por Pedro Martins Rodrigues
[Instituto Superior Técnico]

Uma Expansão-Beta?

Uma expansão-beta é o análogo da usual expansão decimal de um número real, mas numa base não inteira. A letra β (beta) do alfabeto grego é usada habitualmente para designar essa base.

As expansões-beta foram introduzidas pela primeira vez pelo matemático húngaro A. Rényi, em 1957, e têm vindo a ser estudadas até hoje, nomeadamente no contexto da área científica dos Sistemas Dinâmicos. Um aspecto de grande importância e interesse nesse estudo é o da relação entre certas propriedades das expansões-beta e a caracterização aritmética da base β.

Na introdução que se segue começaremos por definir expansão numa base β qualquer, e ao longo deste artigo faremos sistematicamente referência à expansão decimal de um número real, para pôr em evidência as semelhanças e diferenças com o caso de uma base não inteira. Na verdade, a construção da expansão decimal e as suas propriedades são semelhantes às da expansão em qualquer base inteira, pelo que consideraremos este caso mais geral. Quando for necessário frisar que nos estamos a restringir ao caso de a base ser inteira, usaremos a letra b para designar essa base.

Convém, antes de mais, fixar alguma notação: como habitualmente, $\mathbb{R}$ designa o conjunto dos números reais, $\mathbb{Q}$ o dos racionais e $\mathbb{Z}$ o dos inteiros. Dado um número real x, o símbolo $\lfloor x \rfloor$ designa a parte inteira de x, ou seja, $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$.

Por exemplo, $\lfloor 3{,}14 \rfloor = 3$, $\lfloor \sqrt{2}-1 \rfloor = 0$ e $\lfloor -3{,}001 \rfloor = -4$.

Seja então $\beta > 1$. Vamos concentrar-nos primeiro no caso de um real $0 \leq x_0 < 1$; definimos $e_1 = \lfloor \beta x_0 \rfloor$ e $x_1 = \beta x_0 - e_1$.

Temos obviamente: $0 \leq e_1 < \beta$ e $0 \leq x_1 < 1$; além disso $x_0 = e_1\beta^{-1} + x_1\beta^{-1}$.

Repetindo a operação,

$e_2 = \lfloor \beta x_1 \rfloor$ e $x_2 = \beta x_1 - e_2$ e substituindo o valor de x_1 na equação anterior,

$x = e_1\beta^{-1} + x_1\beta^{-1} = e_1\beta^{-1} + \beta^{-1}(e_2\beta^{-1} + x_2\beta^{-1}) = e_1\beta^{-1} + e_2\beta^{-2} + x_2\beta^{-2}$.

O leitor notará facilmente que podemos repetir esta operação um qualquer número n de vezes, obtendo a representação $x_0 = e_1\beta^{-1} + e_2\beta^{-2} + \ldots + e_n\beta^{-n} + x_n\beta^{-n}$ ou, numa notação mais precisa e sucinta,

$$x_0 = \sum_{k=1}^{n} e_k\beta^{-k} + x_n\beta^{-n}$$

em que as sucessões e_n e x_n são definidas a partir do termo inicial x_0 pela recorrência $e_n = \lfloor \beta x_{n-1} \rfloor$; $x_n = \beta x_{n-1}$ e satisfazem portanto as desigualdades $0 \leq e_n < \beta$ e $0 \leq x_n < 1$. Note-se que no caso de β ser não inteiro, a condição nos coeficientes pode ser escrita como $e_n \in \{0, 1, \ldots, \lfloor \beta \rfloor\}$, enquanto para uma base inteira b a condição é $e_n \in \{0, 1, \ldots, b-1\}$.

Outra fórmula que decorre facilmente das anteriores e que permite determinar o coeficiente e_n apenas à custa dos anteriores e de x_0 é a seguinte:

$$e_n = \left\lfloor \beta^n \left(x_0 - \sum_{k=1}^{n-1} e_k\beta^{-k}\right) \right\rfloor$$

De facto, os coeficientes de um x_0, relativamente a uma base β, podem ser definidos como a única sucessão $\{e_n\}$ de inteiros, satisfazendo, para todo o $n \geq 1$, as condições:

1. $0 \leq e_n < \beta$
2. $0 \leq x_0 - \sum_{k=1}^{n} e_k \beta^{-k} < \beta^{-n}$.

Como $\beta > 1$, fica claro que a sucessão $\sum_{k=1}^{n} e_k \beta^{-k}$ converge para x_0, ou seja, x_0 é a soma da série $\sum_{k>1} e_k \beta^{-k}$ que designamos por expansão de x_0 na base β.

A definição da expansão generaliza-se facilmente a qualquer $x>0$: se $x>1$, existe um inteiro (único) m para o qual se verifica $\beta^{m-1} \leq x < \beta^m$; o número real $x\beta^{-m}$, menor do que 1, tem uma expansão na base β, seja ela $\sum_{k>1} e'_k \beta^{-k}$; a série $\sum_{k \geq 1} e'_k \beta^{-k+m} = \sum_{j \geq -m+1} e'_{j+m} \beta^{-j} = \sum_{j \geq -m+1} e_j \beta^{-j}$ é então a expansão, na mesma base, de x, definindo a sucessão dos coeficientes para x por $e_j = e'_{j+m}$.

Se não existem dúvidas sobre qual a base considerada, a expansão pode ser representada pela sequência dos coeficientes $\{e_k\}_{k \geq -m+1}$ e se queremos dar conta apenas dos primeiros termos, podemos escrevê-los ordenadamente, sem os índices, usando um símbolo convencional para separar os coeficientes de índice negativo dos restantes, e indicando, por meio de reticências, por exemplo, que se está a apresentar apenas um bloco inicial da expansão. Temos assim, por exemplo, a familiar expressão
$\sqrt{2} = 1.4142135\ldots = 1 + 4(10)^{-1} + 1(10)^{-2} + 4(10)^{-3} + 2(10)^{-4} + 1(10)^{-5} + \ldots$ para a expansão na base 10 de $\sqrt{2}$, ou
$\sqrt{2} = 1.0110101\ldots = 1 + 2^{-2} + 2^{-3} + 2^{-5} + 2^{-7} + \ldots$ para a expansão do mesmo número na base 2, sendo que na última expressão se omitiram as potências com coeficiente zero.

O leitor é convidado a realizar alguns cálculos do mesmo tipo ou, melhor ainda, a escrever o seu próprio programa para calcular, até uma ordem qualquer, os coeficientes da expansão de um número dado numa determinada base inteira. A possibilidade de desenvolver esse cálculo está, em geral, dependente do grau de precisão com que conhecemos o número em questão (e, muitas vezes, conhecemos esse número exactamente através de uma parte da sua expansão decimal!), mas, no caso, por exemplo, de números da forma $\sqrt[m]{q}$, em que q é racional, o cálculo pode ser feito de modo exacto, com recurso apenas à aritmética dos inteiros. Essa possibilidade estende-se a outros números algébricos (ou seja, raízes de polinómios com coeficientes inteiros) e mesmo ao caso de bases não inteiras mas que sejam números racionais ou algébricos.

Por exemplo, se β for a raiz positiva do polinómio $z^2 - z - 1$, temos $\sqrt{2} = 1 + \beta^{-2} + \beta^{-8} + \beta^{-10} + \ldots$.

Evidentemente, nada impede que a expansão de um certo x_0 na base β seja finita, isto é, que para algum n se tenha $x_n = 0$ e portanto $e_k = 0$ para todo o $k > n$.

Outra situação digna de nota, e a que voltaremos mais adiante, é a de a expansão-β de um determinado número ser (eventualmente) periódica; esclareça-se que uma expansão $\sum_{k>-m+1} e_k \beta^{-k}$ se diz eventualmente periódica, ou apenas periódica, se existir um inteiro n e um inteiro positivo p tais que para todo o $k \geq n$ se tenha $e_{k+p} = e_j$. Se p for o menor inteiro positivo satisfazendo aquela condição, dizemos que a expansão tem período p, e representamos a sucessão dos coeficientes usando uma barra para indicar o bloco periódico, que se repete indefinidamente: $e_{-m} \ldots e_{n-1} e_n \ldots \bar{e}_{n+p-1}$. Se a sucessão se inicia logo com o bloco periódico, diz-se que é puramente periódica.

Expansões e sistemas dinâmicos simbólicos

Antes de continuarmos, vamos fazer uma breve digressão para pôr em evidência a seguinte interpretação da construção da expansão para $x_0 \in [0,1[$: a sucessão $\{x_n\}$ foi definida pela recorrência $x_n = f(x_{n-1})$ onde f designa a função definida por $f : [0,1[\to [0,1[\; f(x) = \beta x - \lfloor \beta x \rfloor$, e, portanto, se designarmos por $f^{(n)}$ a composição de f com ela própria, n vezes, deduzimos que $x_n = f^{(n)}(x_0)$. Descrevemos este facto dizendo que a sucessão $\{x_n\}$ constitui a órbita de x_0 por iteração de f[1].

Este é um exemplo relativamente simples de um sistema dinâmico discreto: temos um conjunto, neste caso o intervalo, cujos elementos estão sujeitos a uma transformação, dada aqui pela aplicação da função f. É sugestivo pensar nas sucessivas aplicações da função exactamente em termos temporais, considerando a órbita como um itinerário em que partimos do ponto x_0 e em que, ao fim de n segundos (ou horas ou dias...), nos encontramos no

[1] A definição da função f depende, de facto, do parâmetro β, pelo que, quando se revelar conveniente, podemos designá-la por f_β. Quando não houver dúvidas sobre o parâmetro em questão, seguimos a prática muito comum de não o incluir na notação. Esta observação aplica-se também, aliás, aos coeficientes e_n.

ponto x_n. O termo discreto refere-se exactamente ao facto de nesta interpretação o tempo ser discreto e não contínuo.

Uma forma de visualizar a órbita de um ponto x_0 passa por traçar, sobre o gráfico de f, a diagonal do quadrado $[0,1[\times[0,1[$ e determinar os pontos da órbita, começando no ponto $(x_0,0)$, subindo na vertical até encontrar o gráfico de f no ponto $(x_0, f(x_0))=(x_0,x_1)$, seguir depois na horizontal até encontrar a diagonal em (x_1,x_1), voltar a deslocar-se na vertical até ao ponto (x_1,x_2), e assim por diante.

Vejamos agora qual o papel dos coeficientes e_n. Aconselhamos o leitor a acompanhar o seu raciocínio com a observação do gráfico da função f para b=2 ou b=3, bem como para uma base não inteira, como a do exemplo apresentado anteriormente. Esse gráfico é constituído por $\lfloor\beta\rfloor$ segmentos de recta, com declive β; os extremos inferiores desses segmentos, todos sobre o eixo y=0, definem uma partição do intervalo $[0,1[$ em intervalos de monotonia de f. No caso de a base ser um inteiro b, $[0,1[=I_0\cup I_1\cup\ldots I_{b-1}$ onde $I_k=\left[\frac{k}{b},\frac{k+1}{b}\right[$, enquanto no caso de uma base não inteira β temos $[0,1[=I_0\cup I_1\cup\ldots I_{\lfloor\beta\rfloor}$ e os intervalos são definidos da mesma forma, com a única diferença de que $I_{\lfloor\beta\rfloor}=\left[\frac{\lfloor\beta\rfloor}{\beta},1\right[$.

O coeficiente e_n é determinado pelo intervalo de monotonia a que x_{n-1} pertence, concretamente, $e_n=k\Leftrightarrow x_{n-1}\in I_k$, e à órbita de x_0, constituída por números reais, corresponde uma outra órbita ou itinerário simbólico, dado pela sucessão de intervalos visitados, ou seja, pela sucessão $\{e_n\}$, e torna-se claro que o itinerário simbólico determina completamente o ponto correspondente x_0 e a sua órbita.

O papel da função f no intervalo é representado para o conjunto dos itinerários simbólicos por um operador deslocamento que designamos σ, e que actua numa sucessão deslocando um lugar para a esquerda os elementos da sucessão e eliminando o primeiro: se x tem o itinerário simbólico $\{e_1, e_2,\ldots\}$, então $f(x)$ tem o itinerário $\sigma(\{e_1, e_2,\ldots\})=\{e_2, e_3,\ldots\}$, o de $f^{(2)}(x)$ é $\sigma^{(2)}(\{e_1, e_2,\ldots\})=\{e_3, e_4,\ldots\}$, etc.

O conjunto das sucessões, munido do operador deslocamento, é usualmente designado por *subshift*. Esta correspondência entre órbitas constituídas por números reais e órbitas simbólicas está no cerne de uma das abordagens mais influentes no estudo dos sistemas dinâmicos, designada por dinâmica simbólica. No nosso caso, temos a propriedade muito particular de esta correspondência se traduzir nas igualdades $x=\sum_{k>1} e_k\beta^{-k}, f(x)=\sum_{k>1} e_{k+1}\beta^{-k}$, e assim por diante.

Tal como os números reais do intervalo $[0,1[$ estão ordenados, também as respectivas sucessões simbólicas o estão através da chamada ordem lexicográfica, que não é mais do que a generalização da ordem alfabética usual: dizemos que a sucessão $\{e_n\}$ é menor do que $\{e'_n\}$ se para o primeiro termo em que diferirem se tiver $e_m<e'_m$. E há uma equivalência entre as duas ordens: $x<x'$ se e só se os respectivos itinerários simbólicos $\{e_n\}$ e $\{e'_n\}$ estiverem na mesma relação.

Ainda uma última observação sobre a correspondência entre x_0 e a sua órbita, por um lado, e o respectivo itinerário simbólico, por outro. A sucessão $\{e_n\}$ indica-nos a sucessão de intervalos $I_{e_1},I_{e_2},\ldots$ sucessivamente visitados pela órbita, mas dá-nos também uma sucessão de intervalos $J_1\supset J_2\supset\ldots$ contendo o ponto inicial x_0: o primeiro termo e_1 indica-nos apenas que $x_0\in I_{e_1}=J_1$; mas $x_1=f(x_0)\in I_{e_2}$ implica que $x_0\in J_1\cap f^{-1}(I_{e_2})=J_2$; e_3 guarda a informação de que $x_2\in I_{e_2}$ mas também de que $x_0\in J_2\cap f^{-2}(I_{e_3})=J_3$.[2] Obtemos assim uma sucessão de intervalos encaixados, cuja intersecção é precisamente o ponto inicial da órbita. Observando o gráfico de f, vemos o que se passa com clareza: em cada passo, dividimos o intervalo J_n já encontrado em β subintervalos e o próximo coeficiente indica-nos qual deles é J_{n+1}. Por exemplo, a expansão na base 2 $\sqrt{2}-1=0\times2^{-1}+1\times2^{-2}+1\times2^{-3}+0\times2^{-4}+1\times2^{-5}+\ldots$ mostra-nos que $x_0=\sqrt{2}-1$ está no intervalo $J_1=I_0=[0,1/2[$, já que e_1=0, mas e_2=1 diz-nos, mais precisamente, que $x_0\in J_2=[1/4,1/2[$ e assim por diante.

Notemos agora a primeira diferença entre expansões em base inteira e não inteira: no caso de uma base inteira b, os intervalos de monotonia de f_b têm todos o mesmo comprimento e a imagem de qualquer um deles

[2] Deve ficar claro o significado dos expoentes negativos: $f^{-1}(I_{e_2})$ é a imagem inversa do intervalo pela função, ou seja, o conjunto $\{x\in[0,1[:f(x)\in I_{e_2}\}$; do mesmo modo, $f^{-2}(I_{e_3})$ é a imagem inversa deste intervalo pela função composta $f^{(2)}$.

pela função é todo o intervalo $[0,1[$; pelo contrário, no caso de uma base não inteira, o $I_{\lfloor\beta\rfloor}$ é menor do que os outros e, mais importante, $f_\beta(I_{\lfloor\beta\rfloor})=[0,\beta-\lfloor\beta\rfloor[$.

Esta distinção traduz-se numa diferença muito significativa entre expansões em bases inteiras e expansões em bases não inteiras. No primeiro caso, os coeficientes são *independentes* no seguinte sentido: sejam quais forem os valores de $e_1, e_2,\dots,e_n$, o coeficiente seguinte e_{n+1} pode assumir qualquer um dos valores admissíveis $0,1,\dots,b-1$.

Pelo contrário, como $\sum_{k\geq 1}\lfloor\beta\rfloor\beta^{-k}=\frac{\lfloor\beta\rfloor}{\beta-1}>1$, existe um natural N tal que $\sum^{N}_{k=1}\lfloor\beta\rfloor\beta^{-k}>1$, e portanto, em qualquer expansão-β, os primeiros N coeficientes não podem ser todos iguais a $\lfloor\beta\rfloor$; os coeficientes não são portanto independentes.

Esta observação, feita já por Rényi, coloca o problema de saber como caracterizar o conjunto, chamemos-lhe X_β, das sucessões $\{e_n\}$ que podem ser a sucessão de coeficientes da expansão-β de algum $x\in[0,1[$.

Comecemos por realçar que para uma base inteira b as expansões construídas nunca terminam com uma infinidade de b-1; este facto, que confirmaremos já adiante, será igualmente familiar no caso b=10: uma dízima infinita terminando por uma infinidade de noves pode ser representada por uma dízima finita (ou, se se preferir, terminando numa infinidade de zeros); por exemplo, 0,7659=0,766.

O mesmo é verdade para expansões em qualquer base inteira b: o "segredo" deste facto está na conhecida fórmula da soma de uma progressão geométrica, já usada no parágrafo anterior: $\sum_{k\geq 0} b^{-k}=\frac{b}{b-1}$. Daqui decorre que se $e_n\neq b-1$ e $e_k=b-1\ \forall k>n$, então $\sum_{k>n}(b-1)b^{-k}=(b-1)b^{-n-1}\sum_{k\geq 0} b^{-k}=b^{-n}$ e portanto
$\sum_{k\geq m} e_k b^{-k}=e_{-m}b^{m}+\dots+e_n b^{-n}+\sum_{k>n}(b-1)b^{-k}=e_{-m}b^{m}+\dots+(e_n+1)b^{-n}$.

O que se passa é que a construção da expansão que descrevemos atrás "evita" este tipo de ambiguidade, escolhendo sempre a expansão finita. Recorde-se, de facto, que para um dado ponto inicial x_0, se tem $x_0=\sum^{n}_{k=1}e_k b^{-k}+x_n b^{-n}$ onde $0\leq x_n<1$; mas, por outro lado, $x_n b^{-n}=\sum_{k>n+1} e_k b^{-k}$ e portanto, pelos cálculos feitos atrás, se os coeficientes nesta última soma fossem todos iguais a b-1, teríamos x_n=1.

Esta é de facto a única restrição: o conjunto X_b pode ser caracterizado pelas condições $\{e_n\}\in X_b$ se e só se

1. $0\leq e_n<b$ para todo o $n\geq 1$
2. $\{e_n\}$ não termina com uma infinidade de b-1.

Esta última condição pode ser enunciada com vantagem do seguinte modo: $\sigma^m(\{e_n\})<b\overline{-1}$, para todo o $m\geq 1$, onde a desigualdade deve ser naturalmente entendida no sentido da ordem lexicográfica já referida e $b\overline{-1}$ designa a sucessão constante.

Como vimos já, $\sum_{k\geq 1}(b-1)b^{-k}$ é a "falsa" expansão na base b de 1. Ela pode também ser vista como a expansão limite quando aproximamos 1 pela esquerda; por exemplo, a expansão na base b de $1-b^{-m}$ é, como o leitor facilmente verificará, $\sum^{m}_{k=1}(b-1)b^{-k}$.

Reencontramos um fenómeno semelhante no caso de uma base não inteira. Seja $\beta-\lfloor\beta\rfloor=\sum_{k>1} e'_k\beta^{-k}$ a expansão-β usual; o leitor verificará facilmente que temos a igualdade $1=\lfloor\beta\rfloor\beta^{-1}+\sum_{k>1} e'_k\beta^{-k-1}=\sum_{k>1} e_k\beta^{-k-1}$; caso esta expansão seja infinita, é ela a "falsa" expansão de 1; caso aquela expansão seja finita, é preciso substituí-la por uma outra: de facto, se se tiver $1=\sum^{m}_{k=1} e_k\beta^{-k}$, naturalmente com $e_m>0$, um cálculo simples envolvendo a soma de uma série geométrica com razão β^{-m}, dá-nos

$$1=\sum_{k\geq 1}\beta^{-km}\sum_{j=1}^{m-1} e_j\beta^{-j}+(e_m-1)\beta^{-m}$$

O que se passa é que, enquanto a expansão finita está associada à sucessão de coeficientes $e_1\dots e_m\bar{0}$, a expansão infinita está associada à sucessão $e_1\dots(\bar{e}_m-1)$, que é menor do que aquela na ordem lexicográfica.

O exemplo mais simples é o da raiz positiva β do polinómio z^2-z-1 : temos nesse caso, como facilmente se verifica,

$$1=\beta^{-1}+\beta^{-2}=\beta^{-1}+\beta^{-3}+\beta^{-5}+\dots=\sum_{k\geq 1}\beta^{-2k+1}$$

Vamos usar a notação especial $\sum_{k>1} c_k\beta^{-k}$ para esta "falsa" expansão de 1 na base β.

Pouco depois do artigo inicial de Rényi, o matemático inglês Bill Parry esclareceu completamente, entre outros aspectos, que sucessões de coeficientes podem ocorrer nas expansões-β; usando a notação anterior, $\{e_n\}\in X_\beta$ se e só se, para todo o $n\geq1$,

1. $0\leq e_n<\beta$
2. $\sigma^m(\{e_n\})<\{c_n\}$.

É claro que os casos em que a sucessão $\{c_n\}$ é eventualmente periódica, e especialmente quando é estritamente periódica, têm um interesse particular. Prova-se que a sucessão $\{c_n\}$ é estritamente periódica se e só se o conjunto X_β das sucessões de coeficientes das expansões-β é um "*subshift* de tipo finito" e que a sucessão $\{c_n\}$ é eventualmente periódica se e só se o conjunto X_β das sucessões de coeficientes das expansões-β é um "*subshift sofic*". Não podemos debruçar-nos aqui sobre a definição e as propriedades destes importantes sistemas dinâmicos simbólicos, que são, num certo sentido, os mais simples e portanto os que se prestam mais facilmente a um estudo detalhado.

É aqui que surge a interacção com as propriedades aritméticas da base β. Curiosamente, não são os números algebricamente mais simples, ou seja, os racionais, os que dão origem a expansões mais fáceis de caracterizar. No momento actual, parece quase tão difícil estudar as expansões-β para $\beta=\pi$ como para $\beta=3/2$! Para caracterizar as bases não inteiras mais simples temos de entrar, ainda que muito superficialmente, no terreno da Teoria dos Números Algébricos.

Um número β diz-se algébrico se for raiz de um polinómio com coeficientes racionais (em que podemos assumir que o coeficiente do termo de maior grau é 1). Entre todos os polinómios desta forma que têm β como raiz, existe um de grau mínimo, o chamado polinómio minimal de β. As restantes raízes (reais ou complexas) do polinómio minimal de β são os seus conjugados. Finalmente, se os coeficientes do polinómio minimal de β forem inteiros, este é um inteiro algébrico.

Um número de Perron é um inteiro algébrico β real maior do que 1 e maior em módulo do que todos os seus conjugados. Por exemplo, $\sqrt{2}$ é um inteiro algébrico mas não um número de Perron, enquanto $\frac{5+\sqrt{5}}{2}$ é um número de Perron.

Um número de Pisot é um número de Perron cujos conjugados têm todos módulo menor do que 1.

Ora, prova-se que se β é um número de Pisot, então a respectiva sucessão $\{c_n\}$ é (eventualmente ou puramente) periódica e portanto o conjunto X_β é de tipo finito ou, pelo menos, *sofic*.

No sentido inverso, prova-se que se a sucessão $\{c_n\}$ é eventualmente ou puramente periódica, então β é um número de Perron.

Mas existem números de Perron cuja sucessão não é eventualmente periódica, tal como existem outros que têm sucessão eventualmente ou até puramente periódica que não são números de Pisot. Em resumo, a caracterização algébrica ou aritmética das bases em relação à complexidade das expansões associadas é ainda, em parte, um problema em aberto.

Resta-nos mencionar um outro aspecto das expansões-β que se mantém igualmente como tema de intensa investigação: a caracterização, em função da base, dos números com expansão finita ou periódica.

No caso de uma base inteira b, o problema é elementar.

A expansão na base b de x_0 é finita se e só se x_0 é um racional $\frac{p}{q}$ tal que todos os factores primos de q dividem b. De facto, se a expansão for finita, temos $x_0=\sum_{k=1}^{n} e_k b^{-k}=\frac{p_n}{b^n}$, o que mostra que a condição é necessária. Reciprocamente, se $x_0=\frac{p}{q}$ e todos os factores primos de q dividem b, é evidentemente possível, multiplicando o numerador e o denominador por um factor apropriado, escrever $x_0=\frac{p'}{b^m}$. Como p' tem uma expansão finita na base b, $p'=\sum_{j=0}^{n} a_j b^j$, em que os coeficientes satisfazem $0\leq a_j<b$, temos $x_0=\sum_{j=0}^{n} a_j b^{j-m}=\sum_{k=m-n}^{m} e_k b^{-k}$, definindo $e_k=a_{m-k}$.

Por outro lado, como o leitor certamente sabe, pelo menos no caso b=10, a expansão de x_0 na base b é (eventualmente) periódica se e só se x_0 for racional.

Deixamos ao leitor a tarefa de verificar que uma expansão eventualmente periódica, numa base inteira b, corresponde a um número racional. Para provar a recíproca, chamamos a atenção para o facto de que se $x_0=\frac{p}{q}$ temos $b\frac{p}{q}=e_1+x_1$, ou seja, $bp=qe_1+qx_1$, o que implica que qx_1 é um inteiro e o mesmo se passa obviamente para todos os x_n; mas há apenas um número finito de números no intervalo [0,1[que satisfazem essa condição, pelo que a sucessão x_n tem de ser eventualmente periódica, o que implica que também a sucessão e_n seja eventualmente periódica.

Os problemas correspondentes para bases não inteiras são consideravelmente mais difíceis e os resultados mais significativos envolvem, mais uma vez, elementos da Teoria dos Números Algébricos.

Seja β um número algébrico. O conjunto dos números da forma $x=\sum_{j=0}^{m} q_j\beta^j$ constitui um corpo, designado por $\mathbb{Q}[\beta]$. Não é difícil deduzir que se $x=\sum_{k>1} e_k\beta^{-k}$ tem expansão-β periódica, então $x\in\mathbb{Q}[\beta]$.

Os matemáticos Anne Bertrand-Mathis e Klaus Schmidt demonstraram que a recíproca é verdadeira se β for um número de Pisot, ou seja, designando por $Per(\beta)$ o conjunto dos números $x\in[0,1[$ que têm expansão-β (eventualmente ou puramente) periódica, tem-se $Per(\beta)=\mathbb{Q}[\beta]\cap[0,1[$.

Esta propriedade não é exclusiva dos números de Pisot, mas Schmidt provou que se $\mathbb{Q}\cap[0,1[\subset Per(\beta)$, então β é necessariamente um número de Pisot ou de Salem.

Nos últimos anos, a investigação sobre as expansões-β e suas generalizações tem-se mantido extremamente activa através do trabalho de matemáticos como Anne-Bertrand Mathis, Christiane Frougny, David Boyd, Nikita Sidorov ou Shigeki Akiyama, entre outros. Os problemas relacionados com a caracterização das expansões periódicas têm sido, sem dúvida, dos que mais têm concentrado as atenções dos investigadores.

Apesar de as expansões-β serem objecto de estudo de numerosos artigos de investigação e um exemplo constantemente referido, quer no contexto da Teoria dos Sistemas Dinâmicos quer em certos aspectos da Teoria dos Números Algébricos, o autor destas linhas não conhece um livro que faça uma apresentação elementar e razoavelmente completa e actualizada destes interessantes objectos matemáticos. Deixamos, no entanto, ao leitor mais interessado algumas referências bibliográficas que podem servir de ponto de partida para o seu estudo. M

Bibliografia:

Blanchard, F., (1989). "β-expansions and Symbolic Dynamics", *Theoretical Computer Science*, 65, 131-141.

Lind, D., Marcus, B., (1995). "Symbolic Dynamics and Coding", *Cambridge University Press*.

Parry, W., (1960). "On the β-expansion of Real Numbers", *Acta Math. Hungar*, 11, 401-416.

Rényi. A., (1957). "Representations for Real Numbers and Their Ergodic Properties", *Acta Math. Hungar.*, 8, 477-493.

Schmidt, K., (1980). "On Periodic Expansions of Pisot Numbers and Salem Numbers", *Bull. London Math. Soc.* 12, no 4, 269-278.

Sidorov, N., (2003). "Arithmetic Dynamics, a survey paper", *Topics in Dynamics and Ergodic Theory*, LMS Notes Ser. 310, 145-189.