Em busca da unidade: conexões de Galois e inversões de Möbius-Rota[1]

Jorge Picado
CMUC, DMat
Universidade de Coimbra
e-mail: picado@mat.uc.pt

Pedro M. Silva
Dep. de Física
Universidade de Coimbra
e-mail: pmsilva@student.fisica.uc.pt

Resumo: Este artigo está organizado em duas partes distintas, desenvolvidas em paralelo, onde ilustramos a utilidade das conexões de Galois na teoria dos números (primeira parte) e das inversões de Möbius-Rota na combinatória (segunda parte). Estas ferramentas permitem abordar problemas aparentemente difíceis, transformando-os, com um simples processo de inversão, em problemas equivalentes mais simples. O fio condutor comum é a visão conceptual da teoria dos reticulados.

Abstract: This paper is organized in two distinct but parallel parts. Our goal is to illustrate the parallel between the usefulness of Galois connections (quasi-inversions) in number theory and of Möbius-Rota inversions in enumerative combinatorics. These tools allow to address apparently hard problems in an illuminating unifying way, by (quasi-)inverting them into much simpler equivalent problems. The common setting is the conceptual point of view of lattice theory.

palavras-chave: Sequências complementares, desencontros, princípio da inclusão-exclusão, conexão de Galois, inversão de Möbius, álgebra de incidência de Rota.

keywords: Complementary sequences, derangements, inclusion-exclusion principle, Galois connection, Möbius inversion, Rota incidence algebra.

> «At a time when mathematical fashion despises generality (seen as gratuitous "generalities", i.e. vacuities) I affirm the principal force in all my work has been the quest for the "general." In truth I prefer to accent "unity" rather than "generality." But for me these are two aspects of one quest. Unity represents the profound aspect, and generality the superficial aspect.»
>
> — Alexandre Grothendieck
> (Récoltes et Semailles, 1986)

[1] Trabalho realizado no âmbito do programa *Novos Talentos em Matemática* da Fundação Calouste Gulbenkian.

1 Um problema elementar de números

«Quem? O infinito?
Diz-lhe que entre.
Faz bem ao infinito
estar entre gente.»

— ALEXANDRE O'NEILL
(De Porta em Porta, 1960)

Denotemos por $\mathbb{N}_0$ o conjunto dos números naturais (onde incluímos o número 0) e consideremos duas funções bem conhecidas da teoria dos números: seja $f(n)$ o n-ésimo número primo e seja $g(n)$ o número de números primos que não excedem n. Será conveniente considerarmos a sucessão f dos números primos a começar em $n = 0$: assumimos que $f(0) = 0$. Claro que $g(0) = 0$. Calculemos agora os primeiros valores das sucessões definidas pelas somas $f(n) + n$ e $g(n) + n + 1$ ($n \in \mathbb{N}_0$):

n	0	1	2	3	4	5	6	7	8	9	10	$\cdots$
$f(n)$	0	2	3	5	7	11	13	17	19	23	29	
$f(n)+n$	0	3	5	8	11	16	19	24	27	32	39	
$g(n)$	0	0	1	2	2	3	3	4	4	4	4	
$g(n)+n+1$	1	2	4	6	7	9	10	12	13	14	15	

Observemos melhor os números nessas duas linhas:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, \ldots$$

Surpreendente, não? Será que o 17 e o 18 aparecerão em seguida na linha correspondente a $g(n) + n + 1$? (o 16 e o 19 já estão na linha de $f(n) + n$.) Isto é, será que estas sequências infinitas de naturais são *complementares*, ou seja, não têm elementos comuns e em conjunto esgotam todos os naturais? Desafiamos o caro leitor a tentar demonstrar tal facto. Não parece ser um exercício fácil, pois não? De facto, está longe de ser óbvio!

E mais: este facto não tem nada a ver com as propriedades dos números primos! Por exemplo, se $f(n)$ for agora o n-ésimo quadrado perfeito (convencionamos mais uma vez $f(0) = 0$) e $g(n)$ for o número de quadrados perfeitos que não excedem n, obtemos a tabela

n	0	1	2	3	4	5	6	7	8	9	10	$\cdots$
$f(n)$	0	0	1	4	9	16	25	36	49	64	81	
$f(n)+n$	0	1	3	7	13	21	31	43	57	73	91	
$g(n)$	1	2	2	2	3	3	3	3	3	4	4	
$g(n)+n+1$	2	4	5	6	8	9	10	11	12	14	15	

E poderíamos continuar com outros exemplos. Mais geralmente, para uma qualquer propriedade P, se $f(n)$ é o n-ésimo número P, $f\colon \mathbb{N}_0 \to \mathbb{N}_0$ uma sucessão crescente, com $f(0)=0$, tendendo para ∞ com n, e $g(n)$ é o número de números P que não excedem n, o problema resume-se a:

Problema 1. *Será verdade que as sequências*

$$F=\{f(n)+n \mid n\in\mathbb{N}_0\} \quad e \quad G=\{g(n)+n+1 \mid n\in\mathbb{N}_0\}$$

são complementares?

Este problema foi originalmente resolvido por J. Lambek e L. Moser em 1954 [10], tendo resposta positiva, como veremos mais adiante. O padrão comum a este e outros problemas análogos, envolvendo funções bem conhecidas da aritmética e da teoria dos números, são as *conexões de Galois* [9]. Isso mesmo, o tipo de conexão entre os corpos intermédios de uma extensão de corpos e os subgrupos do correspondente grupo de automorfismos, base da teoria de Galois moderna reformulada por E. Artin.

Este artigo é constituído por duas partes aparentemente distintas.

Na primeira parte, abrangendo as primeiras cinco secções, veremos como as conexões de Galois [6, 8], combinando grande clareza estrutural e facilidade computacional, resolvem o Problema 1 de uma maneira surpreendentemente simples e elegante. Observaremos como, olhadas como uma generalização de pares de bijecções mutuamente inversas (mais especificamente, pares de funções *quase inversas*), permitem transferir informação do lado onde esta é mais completa para o lado oposto. Aproveitaremos ainda para mostrar a sua utilidade no ensino dos fundamentos da teoria dos conjuntos na possível unificação e sistematização de muitas propriedades .

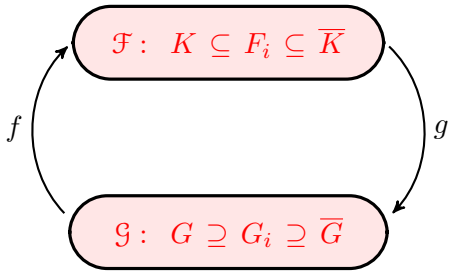
Na segunda parte do artigo (secções 6-10) passamos dos números para a combinatória. O objectivo é apresentar, numa abordagem em tudo paralela à primeira parte, uma ferramenta da combinatória enumerativa que desempenha um papel análogo ao das quase-inversões de Galois descrito na primeira parte: as *inversões de Möbius-Rota* [16].

2 Conexões de Galois

«Adjunctions arise everywhere»

— SAUNDERS MAC LANE
(Categories for the Working Mathematician, 1971)

A famosa solução do problema da resolubilidade algébrica de uma equação polinomial baseia-se na correspondência, descoberta por Galois (1811-1832) em 1830, para uma dada extensão de corpos $K \subseteq \overline{K}$, entre a colecção $\mathcal{F}$ dos subcorpos de $\overline{K}$ contendo K e a colecção $\mathcal{G}$ de todos os subgrupos do grupo de automorfismos de $\overline{K}$ que deixam K invariante (o chamado *grupo de Galois* dessa extensão):



A moderna *teoria de Galois* baseia-se nesta correspondência e no facto essencial de que as funções f e g são, em geral, *quase inversas* uma da outra:

$$\forall G \in \mathcal{G},\ \forall F \in \mathcal{F}\ (f(G) \subseteq F \ \Leftrightarrow\ G \supseteq g(F)).$$

Esta correspondência é um exemplo daquilo a que hoje se chama uma conexão de Galois (dual ou contravariante): Birkhoff, em 1940, associou uma conexão deste tipo a qualquer relação binária (a que chamou *polaridade* [2]) e, em 1944, Ore [14] generalizou o conceito a quaisquer conjuntos parcialmente ordenados. Invertendo a ordem num dos lados chegamos à definição (covariante) moderna de conexão de Galois [2, p. 124]:

Definição 2.1. Sejam $(A, \leq)$ e $(B, \leq)$ dois conjuntos parcialmente ordenados. Um par de funções

$$(A, \leq) \underset{g}{\overset{f}{\rightleftarrows}} (B, \leq)$$

diz-se uma *conexão* (ou *adjunção*[2]) *de Galois* quando

$$\forall a \in A,\ \forall b \in B\ (f(a) \leq b \ \Leftrightarrow\ a \leq g(b)).$$

A função f chama-se o *adjunto à esquerda* enquanto g é o *adjunto à direita*. Abrevia-se tudo escrevendo simplesmente $f \dashv g$.

Note que, no caso em que as relações de ordem em A e B são a relação de igualdade, $f \dashv g$ significa simplesmente que f e g são um par de bijeccções mutuamente inversas.

É agora claro que em todos os exemplos da secção anterior temos precisamente uma conexão de Galois $f \dashv g$ em $(\mathbb{N}_0, \leq)$.

Note que um adjunto à esquerda (resp. à direita) de uma dada função monótona pode não existir, mas caso exista é necessariamente único: se $f_i(a) \leq b \Leftrightarrow a \leq g(b)$, $i = 1, 2$, então, evidentemente, $f_1(a) \leq b \Leftrightarrow f_2(a) \leq b$.

Proposição 2.2. *Sejam $f\colon A \to B$ e $g\colon B \to A$.*

(1) *Se $f \dashv g$ então:*

- (i) *f e g são funções quase-inversas, isto é, $fg \leq \mathrm{id}$ e $\mathrm{id} \leq gf$.*
- (ii) *f e g são monótonas.*[3]
- (iii) *$fgf = f$ e $gfg = g$.*
- (iv) *f e g definem bijecções entre $f[A]$ e $g[B]$, inversas uma da outra.*
- (v) *f preserva supremos*[4]*, g preserva ínfimos,*

$$f(a) = \inf\{b \in B \mid a \leq g(b)\} \quad e \quad g(b) = \sup\{a \in A \mid f(a) \leq b\}.$$

(2) *Reciprocamente, se f e g são monótonas, então:*

- (i) *Se, para cada $b \in B$, existe o supremo de $\{a \in A \mid f(a) \leq b\}$ em A, e f preserva estes supremos, então f possui um (único) adjunto à direita dado por esta fórmula.*

[2] Terminologia influenciada pela teoria das categorias, uma vez que este conceito é um exemplo do conceito mais geral de *adjunção* entre duas categorias, aplicado a conjuntos parcialmente ordenados (considerados como categorias magras, de modo *standard*). Segundo este ponto de vista, toda a teoria de reticulados pode ser vista como teoria das categorias em categorias magras, precisamente $(0,1)$-categorias (ver [*nLab HomePage*, `ncatlab.org/nlab/show/partial+order`]).

[3] As condições (1:i) e (1:ii), em conjunto, caracterizam a adjunção $f \dashv g$: se $f(a) \leq b$ então $gf(a) \leq g(b)$ pelo que $a \leq gf(a) \leq g(b)$; analogamente, se $a \leq g(b)$, então $f(a) \leq fg(b) \leq b$.

[4] Dizemos que *f preserva supremos* quando preserva todos os supremos que existem em A; analogamente para os ínfimos.

(ii) *Se, para cada $a \in A$, existe o ínfimo de $\{b \in B \mid a \leq g(b)\}$ em B, e g preserva estes ínfimos, então g possui um (único) adjunto à esquerda dado por esta fórmula.*

Demonstração. (1) (i) Como $f(a) \leq f(a)$, então $a \leq gf(a)$ para qualquer $a \in A$. Analogamente, $fg(b) \leq b$ para qualquer $b \in B$.
(ii) Se $a \leq a'$ em A, então $a \leq gf(a')$, isto é, $f(a) \leq f(a')$. Analogamente para g.
(iii) A desigualdade $fgf \leq f$ decorre imediatamente da desigualdade $fg \leq \mathrm{id}$, enquanto $\mathrm{id} \leq gf$ e o facto de f ser monótona implicam $f \leq fgf$. De modo análogo, $gfg = g$.
(iv) É consequência imediata da alínea anterior.
(v) Sejam $S \subseteq A$, $x = \sup S$. Temos que mostrar que $f(x)$ é o supremo de $\{f(s) \mid s \in S\}$ em B:

- $f(x) \geq f(s)$ para qualquer $s \in S$ pois f é monótona.
- Se algum $b \in B$ satisfaz $b \geq f(s)$ para todo o $s \in S$ então, pela adjunção, $g(b) \geq s$ para todo o $s \in S$ e, portanto, $g(b) \geq x$. Pela adjunção, isto significa que $b \geq f(x)$.

De modo análogo, pode provar-se que g preserva ínfimos.

Finalmente, mostremos que $g(b) = \sup\{a \in A \mid f(a) \leq b\}$ (o resultado dual para f segue de modo semelhante):

- Denotemos o conjunto $\{a \in A \mid f(a) \leq b\}$ por S. Claramente $g(b) \in S$ pois $fg(b) \leq b$.
- Por outro lado, $g(b) \geq s$ para qualquer $s \in S$, uma vez que, por definição de S, $b \geq f(s)$ para qualquer $s \in S$.

(2) (i) Consideremos a função $h\colon B \to A$ definida por

$$h(b) = \sup\{a \in A \mid f(a) \leq b\}.$$

Trata-se de um adjunto à direita de f: se $f(a) \leq b$, evidentemente $a \leq h(b)$; reciprocamente, se $a \leq h(b)$, então, como f é monótona e preserva supremos, $f(a) \leq \sup\{f(a') \mid a' \in A, f(a') \leq b\} \leq b$.
(ii) Por dualidade. □

Corolário 2.3. *Sejam $f\colon A \to B$ e $g\colon B \to A$ funções monótonas.*

(1) *f é um adjunto à esquerda se e só se preserva supremos e para cada $b \in B$ existe o supremo de $\{a \in A \mid f(a) \leq b\}$ em A.*

(2) *g é um adjunto à direita se e só se preserva ínfimos e para cada $a \in A$ existe o ínfimo de $\{b \in B \mid a \leq g(b)\}$ em B.* □

Portanto, quando A e B são reticulados completos,

(1) f é um adjunto à esquerda se e só se preserva supremos, e

(2) g é um adjunto à direita se e só se preserva ínfimos.

3 Resolvendo o Problema com conexões de Galois

> «The structural contribution of Galois was not so much to do with fields and groups but with their relationship.»
> — Øystein Ore
> (Galois connexions, 1944)

Analisemos agora a situação que nos interessa, do Problema 1:

$$A = B = (\mathbb{N}_0, \leq), \text{ com a ordem usual } \leq .$$

Trata-se de um conjunto totalmente ordenado, onde todo o subconjunto não vazio tem um ínfimo – o seu mínimo –, mas só os subconjuntos *finitos* têm supremo – o máximo do conjunto –; em particular, o máximo do conjunto vazio é o zero.

Lema 3.1. *Seja $f\colon \mathbb{N}_0 \to \mathbb{N}_0$. Então:*

(1) *$f[\mathbb{N}_0]$ é um conjunto infinito se e só se $\{m \in \mathbb{N}_0 \mid f(m) \geq n\} \neq \emptyset$ para qualquer $n \in \mathbb{N}_0$.*

(2) *Se f é monótona então as seguintes afirmações também são equivalentes:*

   (i) *$f[\mathbb{N}_0]$ é um conjunto infinito.*

   (ii) *$f(n) \to \infty$ quando $n \to \infty$ (isto é, $\forall m \in \mathbb{N}_0\ \exists N \in \mathbb{N}_0\colon f(n) \geq m$ para qualquer $n \geq N$).*

   (iii) *$\{n \in \mathbb{N}_0 \mid f(n) \leq m\}$ é finito para qualquer $m \in \mathbb{N}_0$.*

Demonstração. (1) Suponhamos que $f[\mathbb{N}_0]$ é infinito. O caso $n = 0$ é óbvio: $\{m \in \mathbb{N}_0 \mid f(m) \geq 0\} = \mathbb{N}_0$. Os restantes casos ($n \in \mathbb{N}$) também são evidentes: se $\{m \in \mathbb{N}_0 \mid f(m) \geq n\}$ fosse vazio, teríamos $f[\mathbb{N}_0] \subseteq [0, n-1]$, um absurdo.

Reciprocamente, se $f[\mathbb{N}_0]$ fosse finito, igual a, digamos, $\{y_1 < y_2 < \cdots < y_k\} \subseteq \mathbb{N}_0$, para $n = y_k + 1$ existiria, por hipótese, $m \in \mathbb{N}_0$ tal que $f(m) \geq n > y_k$, pelo que $f(m)$ não pertenceria a $f[\mathbb{N}_0]$, um absurdo.

(2) (i)⇒(ii): Para cada $m \in \mathbb{N}_0$, como $f[\mathbb{N}_0]$ é infinito existe N tal que $f(N) \geq m$. Pela monotonia de f, $f(n) \geq f(N) \geq m$ para qualquer $n \geq N$. A implicação recíproca (ii)⇒(i) é óbvia.

(ii)⇒(iii) Seja $m \in \mathbb{N}_0$. Se $\{n \in \mathbb{N}_0 \mid f(n) \leq m\}$ fosse infinito, teríamos uma sucessão

$$n_1 < n_2 < \cdots < n_k < \cdots \quad (k \in \mathbb{N})$$

tal que $f(n_i) \leq m$. Como f é monótona, isto implicaria $f(n) \leq m$ para todo o $n \in \mathbb{N}_0$, o que contraria a hipótese.

(iii)⇒(ii) Seja $m \in \mathbb{N}_0$. Como $A_m := \{n \in \mathbb{N}_0 \mid f(n) \leq m\}$ é finito então existe $N \in \mathbb{N}_0$ tal que $N \notin A_m$, ou seja, $f(N) > m$. □

O Corolário 2.3 reduz-se agora a:

Corolário 3.2. *Sejam $f, g\colon \mathbb{N}_0 \to \mathbb{N}_0$ funções monótonas.*

(1) *f é um adjunto à esquerda se e só se $f(0) = 0$ e $f[\mathbb{N}_0]$ é infinito. Nesse caso, o seu adjunto direito é dado pela fórmula $g(m) = \max\{n \in \mathbb{N}_0 \mid f(n) \leq m\}$.*

(2) *g é um adjunto à direita se e só se $g[\mathbb{N}_0]$ é infinito. Nesse caso, o seu adjunto direito é dado pela fórmula $f(n) = \min\{m \in \mathbb{N}_0 \mid n \leq g(m)\}$.*

Demonstração. (1) Como $(\mathbb{N}_0, \leq)$ é totalmente ordenado, qualquer função monótona $f\colon \mathbb{N}_0 \to \mathbb{N}_0$ preserva supremos binários. Portanto, preserva todos os supremos que existem em $\mathbb{N}_0$, isto é, os supremos finitos, se e só se $f(0) = 0$. Por outro lado, para cada $m \in \mathbb{N}_0$, o conjunto $\{n \in \mathbb{N}_0 \mid f(n) \leq m\}$ tem supremo se e só se é finito. Pelo Lema 3.1(2), a finitude daquele conjunto é equivalente à condição '$f[\mathbb{N}_0]$ é infinito'.

(2) Em primeiro lugar, qualquer função monótona $g\colon \mathbb{N}_0 \to \mathbb{N}_0$ preserva ínfimos de subconjuntos não vazios e, portanto, preserva todos os ínfimos que existem em $\mathbb{N}_0$. Além disso, para cada $n \in \mathbb{N}_0$, o conjunto $\{m \in \mathbb{N}_0 \mid n \leq g(m)\}$ tem ínfimo se e só se não é vazio, ou seja, se e só se a imagem $g[\mathbb{N}_0]$ é infinita (Lema 3.1(1)). □

Daqui decorre, sem dificuldade, o resultado de Lambek-Moser [10, 9] que confirma que o Problema 1 (pág. 3) tem, de facto, resposta positiva. O

Teorema de Lambek-Moser[5] é universal, no sentido em que descreve qualquer partição dos naturais em dois subconjuntos infinitos, em termos de conexões de Galois em $(\mathbb{N}_0, \leq)$:

Teorema 3.3. (1) *Sejam F e G subconjuntos infinitos complementares de $\mathbb{N}_0$ com $0 \in F$. Sendo $F(n)$ o $(n+1)$-ésimo elemento de F e $G(m)$ o $(m+1)$-ésimo elemento de G, as funções $f, g\colon \mathbb{N}_0 \to \mathbb{N}_0$ definidas por*

$$f(n) = F(n) - n, \quad g(m) = G(m) - m - 1$$

determinam uma conexão de Galois $f \dashv g$.

(2) *Seja $f \dashv g$ uma conexão de Galois entre $(\mathbb{N}_0, \leq)$ e ele próprio. Os conjuntos*

$$F = \{f(n) + n \mid n \in \mathbb{N}_0\} \quad e \quad G = \{g(n) + n + 1 \mid n \in \mathbb{N}_0\}$$

são infinitos e formam uma partição de $\mathbb{N}_0$ (com $0 \in F$).

Demonstração. (1) Teremos que mostrar que $f(n) \leq m$ se e só se $n \leq g(m)$, isto é,

$$F(n) \leq n + m \quad \Leftrightarrow \quad n + m + 1 \leq G(m).$$

'$\Rightarrow$': Por absurdo: Se $F(n) \leq n + m$ e $G(m) \leq n + m$, então $F(0), \ldots, F(n)$ e $G(0), \ldots, G(m)$ seriam $m + n + 2$ números naturais distintos $\leq m + n$.

'$\Leftarrow$': Por absurdo: Se $G(m) \not\leq n + m$ e $F(n) \not\leq n + m$, então F teria no máximo n elementos $\leq n+m$ e G teria no máximo m elementos $\leq n+m$, pelo que a sua união teria no máximo $n + m$ elementos $\leq n + m$, contradizendo a existência de $n + m + 1$ naturais $\leq n + m$.

(2) O Corolário 3.2 garante que o subconjunto F é infinito e contém o número 0. O complementar $G' = \mathbb{N}_0 \setminus F$ também é infinito: se não fosse, existiria $n_0 \in \mathbb{N}_0$ tal que $n \in F$ para qualquer $n \geq n_0$, pelo que teríamos $F(n) = F(n_0) + (n - n_0)$, isto é, $f(n) + n = f(n_0) + n_0 + n - n_0$, ou seja, $f(n) = f(n_0)$ para qualquer $n \geq n_0$, contradizendo o facto de que $f[\mathbb{N}_0]$ é infinito (Corolário 3.2).

Assim, por (1), f teria um adjunto à direita g' definido por

$$g'(m) = G'(m) - m - 1$$

onde $G'(m)$ é o $(m+1)$-ésimo elemento de G'. Mas, como vimos na secção anterior, os adjuntos são únicos, logo $g' = g$ e, consequentemente,

$$G'(m) = g(m) + m + 1 = G(m).$$

Portanto, G é o complemento (infinito) de F em $\mathbb{N}_0$. □

[5] O resultado análogo a este para o conjunto parcialmente ordenado $(\mathbb{Z}, \leq)$ encontra-se em [9].

4 Ilustrando as potencialidades do método

> «In their most general sense, adjunctions and/or Galois connections make it possible to relate two "worlds" of (more or less mathematical) objects with each other in order to gain information about one world by passing to the other, perhaps better known world.»
>
> — MARCEL ERNÉ
> (Capítulo 1 de [6], 2004)

Exemplos 4.1. (1) A vantagem que podemos tirar de uma conexão de Galois[6], tal como o próprio Galois fez, é concluir factos novos num dos lados da correspondência (no problema de Galois, no lado das extensões de corpos) a partir de factos mais facilmente prováveis no lado oposto. Por exemplo, a fórmula de cálculo para o n-ésimo termo da sucessão em $\mathbb{N}_0$ dos *quadrados perfeitos* é óbvia: $(n-1)^2$. Mas já não parece tão fácil determinar uma fórmula geral para o n-ésimo número que não é quadrado perfeito. Vejamos como com uma simples conexão de Galois podemos descobrir esta fórmula a partir da primeira.

Seja F a sucessão dos quadrados perfeitos em $\mathbb{N}_0$ e G o respectivo complemento:

n	0	1	2	3	4	5	6	7	8	9	10	$\cdots$
$F\colon f(n)+n$	0	1	4	9	16	25	36	49	64	81	100	
$f(n)$	0	0	2	6	12	20	30	42	56	72	90	
$G\colon g(n)+n+1$	2	3	5	6	7	8	10	11	12	13	14	
$g(n)$	1	1	2	2	2	2	3	3	3	3	3	

O problema resume-se à determinação de uma fórmula para $G(n-1) = g(n-1)+n$. Para isso basta aplicar o Teorema 3.3:

Primeiro,

$$f(n) = n^2 - n, \quad g(n) = G(n) - n - 1.$$

Da adjunção $f \dashv g$ sabemos então que

$$g(n) = \max\{m \in \mathbb{N}_0 \mid f(m) \leq n\} = \max\{m \in \mathbb{N}_0 \mid m^2 - m \leq n\}$$

[6] Para mais informação sobre conexões de Galois consulte, por exemplo, [6, 8, 15].

e, como

$$m^2 - m \leq n \Leftrightarrow m^2 - m + \frac{1}{4} < n + 1$$
$$\Leftrightarrow \left(m - \frac{1}{2}\right)^2 < n + 1 \Leftrightarrow m - \frac{1}{2} < \sqrt{n+1},$$

então $g(n) = \lfloor \sqrt{n+1} + \frac{1}{2} \rfloor = [\sqrt{n+1}]$, onde $\lfloor x \rfloor$ denota a parte inteira do real x (mais adiante usaremos também a notação $\lceil x \rceil$ para referir o menor inteiro que não é menor que x) e $[x]$ designa o inteiro mais próximo de x (onde, no caso de x ser a metade de um inteiro ímpar y, escolhemos $[x] = \frac{y-1}{2}$). Concluindo,

$$G(n-1) = n + [\sqrt{n}]$$

é o n-ésimo número em $\mathbb{N}_0$ que não é quadrado perfeito.

(2) Sejam p e q números irracionais positivos tais que $\frac{1}{p} + \frac{1}{q} = 1$. A função f definida por $f(n) = \lfloor (p-1)n \rfloor$ é um adjunto à esquerda da função g dada por $g(m) = \lfloor (q-1)(m+1) \rfloor$. Portanto,

$$F(n) = f(n) + n = \lfloor pn \rfloor \text{ e } G(m) = g(m) + m + 1 = \lfloor q(m+1) \rfloor,\ n, m \in \mathbb{N}_0,$$

enumeram uma partição de $\mathbb{N}_0$ em subconjuntos infinitos. Esquecendo $F(0) = 0$, temos então que as sequências

$$A = \{\lfloor p \rfloor, \lfloor 2p \rfloor, \lfloor 3p \rfloor, \ldots\} \quad \text{e} \quad B = \{\lfloor q \rfloor, \lfloor 2q \rfloor, \lfloor 3q \rfloor, \ldots\}$$

constituem uma partição de $\mathbb{N}$. Trata-se de um resultado de Beatty [1] com 90 anos; por isso, as sequências em A e B são chamadas *sequências de Beatty*. Por exemplo, $p = \sqrt{2}$ gera a sequência de Beatty

$$A = \{1, 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 16, 18, 19, 21, \ldots\}.$$

A grande surpresa que o resultado de Beatty nos proporciona é que o complemento de A em $\mathbb{N}$,

$$\mathbb{N} \setminus A = \{3, 6, 10, 13, 17, 20, 23, 27, 30, 34, 37, 40, 44, \ldots\},$$

é também uma sequência de Beatty, gerada por $q = \frac{p}{p-1}$:

$$\mathbb{N} \setminus A = B = \left\{ \left\lfloor \frac{n\sqrt{2}}{\sqrt{2}-1} \right\rfloor : n \in \mathbb{N} \right\}.$$

Para mais exemplos elementares em $\mathbb{N}$ (ordenado pela relação de divisibilidade) e em $\mathbb{Z}$ (com a ordem usual) consulte [9].

5 Mais sobre conexões de Galois

> «Since the beginning of the century, computational procedures have become so complicated that any progress by those means has become impossible, without the elegance which modern mathematicians have brought to bear on their research, and by means of which the spirit comprehends quickly and in one step a great many computations.»
>
> — Évariste Galois
>
> (Do prefácio do seu último manuscrito, 1832)

Talvez valha a pena atentarmos em mais alguns exemplos elementares de conexões de Galois [8, 9, 15] que revelam algum do seu potencial o ensino dos fundamentos da teoria dos conjuntos.

Exemplos 5.1. (1) Suponhamos que uma função $f\colon \mathbb{N} \to \mathbb{N}$ pode ser estendida a uma função real crescente $\widetilde{f}$ no intervalo $\langle 1, +\infty)$ e seja ϕ a sua inversa. É evidente que $\lceil \phi(-) \rceil$ é um adjunto à esquerda de f e $\lfloor \phi(-) \rfloor$ é um adjunto à direita de f (consequência do facto óbvio de que $\lceil \phi(m) \rceil \leq n$ sse $\phi(m) \leq n$, e $\lfloor \phi(m) \rfloor \geq n$ sse $\phi(m) \geq n$).

Logo, por exemplo, $\lceil \log_2 \rceil$ e $\lfloor \log_2 \rfloor$ são os adjuntos à esquerda e à direita da exponencial $n \mapsto 2^n$.

(2) Sejam X e Y conjuntos arbitrários e $f\colon X \to Y$ uma função arbitrária. Uma vez que, para quaisquer $A \subseteq X$ e $B \subseteq Y$,

$$f[A] \subseteq B \text{ se e só se } A \subseteq f^{-1}[B], \tag{5.1.1}$$

as funções

$$f[-]\colon \mathcal{P}(X) \to \mathcal{P}(Y) \quad \text{e} \quad f^{-1}[-]\colon \mathcal{P}(Y) \to \mathcal{P}(X)$$

são adjuntas, $f[-]$ à esquerda e $f^{-1}[-]$ à direita. Isto significa que $f[-]$ preserva supremos enquanto $f^{-1}[-]$ preserva ínfimos. Daí as fórmulas básicas da teoria dos conjuntos

$$f[\bigcup_i A_i] = \bigcup_i f[A_i] \quad \text{e} \quad f^{-1}[\bigcap_i B_i] = \bigcap_i f^{-1}[B_i].$$

Compare agora a Proposição 2.2 com as fórmulas *standard* da teoria dos conjuntos

$$f[f^{-1}[B]] \subseteq B \quad \text{e} \quad A \subseteq f^{-1}[f[A]],$$
$$f[f^{-1}[f[A]]] = f[A] \quad \text{e} \quad f^{-1}[f[f^{-1}[B]]] = f^{-1}[B].$$

Mas $f^{-1}[-]$ também tem um adjunto à direita:

$$f^{-1}[B] \subseteq A \text{ se e só se } B \subseteq Y \smallsetminus f[X \smallsetminus A]. \tag{5.1.2}$$

(De facto, se $b \in B$ e, por absurdo, b pertencesse a $f[X \smallsetminus A]$, teríamos $b = f(a')$ para algum $a' \in X \smallsetminus A$, donde $a' \in f^{-1}[B] \subseteq A$, uma contradição; reciprocamente, se $x \in f^{-1}[B]$ então $f(x) \in B \subseteq Y \smallsetminus f[X \smallsetminus A]$, isto é, $f(x) \notin f[X \smallsetminus A]$, pelo que $x \in A$.)

Daqui decorrem as propriedades bem conhecidas da teoria dos conjuntos

$$f^{-1}[\bigcup_i B_i] = \bigcup_i f^{-1}[B_i], \quad B \subseteq Y \smallsetminus f[X \smallsetminus f^{-1}[B]],$$
$$f^{-1}[Y \smallsetminus f[X \smallsetminus A]] \subseteq A, \quad \text{etc.}$$

Em geral, no entanto, a função imagem $f[-]$ não tem nenhum adjunto à esquerda, uma vez que, ao contrário das pré-imagens, as imagens não preservam ínfimos.

De facto, se $f[-]$ preserva ínfimos, f é necessariamente uma bijecção: se não fosse injectiva, existiriam $a \neq b$ em X tais que $f(a) = f(b) = y$, o que implicaria $f[A \cap B] = f[\emptyset] = \emptyset$, para $A := \{a\}$ e $B := \{b\}$, enquanto $y \in f[A] \cap f[B]$; por outro lado, a sobrejectividade de f decorre simplesmente da preservação de ínfimos para famílias *vazias*: em $\mathcal{P}(X)$ (resp. $\mathcal{P}(Y)$) esse ínfimo é precisamente X (resp. Y), pelo que $f[X] = Y$.

Claro que no caso em que f é uma bijecção, com função inversa $g\colon Y \to X$, a equivalência (5.1.1) aplicada à inversa g diz-nos que $g[B] \subseteq A$ se e só se $B \subseteq g^{-1}[A]$, isto é, $f^{-1}[B] \subseteq A$ se e só se $B \subseteq f[A]$, ou seja, neste caso especial tem-se mesmo $f^{-1}[-] \dashv f[-]$.

(3) Sejam X, Y espaços topológicos e $f\colon X \to Y$ uma função contínua. A adjunção (5.1.2) pode ser modificada numa adjunção entre as duas topologias:

$$f^{-1}[B] \subseteq A \text{ se e só se } B \subseteq \operatorname{int}(Y \smallsetminus f[X \smallsetminus A]) \tag{5.1.3}$$

para quaisquer abertos A e B de X e Y, respectivamente. Contudo, a adjunção (5.1.1) não tem correspondente neste contexto, uma vez que, como é fácil de verificar, para funções contínuas e conjuntos abertos, a função imagem não preserva uniões enquanto a pré-imagem preserva uniões mas, em geral, não preserva ínfimos (note que, no caso infinito, estes não coincidem, em geral, com as intersecções mas sim com o interior das intersecções).

6 Um problema elementar de combinatória

> «God created infinity, and man, unable to understand infinity, had to invent finite sets.»

— Gian-Carlo Rota
('Combinatorics', em: Discrete Thoughts, 1969)

«A vida é a arte do encontro
Embora haja tanto desencontro pela vida»

— Vinícius de Moraes
(Samba da benção, 1967)

Como qualquer estudante rapidamente se apercebe, o estudo da combinatória enumerativa torna-se um desafio mais sério a partir do momento em que começamos a impor restrições nalgumas posições das configurações em análise. Por exemplo, no chamado 'jogo dos pares' (de casino):

Problema 2. *As 52 cartas de um baralho são dispostas sequencialmente, com o seu valor à vista. O 'croupier' dispõe então as cartas de um segundo baralho, uma a uma, por cima das primeiras. Ganha-se o jogo caso nenhuma carta do segundo baralho coincida com a carta do primeiro baralho com quem emparelha. Qual é a probabilidade de vitória?*

Uma das fórmulas mais úteis na resolução destes problemas é o chamado Princípio da Inclusão-Exclusão, também conhecido por *fórmula do crivo* ou *fórmula de da Silva-Sylvester*[7]. Este princípio estende a conjuntos arbitrários o princípio óbvio, habitualmente apelidado de *Princípio da Adição*, de que o cardinal da união de conjuntos disjuntos (dois a dois) é a soma dos cardinais de cada um desses conjuntos.

Por exemplo, no caso de três conjuntos A, B e C não necessariamente disjuntos, se somarmos os elementos em A, B e C (Fig. 1)

[7] O Princípio da Inclusão-Exclusão foi publicado pela primeira vez em 1854, num artigo de Daniel da Silva, e redescoberto mais tarde, em 1883, por Sylvester. Por isso, a fórmula do crivo e suas similares são, por vezes, apelidadas de fórmulas de da Silva ou de Sylvester. Realçamos o facto de Daniel da Silva, na opinião de Gomes Teixeira o mais notável matemático português do séc. XIX, ter sido estudante da Universidade de Coimbra; transcrevemos de [J. Silva Oliveira, *Daniel Augusto da Silva*, Boletim da SPM 2 (1979) 3-15]: «Daniel da Silva (1814-1878) foi, além de matemático eminente do seu tempo, oficial da Armada e professor da Escola Naval. Como estudante frequentou primeiro a Academia Real de Marinha e prosseguiu depois os seus estudos na Universidade de Coimbra onde se licenciou em Matemática e acabou por se doutorar.»

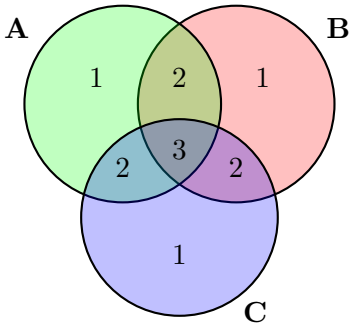


Figura 1: $|A| + |B| + |C|$.

estaremos a contar, uma vez cada um, os elementos de $A \setminus (B \cup C)$, os de $B \setminus (A \cup C)$ e os de $C \setminus (A \cup B)$, mas estaremos a contar por duas vezes os elementos de $(A \cap B) \setminus C$, $(A \cap C) \setminus B$ e $(B \cap C) \setminus A$, e, pior ainda, estaremos a contar por três vezes os elementos da intersecção $A \cap B \cap C$. Podemos começar por descontar os primeiros, subtraindo $|A \cap B|$, $|A \cap C|$ e $|B \cap C|$:

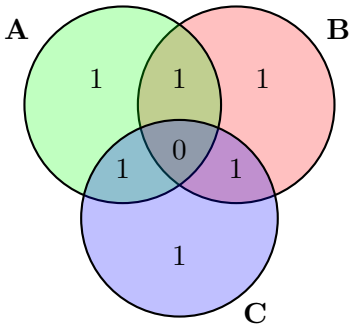


Figura 2: $|A| + |B| + |C| - (|A \cap B| + |A \cap C| + |B \cap C|)$.

Mas agora acabámos por descontar os elementos da intersecção $A \cap B \cap C$ mais do que devíamos (o zero na Fig. 2 indica que os elementos dessa região ainda não foram considerados para a contagem dos elementos de $A \cup B \cup C$), tendo que os repor novamente, para que a contagem fique finalmente certa:

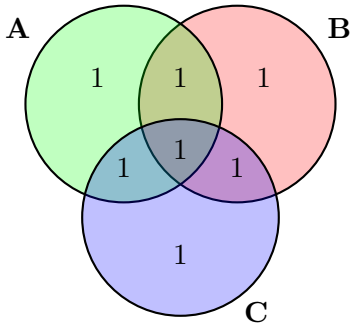


Figura 3: $|A| + |B| + |C| - (|A \cap B| + |A \cap C| + |B \cap C|) + |A \cap B \cap C|$.

No caso geral de n conjuntos $A_1, A_2, \ldots, A_n$, esta fórmula estende-se facilmente a (ver, por exemplo, [3] para uma demostração):

Proposição 6.1. [*Princípio da Inclusão-Exclusão*] *Para cada* $I \subseteq \{1, 2, \ldots, n\}$ *seja* $n(I) = |\cap_{i \in I} A_i|$. *O cardinal da união* $A_1 \cup A_2 \cup \cdots \cup A_n$ *é dado pela fórmula*

$$\sum_{|I|=1} n(I) - \sum_{|I|=2} n(I) + \sum_{|I|=3} n(I) - \cdots + (-1)^{n+1} \sum_{|I|=n} n(I).$$

O problema do jogo dos pares é um caso particular do bem conhecido **problema dos desencontros**:

Problema 2B. *Uma permutação* $a_{j_1} a_{j_2} \ldots a_{j_n}$ *de* $S = \{a_1, a_2, \ldots, a_n\}$ *diz-se um desencontro de* S *caso* $j_k \neq k$ *para qualquer* $k \in \{1, 2, \ldots, n\}$. *Quantas das* $n!$ *permutações de* S *são desencontros?*

De facto, denotando por D_n o número de desencontros de um conjunto com n elementos, a probabilidade de vitória no jogo dos pares é evidentemente igual a

$$\frac{D_{52}}{52!}.$$

Deixamos agora ao cuidado do leitor o exercício (recorrente em qualquer curso de Matemática Discreta) de verificar, com a ajuda do Princípio da Inclusão-Exclusão, que, para cada $n \in \mathbb{N}$,

$$D_n = \sum_{r=0}^{n} (-1)^{n-r} \binom{n}{r} r! = n! \left(\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \right). \tag{6.1.1}$$

n	2	3	4	5	6	7	8	9	10	11
D_n	1	2	9	44	265	1854	14833	133496	1334961	14684570

Aqui, o segredo de aplicação do princípio de da Silva reside na constatação de que, não sendo fácil determinar directamente o número D_n, é, por outro lado, imediato o cálculo, para cada k, do número de permutações $a_{j_1}a_{j_2}\ldots a_{j_n}$ de S tais que $j_k = k$ (portanto aquelas em que a_k está na sua posição original), bem como o cálculo para cada par k, l, do número de permutações nas quais os elementos a_k e a_l estão nas suas posições primitivas k e l, etc.:

$$(n-1)! \text{ no primeiro caso,} \quad (n-2)! \text{ no segundo caso, etc.}$$

7 Inversões de Möbius

> «The apex of mathematical achievement occurs when two or more fields which were thought to be entirely unrelated turn out to be closely intertwined. Mathematicians have never decided whether they should feel excited or upset by such events.»
>
> — GIAN-CARLO ROTA
> ('A Mathematician's Gossip', em: Indiscrete Thoughts, 1997)

Como veremos mais adiante, o Princípio da Inclusão-Exclusão pode obter-se como exemplo de aplicação do processo de inversão de Möbius num conjunto parcialmente ordenado. A inversão de Möbius [13], introduzida originalmente em 1832 por August Ferdinand Möbius (1790-1868) no contexto da teoria dos números, pode ser descrita, nos seus aspectos básicos, do seguinte modo. Uma *função aritmética* é uma função $\mathbb{N} \to \mathbb{R}$ (ou, mais geralmente, $\mathbb{N} \to \mathbb{C}$, mas aqui consideraremos apenas funções reais). Qualquer par de funções aritméticas f, g tem um produto *de convolução* (de Dirichlet) $f * g$, definido por

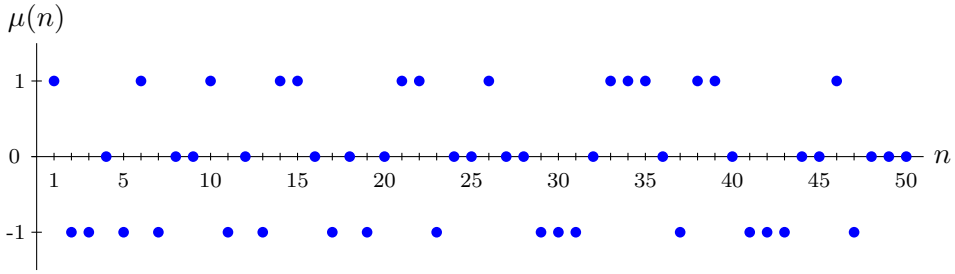
$$(f*g)(n) = \sum_{k,m:\ km=n} f(k)\, g(m) = \sum_{d:\ d|n} f(\frac{n}{d})\, g(d).$$

Este produto tem uma identidade: a *função unidade* δ definida por $\delta(1) = 1$ e $\delta(n) = 0$ para $n \geq 2$. A função constante $\zeta = (1, 1, \ldots)$ tem um inverso: a *função de Möbius* μ, muito usada em teoria dos números[8], definida para

[8] A função μ aparece já implicitamente nos trabalhos de Euler (em 1748) mas foi Möbius o primeiro a investigar de modo sistemático as suas propriedades (em 1832).

cada natural n, com factorização prima $p_1^{n_1}p_2^{n_2}\cdots p_k^{n_k}$ $(n_i \geq 1)$, por

$$\mu(n) \equiv \begin{cases} 1 & \text{se } n = 1 \\ (-1)^k & \text{se } n_1 = n_2 = \cdots = n_k = 1 \\ 0 & \text{se } n_i \geq 2 \text{ para algum } i. \end{cases}$$



Note que

$$\sum_{d:\, d|n} \mu(d) = \begin{cases} 1 & \text{se } n = 1, \\ 0 & \text{senão.} \end{cases} \tag{7.0.1}$$

De facto, para qualquer natural $n > 1$ com factorização prima $p_1^{n_1}p_2^{n_2}\cdots p_k^{n_k}$ $(n_i \in \mathbb{N})$, tem-se

$$\begin{aligned} \sum_{d:\, d|n} \mu(d) &= \mu(1) + \sum_{i=1}^{k} \mu(p_i) + \sum_{1\leq i_1<i_2\leq k} \mu(p_{i_1}p_{i_2}) + \cdots + \mu(p_1p_2\cdots p_k) \\ &= 1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + \cdots + \binom{k}{k}(-1)^k \\ &= (1 + (-1))^k = 0. \end{aligned}$$

É este facto que torna μ tão relevante na teoria das funções aritméticas e está na base da fórmula seguinte da *inversão de Möbius*:

Teorema 7.1. *Se $f, g\colon \mathbb{N} \to \mathbb{R}$ são funções aritméticas tais que*

$$f(n) = \sum_{d:\, d|n} g(d) \quad \textit{para qualquer } n \in \mathbb{N}, \tag{7.1.1}$$

então podemos recuperar g com a identidade

$$g(n) = \sum_{d:\, d|n} f(\frac{n}{d})\, \mu(d) \quad \textit{para qualquer } n \in \mathbb{N}. \tag{7.1.2}$$

Demonstração. Este resultado é trivial quando formulado na linguagem do *anel de Dirichlet* (o conjunto das funções aritméticas com a soma de funções usual, ponto a ponto, e o produto de Dirichlet): a identidade (7.1.1) significa simplesmente $f = g * \zeta$ pelo que, imediatamente, $g = f * \zeta^{-1} = f * \mu$, isto é, (7.1.2). □

Exemplos 7.2. (1) Seja φ a função *tociente* de Euler, definida por

$$\varphi(n) = |\{k \in \{1, 2, \ldots, n\} \colon \text{mdc}(k, n) = 1\}| .$$

A fórmula da inversão de Möbius (7.1.2), aplicada à identidade bem conhecida de Gauss

$$n = \sum_{d \colon d|n} \varphi(d)$$

dá imediatamente

$$\varphi(n) = \sum_{d \colon d|n} \frac{n}{d}\, \mu(d) = n \sum_{d \colon d|n} \frac{\mu(d)}{d}. \tag{7.2.1}$$

Assim, por exemplo,

$$\begin{aligned} \varphi(36) = \varphi(2^2 \times 3^2) &= 36 \left(\frac{\mu(1)}{1} + \frac{\mu(2)}{2} + \frac{\mu(3)}{3} + \frac{\mu(2 \times 3)}{2 \times 3} \right) \\ &= 36 \left(1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{6} \right) = 12. \end{aligned}$$

(2) Vejamos agora um exemplo um pouco mais complicado: determinemos o número de elementos do conjunto $C(n)$ de sequências *circulares* de zeros e uns de comprimento n (ou seja, sequências $a_1 a_2 \ldots a_n \in \{0, 1\}^n$ onde quaisquer duas sequências, uma das quais se obtém da outra por rotação, são consideradas iguais). Para isso, seja $\overline{P}(n)$ o conjunto de sequências circulares de comprimento n que não são periódicas (por exemplo, 010011 não é periódica, enquanto 010010 o é). Dada uma sequência arbitrária $S \in C(n)$, das duas uma: ou S não é periódica, isto é, $S \in \overline{P}(n)$, ou S é periódica, de período $d \mid n$. No segundo caso,

$$S = a_1 a_2 \ldots a_d a_1 a_2 \ldots a_d \cdots a_1 a_2 \ldots a_d,$$

pelo que podemos supor $S = a_1 a_2 \ldots a_d \in \overline{P}(d)$. Portanto,

$$|C(n)| = \sum_{d \colon d|n} |\overline{P}(d)|.$$

O problema resume-se então à determinação do número $|\overline{P}(d)|$. Mas como cada sequência em $\overline{P}(d)$ é igual às suas d permutações cíclicas (obtidas por rotação), tem-se

$$\sum_{d\colon\, d|n} d\,|\overline{P}(d)| = 2^n.$$

Logo, pelo Teorema 7.1,

$$n\,|\overline{P}(n)| = \sum_{d\colon\, d|n} 2^{\frac{n}{d}}\,\mu(d) = \sum_{d\colon\, d|n} 2^d\,\mu(\frac{n}{d})$$

e, consequentemente,

$$\begin{aligned}|C(n)| &= \sum_{d\colon\, d|n} |\overline{P}(d)| = \sum_{d\colon\, d|n} \frac{1}{d} \sum_{k\colon\, k|d} 2^k\,\mu(\frac{d}{k}) \\ &= \sum_{d\colon\, d|n} \sum_{k\colon\, k|d} \frac{2^k}{k\,\frac{d}{k}}\,\mu(\frac{d}{k}) = \sum_{k\colon\, k|n} \frac{2^k}{k} \sum_{l\colon\, l|\frac{n}{k}} \frac{\mu(l)}{l}.\end{aligned}$$

Finalmente, por (7.2.1),

$$|C(n)| = \frac{1}{n} \sum_{k\colon\, k|n} \varphi(\frac{n}{k})\,2^k. \tag{7.2.2}$$

8 Resolvendo o Problema 2 com inversões de Möbius-Rota

> «We often hear that mathematics consists mainly of 'proving theorems'. Is a writer's job mainly that of 'writing sentences'?»
>
> — Gian-Carlo Rota
>
> (Prefácio a [P. Davis e R. Hersh, *The Mathematical Experience*], 1981)

Em 1964, num artigo revolucionário [16], Gian-Carlo Rota (1932-1999) generalizou a inversão de Möbius a quaisquer conjuntos parcialmente ordenados *localmente finitos*, com o intuito de a tornar útil também na combinatória (e na teoria dos grupos). Estes conjuntos localmente finitos são aqueles $(P, \leq)$ nos quais qualquer intervalo

$$[x, y] := \{z \in P \mid x \leq z \leq y\}$$

é finito.

O par $(\mathbb{N}_0, \leq)$ é um exemplo de conjunto parcialmente ordenado localmente finito; outro é o par $(\mathbb{N}, |)$ dos inteiros positivos com a relação de divisibilidade.

Definição 8.1. Seja $(P, \leq)$ um conjunto parcialmente ordenado localmente finito e denotemos por $\text{int}(P)$ o respectivo conjunto de intervalos. A *álgebra de incidência* [16, 7], $\mathcal{I}(P)$, é o conjunto das funções

$$f\colon \text{int}(P) \to \mathbb{R}.$$

Abreviaremos $f([x,y])$ por $f(x,y)$. Se convencionarmos que $f(x,y) = 0$ sempre que $x \not\leq y$, podemos considerar cada $f \in \mathcal{I}(P)$ como uma função $P \times P \to \mathbb{R}$.

Qualquer álgebra de incidência $\mathcal{I}(P)$ é um espaço vectorial real com as operações de adição e multiplicação escalar definidas ponto a ponto. Com o produto de convolução

$$(f * g)(x,y) = \sum_{z \in [x,y]} f(x,z)g(z,y)$$

torna-se uma álgebra associativa. A identidade de $\mathcal{I}(P)$ é a função de Kronecker

$$\delta_P(x,y) \equiv \begin{cases} 1 & \text{se } x = y \\ 0 & \text{caso contrário} \end{cases}$$

e $f \in \mathcal{I}(P)$ é invertível se e só se $f(x,x) \neq 0$ para qualquer x. A função de Möbius μ_P é definida recursivamente sobre o comprimento dos intervalos:

(M1) $\mu_P(x,x) = 1$ para qualquer $x \in P$.

(M2) Se $x \not\leq y$, então $\mu_P(x,y) = 0$.

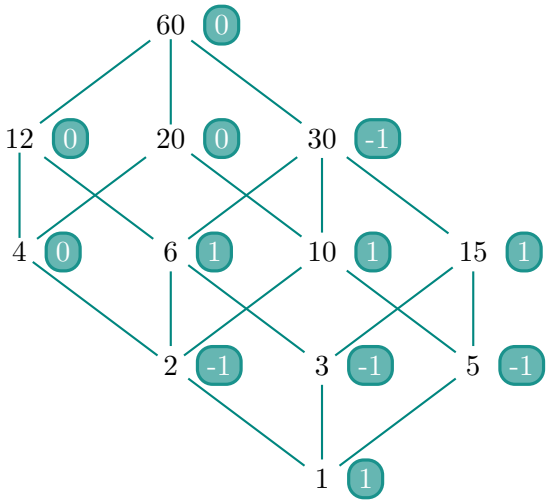
(M3) Se $x < y$, então $\mu_P(x,y) = -\sum_{z\colon x \leq z < y} \mu_P(x,z)$.

É surpreendente o dom da ubiquidade, na combinatória, da função de Möbius de um conjunto parcialmente ordenado.

Exemplos 8.2. (1) Calculemos alguns valores particulares de $\mu_P(1,n)$, no caso

$$(P, \leq) := (\mathbb{N}, |),$$

enumerados nos círculos da figura seguinte:



Estes valores correspondem precisamente à função de Möbius clássica:

$$\mu_P(1,n) = \mu(n) \text{ para qualquer } n \in \mathbb{N}.$$

De facto, $\mu_P(1,1) = 1 = \mu(1)$; para $n > 1$, supondo, por hipótese de indução, que $\mu_P(1,d) = \mu(d)$ para qualquer $d < n$ obtemos, usando (M3) e (7.0.1):

$$\mu_P(1,n) = -\sum_{d:\ 1\leq d<n} \mu_P(1,d) = -\sum_{d|n,\, d\neq n} \mu(d) = \mu(n).$$

Observe também como a função de Kronecker corresponde à função δ clássica: $\delta_P(1,n) = \delta(n)$ para qualquer natural n.

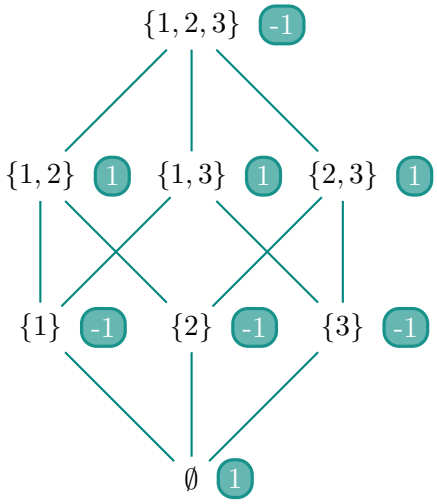
(2) No caso da álgebra de Boole

$$(P,\leq) = (\mathcal{P}(X_n),\subseteq)$$

sobre $X_n = \{1,2,\ldots,n\}$, para cada par $S \subseteq T$ de elementos de P,

$$\mu_P(S,T) = (-1)^{|T|-|S|}. \tag{8.2.1}$$

A figura seguinte enumera os diferentes valores de $\mu_P(\emptyset,T)$ no caso $n = 3$:



A prova de (8.2.1) segue por indução[9] sobre $|T| - |S|$:

Se $S = T$ então, por (M1), $\mu_P(S,T) = 1$ e (8.2.1) confirma-se; se $S \neq T$ e $p = |T \smallsetminus S| = |T| - |S|$, por (M3) e pela hipótese de indução obtemos

$$\mu_P(S,T) = -\sum_{R:\ S\subseteq R\subset T} \mu_P(S,R) = -\sum_{R:\ S\subseteq R\subset T} (-1)^{|R|-|S|} = -\sum_{k=0}^{p-1} (-1)^k \binom{p}{k}.$$

Como $0 = (1-1)^p = \sum_{k=0}^{p} (-1)^k \binom{p}{k}$, então, finalmente,

$$\mu_P(S,T) = (-1)^p \binom{p}{p} = (-1)^p = (-1)^{|T|-|S|}.$$

Rota mostrou que a função zeta[10] definida por $\zeta_P(x,y) = 1$ para quaisquer $x \leq y$ é a inversa de μ_P. Portanto, mais uma vez,

$$f = g * \zeta_P \quad \Rightarrow \quad g = f * \mu_P,$$

[9] Outra maneira de concluir isto é observar que: (1) P é isomorfo a $\mathcal{P}(\{1\})^n = \mathbf{2}^n$; (2) $\mathbf{2}$ pode ser visto como o intervalo $[0,1] \subseteq (\mathbb{N}, \leq)$; (3) portanto, a função de Möbius de $\mathbf{2}$ é precisamente $(-1)^n$, $n \in \{0,1\}$; (4) a função de Möbius do produto directo de dois conjuntos parcialmente ordenados P_1, P_2 é o produto das funções de Möbius de P_1 e P_2, ou seja, é dada por $\mu_{P_1\times P_2}((x,y),(x',y')) = \mu_{P_1}(x,x')\mu_{P_2}(y,y')$ para quaisquer $(x,y) \leq (x',y')$ em $P_1 \times P_2$ (regra do produto, uma das ferramentas mais úteis para calcular funções de Möbius em conjuntos parcialmente ordenados [16]).

[10] Sim, ζ como na função de Riemann, e não é por acaso!

Para mais informação ver as primeiras postagens em [*The n-Category Café*, `golem.ph.utexas.edu/category/2011/05/mbius_inversion_for_categories.html`].

Sobre a motivação e desenvolvimentos das ideias de Rota na transferência de conceitos da teoria dos números para a combinatória, ver [4].

ou seja:

Teorema 8.3 (Inversão de Möbius-Rota)**.** *Sejam $f, g \in \mathcal{I}(P)$. Se $f(x,y) = \sum_{z\in[x,y]} g(x,z)$, então $g(x,y) = \sum_{z\in[x,y]} f(x,z)\,\mu_P(z,y)$.* □

Quando P possui primeiro elemento 0, a restrição de $f\colon P \times P \to \mathbb{R}$ a $\{0\} \times P$, que continuaremos a designar por f, pode ser vista como uma função $f\colon P \to \mathbb{R}$, $f(y) = f(0,y)$. O Teorema anterior, no caso particular $x = 0$, garante então o seguinte:

Corolário 8.4. *Sejam $f, g\colon P \to \mathbb{R}$. Se $f(y) = \sum_{z\le y} g(z)$, então $g(y) = \sum_{z\le y} f(z)\,\mu_P(z,y)$.* □

De modo perfeitamente análogo ao que fizemos com as quase-inversões de Galois no Problema 1, a inversão de Möbius-Rota (Corolário 8.4) resolve o problema dos desencontros (Problema 2) de modo trivial a partir da correspondente identidade inversa, que é óbvia. Com efeito, denotemos por $\mathrm{Per}(X_n)$ o conjunto das $n!$ permutações dos elementos de $X_n = \{1, 2, \ldots, n\}$. Basta então tomar para P a álgebra de Boole $(\mathcal{P}(X_n), \subseteq)$ do Exemplo 8.2(2). Se considerarmos, para cada $S \in P$, o conjunto $\mathrm{Des}(S)$ das permutações $(p_1, p_2, \ldots, p_n)$ de X_n nas quais $p_i \neq i$ para todo o $i \in S$, é evidente que

$$\mathrm{Per}(X_n) = \bigcup_{S\subseteq X_n} \mathrm{Des}(S) \quad \text{(união disjunta)}$$

pelo que

$$|\mathrm{Per}(X_n)| = \sum_{S\subseteq X_n} |\mathrm{Des}(S)|\,.$$

Logo, pela inversão de Möbius-Rota, tomando para $f, g\colon P \to \mathbb{R}$ as funções $S \mapsto |\mathrm{Per}(S)|$ e $S \mapsto |\mathrm{Des}(S)|$, respectivamente, obtemos

$$\begin{aligned}
|\mathrm{Des}(X_n)| &= \sum_{S\subseteq X_n} |\mathrm{Per}(S)|\ \mu_P(S, X_n) = \sum_{S\subseteq X_n} (-1)^{n-|S|}\ |\mathrm{Per}(S)| \\
&= \cancel{n!} - \binom{n}{n-1}\cancel{(n-1)!} + \binom{n}{n-2}(n-2)! - \cdots + (-1)^n\, 0! \\
&= \frac{n!}{2!} - \frac{n!}{3!} + \frac{n!}{4!} - \cdots + (-1)^n\,\frac{n!}{n!},
\end{aligned}$$

precisamente a fórmula (6.1.1).

O paralelismo com as conexões de Galois é evidente: resolver problemas complicados, por inversão, a partir de identidades mais óbvias no lado oposto.

Para terminar esta secção, vejamos como se pode obter o Princípio da Inclusão-Exclusão como caso particular de aplicação da inversão de Möbius-Rota. Sejam $A_1, A_2, \ldots, A_n$ subconjuntos de um conjunto finito X. Designaremos por $\overline{A_i}$, $i = 1, 2, \ldots, n$, o complementar de A_i em X. Consideremos mais uma vez $P = (\mathcal{P}(X_n), \subseteq)$, e a função $f\colon P \to \mathbb{R}$ definida por

$$f(I) = \left|\{x \in X \mid x \in \overline{A_i} \text{ sse } i \in I\}\right| = \left|\bigcap_{i\in I} \overline{A_i} \cap \bigcap_{i\in X_n \smallsetminus I} A_i\right|.$$

Claro que $f(X_n) = \left|\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}\right|$. Além disso, seja $g\colon P \to \mathbb{R}$ a função definida por $g(I) = \sum_{J\subseteq I} f(J)$. Não é difícil provar que

$$g(I) = |\bigcap_{i\in X_n \smallsetminus I} A_i|.$$

Aplicando a inversão de Möbius-Rota, podemos então concluir que

$$f(I) = \sum_{J\subseteq I} g(J)\, \mu_P(J, I) = \sum_{J\subseteq I} (-1)^{|I|-|J|} g(J).$$

Em particular,

$$\begin{aligned}\left|\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}\right| = f(X_n) &= \sum_{J\subseteq X_n} (-1)^{n-|J|}\, g(J) \\ &= \sum_{J\subseteq X_n} (-1)^{n-|J|}\, |\bigcap_{j\in X_n \smallsetminus J} A_j| = \sum_{I\subseteq X_n} (-1)^{|I|}\, |\bigcap_{i\in I} A_i|.\end{aligned}$$

Concluindo, $|A_1 \cup A_2 \cup \cdots \cup A_n|$ é igual a

$$\begin{aligned}|X| - \left|\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}\right| &= |X| - \left(\sum_{I\subseteq X_n} (-1)^{|I|}\, |\bigcap_{i\in I} A_i|\right) \\ &= |X| - \left(|X| + \sum_{\varnothing\neq I\subseteq X_n} (-1)^{|I|}\, |\bigcap_{i\in I} A_i|\right) \\ &= \sum_{\varnothing\neq I\subseteq X_n} (-1)^{|I|+1}\, |\bigcap_{i\in I} A_i|,\end{aligned}$$

que é precisamente a fórmula do Princípio da Inclusão-Exclusão em 6.1.

9 Ilustrando as potencialidades do método

> «Mathematics is the study of analogies between analogies. All science is. Scientists want to show that things that don't look alike are really the same. That is one of their innermost Freudian motivations. In fact, that is what we mean by understanding.»
> — GIAN-CARLO ROTA
> ('A Mathematician's Gossip', em: Indiscrete Thoughts, 1997)

Exemplos 9.1. (1) De modo análogo, o número de funções sobrejectivas de um conjunto X num conjunto Y pode ser calculado observando primeiro que

$$\mathrm{Func}(X,Y) = \bigcup_{S\subseteq Y} \mathrm{Sobrej}(X,S) \quad \text{(união disjunta)}$$

(uma vez que cada função $f\colon X \to Y$ pode ser identificada pela função sobrejectiva $f\colon X \to f[X]$). Portanto,

$$|\mathrm{Func}(X,Y)| = \sum_{S\subseteq Y} |\mathrm{Sobrej}(X,S)|\,.$$

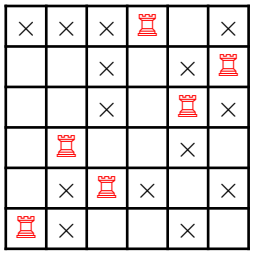
Pela inversão de Möbius-Rota, considerando $f, g\colon \mathcal{P}(Y) \to \mathbb{R}$ dadas respectivamente por $S \mapsto |\mathrm{Func}(X,S)|$ e $S \mapsto |\mathrm{Sobrej}(X,S)|$, obtemos

$$|\mathrm{Sobrej}(X,Y)| = \sum_{S\subseteq Y} (-1)^{|Y|-|S|}\ |\mathrm{Func}(X,S)|$$

pelo que, se X tem m elementos e Y tem n elementos ($m \geq n$), então

$$|\mathrm{Sobrej}(X,Y)| = \sum_{r=0}^{n} (-1)^{n-r} \binom{n}{r} r^m. \tag{9.1.1}$$

(2) Apliquemos agora a inversão de Möbius-Rota ao problema, mais complicado, do cálculo do número de maneiras de dispor n torres num tabuleiro de xadrez $n \times n$ com *posições proibidas*, de modo a não se atacarem mutuamente (isto é, de modo a que nenhum par de torres esteja numa linha ou coluna comuns). Por exemplo, no caso $n = 6$, com posições proibidas marcadas com $\times$, uma dessas configurações é a seguinte:



Representemos o tabuleiro com posições proibidas pela matriz binária (0: posição proibida)

$$T = [t_{ij}] = \begin{bmatrix} 0 & 0 & 0 & \mathbf{1} & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & \mathbf{1} \\ 1 & 1 & 0 & 1 & \mathbf{1} & 0 \\ 1 & \mathbf{1} & 1 & 1 & 0 & 1 \\ 1 & 0 & \mathbf{1} & 0 & 1 & 0 \\ \mathbf{1} & 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

É claro que colocar seis torres no tabuleiro, sem se atacarem, corresponde a uma colecção de seis 1's em T (marcados a carregado na matriz) com a seguinte propriedade: cada linha e cada coluna contém exactamente um desses 1's. Podemos ainda representar esta colecção pela sequência (permutação) dos números das colunas onde estão esses 1's (começando de cima para baixo, a partir da linha 1): $(4, 6, 5, 2, 3, 1)$. Em geral, n torres colocadas num tabuleiro $n \times n$ correspondem a uma permutação σ (no grupo simétrico S_n de todas as permutações de X_n) com $t_{i\,\sigma(i)} = 1$ $(i = 1, 2, \ldots, n)$ e, portanto,

$$\prod_{i=1}^{n} t_{i\,\sigma(i)} = t_{1\,\sigma(1)} t_{2\,\sigma(2)} \cdots t_{n\,\sigma(n)} = 1.$$

Como este produto só poderá ser, além de 1, igual a 0 (precisamente quando pelo menos uma das torres for colocada numa posição proibida), torna-se evidente que o número que queremos calcular é precisamente o chamado *permanente* da matriz T, ou seja, a soma

$$\sum_{\sigma \in S_n} \prod_{i=1}^{n} t_{i\,\sigma(i)}. \tag{9.1.2}$$

Consideremos novamente $P = (\mathcal{P}(X_n), \subseteq)$. Cada subconjunto S de cardinalidade k de X_n corresponde a uma escolha de k colunas de T. Seja $\mathrm{Func}(X_n, S)$ o conjunto de todas as funções $\sigma\colon X_n \to S$ e seja $\mathrm{Sobrej}(X_n, S)$

o respectivo subconjunto de funções sobrejectivas. Como já observámos no exemplo anterior, $\mathrm{Func}(X_n, S)$ é a união disjunta $\bigcup_{R\subseteq S} \mathrm{Sobrej}(X_n, R)$.

Consideremos agora a função $f\colon \mathcal{P}(X_n) \to \mathbb{R}$ definida por

$$f(S) = \sum_{\sigma\in \mathrm{Sobrej}(X_n,S)} \prod_{i=1}^{n} t_{i\,\sigma(i)}.$$

Note que $f(\varnothing) = 0$ e que $f(X_n)$ é a soma (9.1.2), uma vez que qualquer função sobrejectiva $\sigma\colon X_n \to X_n$ é uma bijecção.

Finalmente, definindo

$$g(S) = \sum_{R\subseteq S} f(R) \quad (S \in \mathcal{P}(X_n)),$$

a inversão de Möbius-Rota diz-nos que

$$f(X_n) = \sum_{S\subseteq X_n} (-1)^{n-|S|} g(S).$$

Mas, contrariamente a $f(X_n)$ (mais geralmente $f(S)$), não é difícil calcular o valor de $g(S)$ directamente, pelo que conseguiremos ter assim uma forma de cálculo para $f(X_n)$. De facto, não é difícil concluir que

$$g(S) = \sum_{\theta\in \mathrm{Func}(X_n,S)} t_{1\,\theta(1)}\, t_{2\,\theta(2)} \cdots t_{n\,\theta(n)} \quad (S \in \mathcal{P}(X_n))$$

é igual a

$$\Big(\sum_{j\in S} t_{1j}\Big) \cdot \Big(\sum_{j\in S} t_{2j}\Big) \cdots \Big(\sum_{j\in S} t_{nj}\Big).$$

Logo

$$f(X_n) = \sum_{S\subseteq X_n} (-1)^{n-|S|} \prod_{i=1}^{n} \Big(\sum_{j\in S} t_{ij}\Big). \tag{9.1.3}$$

Temos assim uma fórmula de cálculo do número de maneiras de dispor n torres num tabuleiro de xadrez $n \times n$ com *posições proibidas*, de modo a não se atacarem mutuamente, de fácil implementação computacional: escolher um conjunto S de colunas, calcular a soma dos elementos de cada linha nessas colunas, multiplicar estas somas todas, justapor-lhe o sinal apropriado, e somar os resultados sobre todas as escolhas de S (o número de parcelas é igual a 2^n mas algumas são nulas, precisamente as correspondentes a conjuntos S para os quais a matriz tem uma linha que só tem zeros nas colunas de S). Deixamos a cargo do leitor o cálculo (um pouco fastidioso...) desse número no tabuleiro do exemplo inicial com matriz associada T.

10 Mais sobre funções de Möbius-Rota

> «We tend to think of generating functions as related to combinatorics, and Dirichlet series as related to number theory. But this is because combinatorists prefer adding finite sets, while number theorists get more excited about multiplying them. (Primes and all that.)»
>
> — John Baez
> (The n-Category Café, Maio de 2011)

Terminamos o artigo com o retorno às conexões de Galois num resultado de Rota [16] que mostra como se relacionam as funções de Möbius disponíveis em cada um dos lados da conexão.

Teorema 10.1. *Sejam P e Q conjuntos parcialmente ordenados finitos, P com primeiro e último elementos 0 e 1 (diferentes) e Q com último elemento 1. Sejam μ_P e μ_Q as respectivas funções de Möbius. Se $f\colon P \to Q$ e $g\colon Q \to P$ constituírem uma conexão de Galois $f \dashv g$ tal que*

(G1) *$f(a) = 1$ se e só se $a = 1$,*

(G2) *$g(1) = 1$,*

então

$$\mu_P(0,1) = \sum_{y<1} \mu_Q(y,1)\,\zeta(g(y),0) = \sum_{y\colon g(y)=0} \mu_Q(y,1).$$

Demonstração. Como $f(a) \leq b$ se e só se $a \leq g(b)$, então, para cada $b \in Q$,

$$\sum_{y\colon y\leq b} \delta(f(a),y) = \zeta_Q(a,g(b)). \tag{10.1.1}$$

Aplicando a (10.1.1) a inversão de Möbius-Rota relativamente a Q obtemos

$$\delta(f(a),1) = \sum_{y\leq 1} \mu_Q(y,1)\,\zeta(a,g(y)). \tag{10.1.2}$$

$\delta(f(a),1)$ toma o valor 1 se e só se $f(a) = 1$, isto é, $a = 1$, por (G1). Para os restantes valores de a, $\delta(f(a),1) = 0$. Assim, $\delta(f(a),1) = 1-\zeta(a,1)+\delta(a,1)$. Denotando a função de incidência $\zeta - \delta$ por n, temos $\delta(f(a),1) = 1 - n(a,1)$ e a identidade (10.1.2) pode então ser reescrita como

$$1 - n(a,1) = \zeta(a,g(1)) + \sum_{y<1} \mu_Q(y,1)\,\zeta(a,g(y)).$$

Mas a condição (G2) implica $\zeta(a, g(1)) = \zeta(a, 1) = 1$ para qualquer $a \in P$. Portanto,

$$-n(a, 1) = \sum_{y<1} \mu_Q(y, 1)\, \zeta(a, g(y)).$$

Agora, como $\zeta = \delta + n$, temos $\delta - \mu * n = \delta - \mu * (\zeta - \delta) = \delta - \delta + \mu * \delta = \mu$. Logo

$$\mu_P(0, 1) = - \sum_{0 \le a \le 1} \mu_P(0, a)\, n(a, 1) = \sum_{0 \le a \le 1} \sum_{y<1} \mu_Q(y, 1) \mu_P(0, a) \zeta(a, g(y)).$$

Finalmente, trocando a ordem dos somatórios,

$$\mu_P(0, 1) = \sum_{y<1} \mu_Q(y, 1) \sum_{0 \le a \le 1} \mu_P(0, a) \zeta(a, g(y))$$

e o somatório mais à direita é igual a $(\mu_P * \zeta)(0, g(y)) = \delta(0, g(y)) = \zeta(g(y), 0)$. □

Com este resultado, fixando um dos conjuntos parcialmente ordenados, e variando o outro entre ordens parciais nos quais a função de Möbius é bem conhecida, podemos obter informação sobre a função de Möbius no conjunto previamente fixado.

Comentários finais. Do mesmo modo que uma conexão de Galois é um exemplo muito particular de um conceito fundamental da moderna teoria das categorias – o conceito de adjunção –, base de muitos teoremas importantes que relacionam áreas distintas da matemática, as inversões de Möbius também podem ser formuladas categorialmente em contextos mais gerais que os clássicos (o conjunto parcialmente ordenado dos inteiros positivos, ordenado pela relação de divisibilidade, no caso de Möbius, e qualquer conjunto parcialmente ordenado *localmente finito*, na generalização de Rota). As ideias de Rota criaram as condições para essas recentes extensões (a categorias com alguma condição de *finitude*). Isto foi feito essencialmente de dois modos distintos, por Content, Lemay e Leroux [5] e, mais recentemente, de um modo mais geral, por Leinster [11, 12]. A segunda abordagem faz parte da teoria da característica de Euler de uma categoria [11], que coincide com a característica de Euler topológica quando esta existe (mas é também válida em situações diversas nas quais esta não existe). Curiosamente, neste contexto categorial é possível generalizar ainda mais o princípio da inclusão-exclusão a fórmulas sobre cardinais de colimites de conjuntos!

«Go to the roots, of these calculations! Group the operations. Classify them according to their complexities rather than their appearances! This, I believe, is the mission of future mathematicians. This is the road on which I am embarking in this work.»

— Évariste Galois

(Do prefácio do seu último manuscrito, 1832)

Referências

[1] S. Beatty, Problem 3173, *Amer. Math. Monthly* 33 (1926) 159.

[2] G. Birkhoff, *Lattice Theory*, Amer. Math. Soc. Colloq. Publ. 25 (1967).

[3] R. Brualdi, *Introductory Combinatorics*, 5^a edição, Prentice Hall (2010).

[4] T. Y. Chow, The combinatorics behind number-theoretic sieves, *Adv. Math.* 138 (1998) 293-305.

[5] M. Content, F. Lemay e P. Leroux, Catégories de Möbius et fonctorialités: Un cadre général pour l'inversion de Möbius, *Journal of Combinatorial Theory Series A* 28 (1980) 169-190.

[6] K. Denecke, M. Erné e S. L. Wismath, *Galois Connections and Applications*, Mathematics and its Applications, vol. 565, Kluwer, Dordrecht (2004).

[7] P. Doubilet, G.-C. Rota e R. P. Stanley, On the foundations of combinatorial theory VI: The idea of generating function, *Berkeley Symp. on Math. Statist. and Prob.*, vol. 2, pp. 267-318, Univ. of Calif. Press (1972).

[8] M. Erné, J. Koslowski, A. Melton e G. Strecker, A primer on Galois connections, *Papers on general topology and applications* (Madison, WI, 1991), pp. 103–125, Ann. New York Acad. Sci., 704 (1993).

[9] J. Lambek, Some Galois connections in elementary number theory, *J. Number Theory* 47 (1994) 371–377.

[10] J. Lambek e L. Moser, Inverse and complementary sequences of natural numbers, *Amer. Math. Monthly* 61 (1954) 454-458.

[11] T. Leinster, The Euler characteristic of a category, *Documenta Math.* 13 (2008) 21-49.

[12] T. Leinster, Notions of Möbius inversion, *Bulletin of the Belgian Mathematical Society* 19 (2012) 911-935.

[13] A. F. Möbius, Über eine besondere Art von Umkehrung der Reihen, *J. reine angew. Math.* 9 (1832) 105-123.

[14] O. Ore, Galois connexions, *Trans. Amer. Math. Soc.* 55 (1944) 493-513.

[15] J. Picado e A. Pultr, *Frames and locales: Topology without points*, Frontiers in Mathematics, vol. 28, Springer, Basel (2012).

[16] G.-C. Rota, On the foundations of combinatorial theory I: Theory of Möbius functions, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 2 (1964) 340-368.