\sqrt{-b})$ o corpo valorado (por Teorema 6.4) obtido por adjunção dos zeros de $X^2 - b$ à $\mathbf{K}$. Recordemos a *norma* $\mathrm{N}\colon \mathbf{K}(\sqrt{-b})^* \to \mathbf{K}^*$ definida em Equação (6.1). Neste caso,

$$\mathrm{N}(x + \sqrt{-b}y) = |x^2 - by^2|.$$

O que importará particularmente para nós é que a norma é multiplicativa, isto é:

$$\mathrm{N}(z'z'') = \mathrm{N}(z') \cdot \mathrm{N}(z'')$$

para $z' = x' + \sqrt{-b}y'$ e $z'' = x'' + \sqrt{-b}y''$ em $\mathbf{K}(\sqrt{-b})^*$.

Proposição E.3. *Sejam a e b em $\mathbf{K}$. A equação*

$$X^2 = aY^2 + bZ^2$$

tem uma solução $\neq (0,0,0)$ se, e somente se, a é uma norma em $\mathbf{K}(\sqrt{b})$.

Para a demonstração do Teorema de Hasse-Minkowski, usaremos que se esta equação tem uma solução não-trivial para os pares a' e b e a'' e b, então tem uma solução para o par $a'a''$ e b.

E.3. Formas Quadráticas sobre $\mathbb{Q}$

Seja $f = \sum_{i,j=1,\ldots,n} a_{i,j} X_i X_j$ uma forma quadrática com coeficientes em $\mathbb{Q}$. Dizemos que

- a forma f *representa zero* não-trivial se existe $(x_1, \ldots, x_n) \neq 0$ com entradas em $\mathbb{Q}$ tal que $f(x) = 0$,
- a forma f_p para um número primo p *representa zero* se existe $(x_1, \ldots, x_n) \neq 0$ com entradas em $\mathbb{Q}_p$ tal que $f(x) = 0$, e

- a forma f_∞ *representa zero* se existe $(x_1, \ldots, x_n) \neq 0$ com entradas em $\mathbb{R}$ tal que $f(x) = 0$.

Teorema E.4 (Hasse-Minkowski). *A forma f representa zero se, e somente se,*

- *a forma f_p representa zero para todo número primo p e*
- *a forma f_∞ representa zero.*

Nota. O teorema de Hasse-Minkowski estende-se a extensões finitas de $\mathbb{Q}$, como foi demonstrado por Hasse por métodos p-ádicos (como o caso $\mathbb{Q}$, que, porém, foi demonstrado pouco antes por Minkowski pela sua teoria de géneros). Porém, não se estende a polinômios homogêneos de grau > 2. Por exemplo, Selmer demonstrou que

$$3X^3 + 4Y^3 + 5Z^3$$

tem um zero (diferente) sobre cada $\mathbb{Q}_p$ e sobre $\mathbb{R}$, mas não sobre $\mathbb{Q}$.

Quanto à necessidade, se f representa zero, o mesmo zero mostra que f_p para todos números primos p e f_∞ representa zero.

Para demonstrar a suficiência, observamos que se f é degenerada sobre $\mathbb{Q}_p$ ou $\mathbb{R}$, então sobre $\mathbb{Q}$ (visto, por exemplo, pelo critério da nulidade da determinante da matriz $(a_{i,j})$). Por isso, podemos supor que f é não-degenerada e escrever por Teorema E.2'

$$f \sim a_1 X_1^2 + \cdots + X_n^2.$$

Substituindo f por $a_1 f$, podemos supor que $a_1 = 1$. Demonstramos somente os casos de posto $n = 2$ e $n = 3$; mencionamos apenas que o caso $n = 4$ pode ser reduzido ao caso $n = 3$, e o caso $n \geq 5$ pode ser reduzido por indução ao caso $n = 4$.

(i) O caso $n = 2$:

Tem-se $f = X^2 - aY^2$. Como f_∞ representa zero, vale $a > 0$. Escrevemos

$$a = \prod_p p^{v_p(a)}.$$

Como cada f_p representa zero, segue que cada $v_p(a)$ é par. Por isso, a é quadrático (e > 0). Logo, f representa zero.

(ii) O caso $n = 3$ (por Legendre):

Seja $f = \mathrm{X}^2 - a\mathrm{Y}^2 - b\mathrm{Z}^2$.

Dizemos que x em $\mathbb{Z}$ é *livre de quadrados* se não existe y em $\mathbb{Z}$ tal que y^2 divida x; isto é, $v_p(x) = 0$ ou 1 para todo número primo p. Por escolha judiciosa da base, suponhamos que a e b são inteiros e *livre de quadrados*, isto é, $v_p(a)$ e $v_p(b) = 0$ ou 1 para todo número primo p. (Se um dos coeficientes c, ou a, ou b, não é inteiro, isto é, tem denominador d, então multiplicamos o vetor de base correspondente por d; logo o coeficiente correspondente da forma quadrática na nova base é multiplicado por d^2, em particular inteiro. Se um dos coeficientes c, ou a, ou b não é livre de quadrados, $c = \tilde{c}^2 c_0$ onde c_0 é livre de quadrados, então multiplicamos o vetor de base correspondem por $\tilde{c}^{-1}$; logo o coeficiente correspondente da forma quadrática na nova base é dividido por $\tilde{c}^2$, isto é, livre de quadrados.) Por escolha, suponhamos que $|a| \leq |b|$.

Demonstramos a existência de um zero de f por indução sobre $m = |a|+|b|$. Para iniciá-la, se $m = 2$, então

$$f = \mathrm{X}^2 \pm \mathrm{Y}^2 \pm \mathrm{Z}^2.$$

Como f_∞ representa zero, é impossível que $f = \mathrm{X}^2+\mathrm{Y}^2+\mathrm{Z}^2$ representa zero. Em todos os outros casos, f representa zero (por exemplo, $10^2 = 6^2 + 8^2$).

Seja $m > 2$, isto é, $|b| \geq 2$. Escreve $b = \pm p_1 \cdots p_k$ para números primos distintos. Seja p um destes.

Afirmação: O inteiro a é um quadrado módulo p.

Demonstração: Ou $a \equiv 0 \bmod p$, e a asserção vale, ou a é uma unidade módulo p. Por hipótese, existem x, y, z em $\mathbb{Q}_p$ tais que

$$x^2 - ay^2 - bz^2 = 0. \qquad (*)$$

Como f é um polinômio homogêneo, podemos supor que x, y, z é *primitivo* (isto é, todas as entradas são em $\mathbb{Z}_p$ e uma delas é em $\mathbb{Z}_p^*$, em outras palavras, indivisível por p). Como $b \equiv 0 \bmod p$, obtemos por $(*)$ que $x^2 - ay^2 \equiv 0 \bmod p$.

Se $y \equiv 0 \bmod p$, então também $x \equiv 0 \bmod p$, logo bz^2 seria divisível por p^2. Como $v_p(b) = 1$, obteríamos $z \equiv 0 \bmod p$. Isto contradiria a primitividade de x, y, z.

Logo $y \not\equiv 0 \bmod p$, logo a é um quadrado módulo p. Como pelo Teorema Chinês dos Restos (Teorema B.5) $\mathbb{Z}/b\mathbb{Z} = \mathbb{Z}/p_1\mathbb{Z} \times \cdots \times \mathbb{Z}/p_1\mathbb{Z}$, segui que a é um quadrado módulo b.

Isto é, existem inteiros t e b' tal que

$$a = t^2 + b'b. \tag{**}$$

Podemos escolher t tal que $|t| \leq |b|/2$. Denote $\mathbf{K} = \mathbb{Q}$ ou $\mathbb{Q}_p$ ou $\mathbb{R}$. A equação $(**)$ diz que $b'b$ é uma *norma* da extensão $\mathbf{K}(\sqrt{-a})$; a equação $(*)$ diz que b é uma *norma* da extensão $\mathbf{K}(\sqrt{-a})$. Diante de Proposição E.3 e da multiplicidade do grupo de normas, concluímos que f representa zero em $\mathbf{K}$ se, e somente se,

$$f' = \mathrm{X}^2 - a\mathrm{Y}^2 - b'\mathrm{Z}^2$$

representa zero. Em particular, f' representa zero em cada $\mathbb{Q}_p$ e $\mathbb{R}$.

Vale, pela desigualdade triangular,

$$|b'| = \left|\frac{t^2 - a}{b}\right| \leq |b|/4 + 1 < |b|$$

pois $|b| \geq 2$ pela escolha de t. Escreve $b' = b''u^2$ com b'' e u inteiros e b'' livre de quadrados. Vale $|b''| < |b|$.

Assim a hipótese de indução aplica-se à forma $f'' = \mathrm{X}^2 - a\mathrm{Y}^2 - b''\mathrm{Z}^2$ que é equivalente a f'. Logo esta forma representa zero em $\mathbb{Q}$, e o mesmo vale para f.

(iii) O caso $n = 4$:

É reduzido ao caso $n = 3$.

(iv) O caso $n \geq 5$:

É provado por indução sobre n.

E.4. Existência de zeros para muitas variáveis

Teorema E.5 (de Chevalley-Warning). *Seja* $\mathrm{P}(\mathrm{X}_1, \ldots, \mathrm{X}_d)$ *um polinómio sobre um corpo finito* $\mathbb{F}_q$ *de característica* $p > 0$. *Se* $d > \deg \mathrm{P}$, *então* p *divide o número de zeros de* P. *Em particular, (o Teorema de Chevalley), se* 0 *é um zero de* P, *então existe outro.*

Demonstração: Como (com a convenção $0^0 = 1$) para qualquer expoente $i < q-1$

$$\sum_{x \in \mathbb{F}} x^i = 0,$$

vale para todo polinômio $\mathrm{P}(x_1, ..., x_d)$ de grau $\deg \mathrm{P} < d(q-1)$

$$\sum_{(x_1,\ldots,x_n) \in \mathbb{F}_q^n} \mathrm{P}(x_1, \ldots, x_n) = 0. \qquad (*)$$

Com efeito, basta por linearidade verificar isto para todos os monômios.

Como $\#\mathbb{F}_q^* = q - 1$, a função

$$\chi = 1 - \mathrm{P}^{q-1}$$

é a função característica dos zeros de P, isto é, tem o valor 1 em cada zero e 0 além. Se $d > \deg \mathrm{P}$, então

$$\deg \chi = \deg \mathrm{P}(q-1) < d(q-1).$$

Logo, por Equação $(*)$,

$$\sum_{(x_1,\ldots,x_n) \in \mathrm{F}^n} \chi(x_1, \ldots, x_n) = \#\{\text{zeros de P}\} = 0$$

em $\mathbb{F}_q$. Em particular, p divide o número de zeros de P. □

Lema E.6 (de Hensel para múltiplas varíaveis). *Seja f em $\mathcal{O}_{\mathbf{K}}[\mathrm{X}_1, \ldots, \mathrm{X}_d]$. Se existe um x_0 em $\mathcal{O}_{\mathbf{K}} \times \cdots \times \mathcal{O}_{\mathbf{K}}$ tal que*

$$f(x_0) \equiv 0 \bmod \pi^{2k+1} \quad e \quad \partial f / \partial \mathrm{X}_i(x_0) \not\equiv 0 \bmod \pi^{k+1} \text{ para algum } i = 1, \ldots, d,$$

então existe um único x em $\mathcal{O}_{\mathbf{K}} \times \cdots \times \mathcal{O}_{\mathbf{K}}$ tal que $f(x) = 0$, e vale $x \equiv x_0 \bmod \pi^{k+1}$.

Demonstração: Suponhamos que $i = 1$. Considera $\mathrm{F}(\mathrm{X}) = f(\mathrm{X}, x_{0,2}, \ldots, x_{0,d})$ e seja $y_0 \coloneqq x_{0,1}$. Vale $\mathrm{F}'(y_0) = \partial f / \partial \mathrm{X}_1(x_0)$; pelo Lema de Hensel, existe um único y tal que $f(y) = 0$ e $y \equiv y_0 \bmod \pi^{k+1}$. Logo $x = (y, x_{0,2}, \ldots, x_{0,d})$ satisfaz as condições desejadas. □

Proposição E.7. *Seja $p > 2$.*

- *Seja* F *uma forma quadrática regular sobre $\mathbb{Q}_p$ de $d > 2$ incógnitas. Se* F *tem forma diagonal cujos coeficientes são unidades em $\mathbb{Z}_p$, então* F *é isótropo, isto é, existe um zero não-trivial.*

- *Toda forma quadrática de $d > 4$ incógnitas sobre $\mathbb{Q}_p$ é isótropa.*

Demonstração:

- Pelo Teorema de Chevalley, existe um zero x sobre $\mathbb{F}_p$ para qualquer p. Seja $x_1 \neq 0$. Vale $\partial \mathrm{F}/\partial \mathrm{X}_1(x) = 2x_1 \not\equiv 0 \bmod p$ (porque $p > 2$). Pelo Lema de Hensel para múltiplas variáveis, este zero pode ser levado a um zero sobre $\mathbb{Z}_p$.

- Podemos supor que $n = 5$ e que $\mathrm{F} = a_1\mathrm{X}_1^2 + \cdots + a_5\mathrm{X}_5^2$ tenha forma diagonal. Pela escolha de uma base, podemos supor que $v_p(a_1), \ldots, v_p(a_5)$ em $\{0,1\}$. Logo $\mathrm{F} = \mathrm{F}' + p\mathrm{F}''$ onde os coeficientes de F' e F'' são unidades e, ou F', ou F'', tem > 2 incógnitas. Logo, uma delas é isótropa pela primeira proposição. □

A primeira parte da proposição é errada para $p = 2$, porém a segunda vale sem esta condição.

Corolário E.8. *Seja* F *uma forma quadrática sobre $\mathbb{Z}$ de $d > 4$ incógnitas (e $p > 2$). Se* F *tem um zero não-trivial sobre $\mathbb{R}$, então* F *tem um zero não-trivial sobre $\mathbb{Z}$.*

Demonstração: Por Proposição E.7, existe um zero não-trivial sobre $\mathbb{Q}_p$ para todo primo p. Pela hipótese, também sobre $\mathbb{R}$. Pelo Teorema de Hasse-Minkowski, existe um zero sobre $\mathbb{Q}$, ou, equivalentemente, por ser um polinômio homogêneo, sobre $\mathbb{Z}$. □

Referências

[GS92] M. Gromov and R. Schoen, *Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one*, Inst. Hautes Études Sci. Publ. Math. (1992), no. 76, 165–246. MR 1215595. Confer http://www.numdam.org/item?id=PMIHES_1992__76__165_0.

[Kob84] N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, second ed., Graduate Texts in Mathematics, vol. 58, Springer-Verlag, New York, 1984. MR 754003. DOI 10.1007/978-1-4612-1112-9.

[Nag11] E. Nagel, *Fractional non-Archimedean differentiability*, Univ. Münster, Mathematisch-Naturwissenschaftliche Fakultät (Diss.), 2011. zbMATH 1223.26011. Confer http://nbn-resolving.de/urn:nbn:de:hbz:6-75409405856.

[Nag12] _____, *Fractional non-Archimedean calculus in one variable*, p-Adic Numbers Ultrametric Anal. Appl. **4** (2012), no. 4, 271–305. MR 2992413. DOI 10.1134/S2070046612040036.

[Nag16] _____, *p-adic Taylor polynomials*, Indag. Math. (N.S.) **27** (2016), no. 3, 643–669. MR 3505986. DOI 10.1016/j.indag.2015.12.003. Confer http://www.sciencedirect.com/science/article/pii/S0019357715001275.

[PGS10] C. Perez-Garcia and W. H. Schikhof, *Locally convex spaces over non-Archimedean valued fields*, Cambridge Studies in Advanced Mathematics, vol. 119, Cambridge University Press, Cambridge, 2010. MR 2598517. DOI 10.1017/CBO9780511729959.

[RL03] J. Rivera-Letelier, *Espace hyperbolique p-adique et dynamique des fonctions rationnelles*, Compositio Math. **138** (2003), no. 2, 199–231. MR 2018827. DOI 10.1023/A:1026136530383.

[RRS11] J. Ripoll, C. Ripoll, and J. Silveira, *Números racionais, reais e complexos*, Porto Alegre: UFRGS, 2011.

[Sch84] W. H. Schikhof, *Ultrametric calculus*, Cambridge Studies in Advanced Mathematics, vol. 4, Cambridge University Press, Cambridge, 1984, An introduction to p-adic analysis. MR 791759.

[Sch95] ———, *A perfect duality between p-adic Banach spaces and compactoids*, Indag. Math. (N.S.) **6** (1995), no. 3, 325–339. MR 1351151. DOI 10.1016/0019-3577(95)93200-T.

[Sch99] P. Schneider, *p-adic Representation Theory*, November 1999, Britton Lectures at McMaster University. Confer http://wwwmath.uni-muenster.de/u/pschnei/publ/lectnotes/hamilton.dvi.

[Sch02] ———, *Nonarchimedean functional analysis*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2002. MR 1869547. DOI 10.1007/978-3-662-04728-6.

[Sch03] W. H. Schikhof, *A crash course in p-adic analysis*, 2003. Confer http://pucuch.mat.uc.cl/docume/100304120442.pdf.

[Ser70] J.-P. Serre, *Cours d'arithmétique*.

[ST02] P. Schneider and J. Teitelbaum, *Banach space representations and Iwasawa theory*, Israel J. Math. **127** (2002), 359–380. MR 1900706. DOI 10.1007/BF02784538.