

Sobrevoando Cumes e Vales

PEDRO DUARTE E ALESSANDRO MARGHERI

UNIVERSIDADE DE LISBOA

pedromiguel.duarte@gmail.com, margheri@ptmat.fc.ul.pt

Recorrendo a analogias geográficas e à imagem dum planeta imaginário, procuramos explicar o conteúdo da Teoria de Morse, uma atraente teoria matemática do século XX que relaciona os pontos críticos duma função suave com a topologia do seu domínio.

Imagine a superfície acidentada dum planeta fictício, completamente submersa num oceano pacífico. Um dia, o nível da água do mar começa a descer, fazendo aparecer as primeiras ilhas. À medida que a terra se eleva acima do mar, são formadas baías, que depois viram lagos. Surgem istmos ligando terras próximas, penínsulas e continentes vão sendo descobertos pelas águas. A descida das águas faz com que lagos sequem, descobrindo vales profundos. Neste mundo todas as águas estão niveladas porque a terra é formada de matéria porosa. Entendendo “geografia”, em sentido restrito, como a simples contagem do número de ilhas ou continentes, por um lado, e do número de mares ou lagos, por outro, será possível compreender as mudanças que ocorrem na “geografia” do planeta em função da descida do mar? Uma teoria desenvolvida pelo matemático americano Marston Morse nos anos 30 do século XX responde a esta questão. A Teoria de Morse relaciona mudanças “geográficas” globais com certos pontos especiais, ditos *pontos críticos*, em torno dos quais ocorrem transformações qualitativas na “geografia”. Existem essencialmente três tipos de metamorfose causadas pelo abaixamento do nível das águas, que são: (1) o nascimento duma ilha, (2) a emergência dum istmo, ligando duas porções de terra e (3) o desaparecimento dum lago, porque o fundo dum vale ficou acima do nível do mar.

O objetivo principal deste artigo é o de mostrar algumas relações quantitativas entre o número de pontos críticos associados a estas metamorfoses (chamados, respetivamente, cumes, selas e poços) e o número de ilhas e de lagos do planeta. Provaremos estas relações considerando uma região suficientemente pequena do planeta fictício, de forma a po-

dermos considerar aceitável a sua aproximação com a superfície de um plano. Para nos convenceremos da correção desta aproximação local, é suficiente lembrar que, no caso da Terra, as civilizações mais antigas achavam que o nosso planeta, na realidade aproximadamente esférico, fosse mesmo plano (e não só as mais antigas, em meados do século XX, foi fundada uma sociedade, The Flat Earth Society, que defende que a Terra é plana!). Na parte final do trabalho, um pouco mais técnica, iremos explicar brevemente como uma das relações referidas acima se estende a contextos mais gerais. Em particular, consideraremos uma generalização que envolve a geografia de todo o planeta, esférico ou de outro tipo. De facto, o tipo de relação que se obtém neste contexto global depende da forma (tecnicamente, da topologia) do nosso planeta fictício, que poderá ser uma esfera, um *donut* (tecnicamente, um toro), ou ter outras formas ainda mais complicadas, como a de um oito “insuflado” (tecnicamente, um bitoro).

Começamos então por introduzir as notações e os conceitos que vão levar-nos ao nosso resultado principal.

Se quisermos descrever a geografia de uma pequena região do planeta, será suficiente considerar a função $f(x,y)$, definida no plano cartesiano \mathbb{R}^2 que mede a cota do lugar em função das suas coordenadas geográficas x e y . O gráfico da função $f(x,y)$ reproduz localmente a orografia do planeta. Seja $c \in \mathbb{R}$ um número real. O conjunto $N_c = \{ (x,y) \in \mathbb{R}^2 : f(x,y) = c \}$ representa a *curva de nível orográfica*, formada pelos pontos de cota c . Se c medir o nível do mar, N_c representa a linha de água, enquanto $M_c = \{ (x,y) \in \mathbb{R}^2 : f(x,y) \geq c \}$ representa a “terra” descoberta pelo oceano. É claro que o conjunto de nível N_c é o bordo da região M_c . Vamos agora enunciar as hipóteses sobre f .

(I) *Assumimos que a função $f(x,y)$ seja suave, isto é, que admita derivadas parciais de todas as ordens.* Do ponto de vista geométrico, isto implica, em particular, que existe um plano tangente ao gráfico de f por cada um dos seus pontos e que, ao variar do ponto sobre o gráfico, o correspondente plano tangente varia de uma forma contínua.

(II) *Assumimos que o conjunto $M_c = \{ (x,y) \in \mathbb{R}^2 : f(x,y) \geq c \}$ seja limitado para todo o $c \in [c_0, +\infty[$, ou seja, que a “geografia” é limitada para estes níveis do mar.*

(III) *Assumimos que f tem apenas um número finito de pontos críticos.* Chama-se *ponto crítico* de f a qualquer ponto em que o plano tangente ao gráfico de f seja horizontal. Os máximos

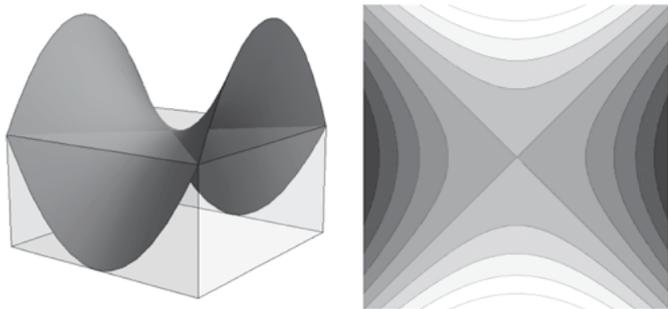


Figura 1: Um ponto de sela.

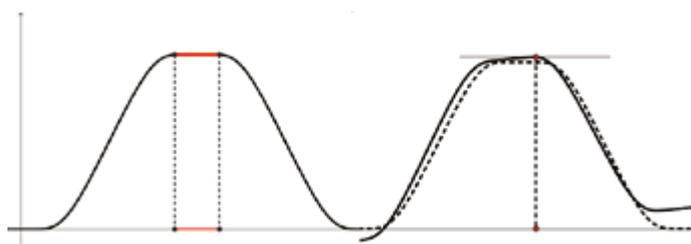


Figura 2: À esquerda, uma função com infinitos pontos críticos. À direita, uma sua pequena perturbação com um único ponto crítico.

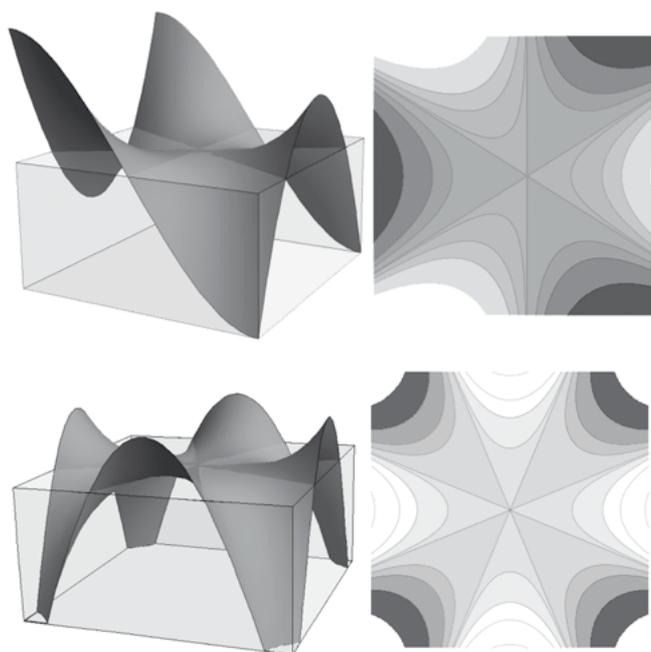


Figura 3: Duas selas degeneradas.

e mínimos locais são os exemplos mais óbvios de pontos críticos de f , que correspondem aos cumes de montanhas e aos fundos de vales. Nasce uma nova ilha em torno do ponto p (metamorfose 1) quando o nível das águas desce abaixo do valor $f(p)$ dum máximo local p . Analogamente, um lago desaparece num ponto q (metamorfose 3) sempre que o nível do mar desce abaixo do valor $f(q)$ dum mínimo local q . Chama-se *sela* a qualquer ponto crítico que não seja um máximo nem um mínimo local. Uma função cujo gráfico tenha a forma duma *sela de cavalo* tem um ponto crítico no centro da sela, que não é um máximo nem um mínimo local. Quando o nível do mar desce abaixo do nível crítico, forma-se um *istmo* (metamorfose 2). Veja a figura 1.

Uma função pode ter infinitos pontos críticos. Por exemplo, um *planalto* cujo topo seja perfeitamente horizontal corresponde a uma coleção infinita de pontos críticos. Esta situação é atípica. Qualquer perturbação, causada por um pequeno abalo sísmico, fará com que o planalto se incline muito ligeiramente e a função f passe a ter um único máximo local onde antes existia o planalto. Veja a figura 2.

(IV) *Assumimos que todos os pontos críticos de f são não degenerados.* Chama-se *sela degenerada* a uma sela onde três ou mais regiões se conectam num istmo múltiplo, logo que o mar desce abaixo do nível crítico. Veja dois exemplos na figura 3. Ao perturbar estas selas, pense num abalo sísmico provocando uma pequena perturbação aleatória no gráfico de f , duas ou mais selas (não degeneradas) aparecem no lugar da sela degenerada original. Veja na figura 4 exemplos de perturbações das selas da figura 3.

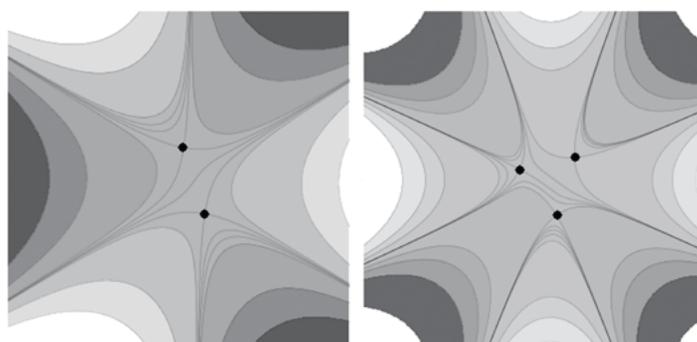


Figura 4: Perturbações de selas degeneradas.

Quanto aos máximos e mínimos não degenerados, sem entrar em pormenores demasiado técnicos, podemos dizer que são pontos críticos tais que, em torno deles, o gráfico de f tem a forma de uma superfície esférica em torno, respectivamente, do seu ponto com cota maior (polo norte) ou do seu ponto com cota menor (polo sul). Estes pontos são “robustos” no sentido que persistem para toda a pequena perturbação do gráfico de f .

As hipóteses (III) e (IV) são *genéricas*.

Uma propriedade genérica é uma propriedade que é partilhada por “quase todas as funções”. Intuitivamente, isto significa que se tivermos um saco que contém todas as possíveis funções suaves e extrairmos aleatoriamente uma função do saco, ela satisfará as hipóteses (III) e (IV).

Vamos chamar *cume* a um máximo local, *poço* a um mínimo local, e *sela* a um ponto crítico de tipo sela. Um nível $c \in \mathbb{R}$ diz-se *regular* se o conjunto de nível N_c não tiver pontos críticos. Imagine c como o nível do mar. Quando c é um nível regular, designamos por $\mu_0(M_c)$, resp. $\mu_1(M_c)$ e $\mu_2(M_c)$, o número de cumes, resp. selas e poços, em M_c . As *componentes conexas* de M_c são as regiões, as ilhas, em que M_c se decompõem. Analogamente, as componentes conexas de N_c são as “curvas” do conjunto de nível N_c . A figura 5 ilustra estes conceitos. Sejam agora

$$\beta_0(M_c) = \#(\text{componentes conexas de } M_c) = \# \text{ ilhas,}$$

$$\beta_1(M_c) = \#(\text{componentes conexas de } N_c) - 1 = \# \text{ lagos.}$$

Estamos agora prontos para enunciar e provar o resultado principal deste artigo.

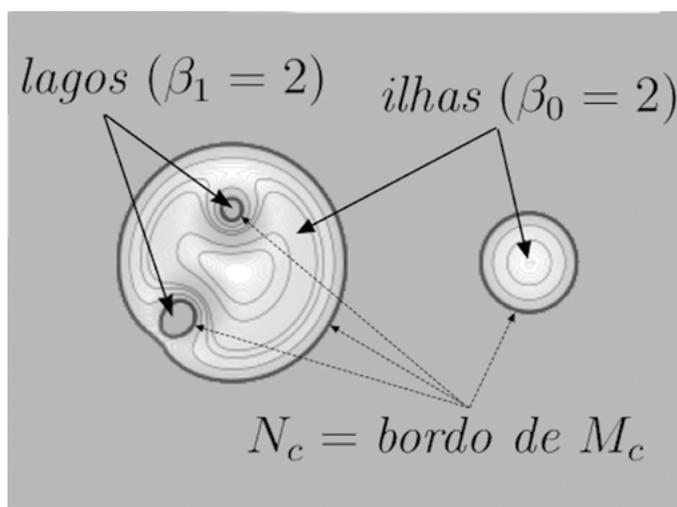


Figura 5: Ilhas e lagos.

Teorema Seja $f(x,y)$ uma função em \mathbb{R}^2 satisfazendo as hipóteses (I),(II),(III) e (IV). Então, para todo o nível regular $c \geq c_0$,

$$\begin{aligned} \beta_0(M_c) - \beta_1(M_c) &= \mu_0(M_c) - \mu_1(M_c) + \mu_2(M_c), \\ \mu_0(M_c) &\geq \beta_0(M_c), \\ \mu_1(M_c) &\geq \beta_1(M_c). \end{aligned} \quad (1)$$

A figura 6 e a tabela seguinte ilustram este resultado na emersão duma ilha com um cume, uma sela e um poço, desenrolada em três etapas: A, B e C.

	β_0	β_1	μ_0	μ_1	μ_2	$\beta_0 - \beta_1$	$\mu_0 - \mu_1 + \mu_2$
Etapa A	1	0	1	0	0	1	1
Etapa B	1	1	1	1	0	0	0
Etapa C	1	0	1	1	1	1	1

Vamos supor, para simplificar, que cada nível crítico de f tem um único ponto crítico. Esta é uma hipótese fácil de contornar, pelo que não é necessário assumi-la no enunciado do teorema.

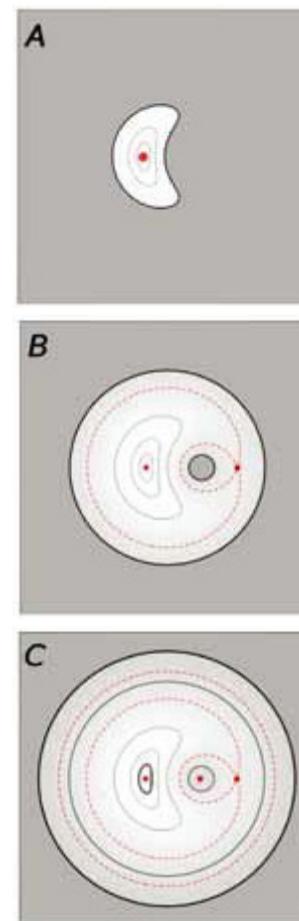


Figura 6: Emersão duma ilha.

Quando c decresce, as mudanças “geográficas” qualitativas correspondem a mudanças na “topologia” dos conjuntos M_c , que são detetadas por variações dos números $\beta_0(M_c)$ e $\beta_1(M_c)$. Estas mudanças ocorrem de cada vez que o nível do mar c desce abaixo dum nível crítico. Temos assim o “nascimento duma ilha” quando c passa pelo nível dum cume, temos a “emergência dum istmo” quando c cruza o nível duma sela, e o “desaparecimento dum lago” quando c desce abaixo do nível dum poço. Se agora virmos que em cada um destes três casos as relações (1) são preservadas, por um argumento indutivo elas terão de permanecer sempre válidas.

Nostrês casos que vamos analisar, c_0 designará um nível crítico, c e c' serão dois níveis regulares tais que $c > c_* > c'$, sendo c_* o único nível crítico entre c e c' . Os números $\beta_0, \beta_1, \mu_0, \mu_1$ e μ_2 reportam-se ao conjunto M_c , i.e., $\beta_0 = \beta_0(M_c)$, etc., enquanto os números $\beta'_0, \beta'_1, \mu'_0, \mu'_1$ e μ'_2 se referem ao conjunto $M_{c'}$, i.e., $\beta'_0 = \beta_0(M_{c'})$, etc. Vamos descrever os vários casos possíveis, deixando ao leitor a verificação em cada caso de que se $\beta_0, \beta_1, \mu_0, \mu_1, \mu_2$ satisfazem (1), então também $\beta'_0, \beta'_1, \mu'_0, \mu'_1, \mu'_2$ satisfazem (1).

No “nascimento duma ilha” aparece uma nova ilha mas nenhum lago novo, pelo que $(\beta'_0, \beta'_1) = (\beta_0, \beta_1) + (1, 0)$. Por outro lado, $M_{c'}$ contém apenas mais um cume do que M_c , pelo que $(\mu'_0, \mu'_1, \mu'_2) = (\mu_0, \mu_1, \mu_2) + (1, 0, 0)$.

No “desaparecimento dum lago”, o número de ilhas mantém-se, havendo apenas menos um lago, pelo que $(\beta'_0, \beta'_1) = (\beta_0, \beta_1) + (0, -1)$. Por outro lado, $M_{c'}$ contém apenas mais um poço do que M_c , pelo que

$$(\mu'_0, \mu'_1, \mu'_2) = (\mu_0, \mu_1, \mu_2) + (0, 0, 1).$$

Na “emergência dum istmo”, temos dois casos a considerar. Se o istmo liga duas ilhas distintas, temos $(\beta'_0, \beta'_1) = (\beta_0, \beta_1) + (-1, 0)$. Como $M_{c'}$ contém apenas mais uma sela do que M_c , segue que $(\mu'_0, \mu'_1, \mu'_2) = (\mu_0, \mu_1, \mu_2) + (0, 1, 0)$. Se o istmo liga dois cabos da mesma ilha, temos a formação dum lago, mantendo-se inalterado o número de ilhas. Pode também acontecer que o istmo ligue dois cabos dum lago ou mar interior, que assim se divide ao meio. Em qualquer caso, o número de ilhas mantém-se e o número de lagos vem acrescido dum unidade, pelo que $(\beta'_0, \beta'_1) = (\beta_0, \beta_1) + (0, 1)$. Como $M_{c'}$ contém apenas mais uma sela do que M_c , segue que $(\mu'_0, \mu'_1, \mu'_2) = (\mu_0, \mu_1, \mu_2) + (0, 1, 0)$.

□

O resultado anterior lembra a conhecida *fórmula de Euler* para poliedros convexos, $(F - A + V = 2)$, onde $F = \#$ faces, $V = \#$ arestas, $V = \#$ vértices do poliedro), o que não é por mera coincidência. Ambas as fórmulas descrevem a “topologia” dum certo objeto: dos poliedros no caso da fórmula de Euler, e da região emersa M_c no caso da fórmula de Morse. O número 2 é o valor da *característica de Euler* de todos os poliedros convexos, um importante invariante topológico que reflete o facto de estes objetos, se os imaginarmos de borracha, poderem ser todos deformados de forma contínua numa esfera. Os números $\beta_0(M_c)$ e $\beta_1(M_c)$ são também invariantes topológicos de M_c , conhecidos como os *números de Betti* de M_c .

O teorema enunciado pode ser generalizado a contextos mais gerais. Em particular, podemos substituir o plano \mathbb{R}^2 por uma superfície compacta, o que nos permite passar de um resultado essencialmente local a um que tem em conta a topologia da totalidade da superfície. O adjetivo “compacta” refere-se ao facto de a superfície ser limitada e de conter todos os seus pontos de fronteira (isto é, os pontos que podem ser aproximados tanto quanto se queira por pontos da superfície). Pense-se, por exemplo, na superfície de uma esfera, ou na superfície de um *donut*, ou na de um oito “insuflado”, ou, ainda, num disco com a circunferência do bordo incluída. Analogamente ao feito acima quando descrevemos a orografia local do planeta através de uma função definida num plano, podemos descrever a sua orografia global utilizando uma função que mede a cota dos seus pontos relativamente à superfície que aproxima a forma média do astro (no caso da Terra, esta superfície será uma esfera cujo raio é o raio médio do nosso planeta).

A primeira relação do teorema enunciado permanece válida para funções definidas em superfícies compactas se na fórmula usarmos $\beta_0 - \beta_1 + \beta_2$ em vez de $\beta_0 - \beta_1$. No caso em que a superfície M é compacta e orientável¹, os números de Betti para uma podem ser caracterizados do modo seguinte: $\beta_0(M)$ é o número de componentes conexas de M . $\beta_1(M)$ é o maior número de “cortes” que podemos fazer sem dividir M em duas componentes separadas. Cada “corte” deve ser efetuado ao longo dum curva simples (isto é, sem autointersecções) fechada, ou dum curva simples aberta entre dois pontos do bordo de M . Finalmente, $\beta_2(M) = 1$ se M não tiver bordo, e $\beta_2(M) = 0$ no caso contrário. Se M for um disco ou um retângulo, então $\beta_1(M) = 0$, porque todo o corte divide M em duas partes. Se M for o toro, então $\beta_1(M) = 2$ porque cortando M ao longo dos dois círculos ortogonais indicados na figura 7 obtemos um “retângulo”, que não pode mais ser cortado sem ser separado. Imagine agora um planeta fictício na forma dum toro, com uma orografia à sua escolha, e procure verificar a validade do teorema.

O teorema é ainda válido em superfícies de dimensão finita arbitrária, e mesmo de dimensão infinita, sob hipóteses adequadas. Na sua essência, a Teoria de Morse mostra-nos como a topologia dum superfície condiciona, e é ao mesmo tempo condicionada, pelo número e o tipo de pontos críticos dum função suave sobre ela definida. A Teoria de Morse, e



Figura 7: Dissecando um toro.

as suas múltiplas generalizações enquadram-se e expandem o chamado *Cálculo das Variações* com um arsenal de novas técnicas. No seu trabalho [1], além de estabelecer as bases desta teoria, M. Morse faz uma bela aplicação à resolução dum problema da Geometria Riemanniana. Os trabalhos de M. Morse tiveram um impacto profundo numa boa parte da matemática do século XX, em áreas como a Topologia, a Geometria, as Equações Diferenciais e os Sistemas Dinâmicos. A Teoria de Morse influenciou de forma decisiva o trabalho de Stephen Smale, nomeadamente em Sistemas Dinâmicos, como a Teoria dos Sistemas Morse-Smale, e a Teoria Hiperbólica, e na sua demonstração da Conjectura de Poincaré Generalizada para dimensão ≥ 5 , publicada em 1960, e que até então se acreditava ser falsa em dimensões superiores. Como curiosidade, referimos que, segundo o próprio Smale, uma parte desta conjectura foi demonstrada na praia de Copacabana, no Rio de Janeiro. Pode ler mais sobre a vida e o trabalho de Smale, e as suas ligações à escola brasileira de Sistemas Dinâmicos, em [2]. A versão final da Conjectura de Poincaré foi finalmente demonstrada em 2002 por Grigori Perelman para a dimensão 3. O livro [3] de John Milnor contém uma excelente exposição do trabalho de M. Morse. No artigo de *survey* [4], R. Bott, que fez importantes generalizações da Teoria de Morse, faz uma bela síntese desta teoria. O nível das duas referências anteriores deve ser considerado avançado. Recomendamos o vídeo [5] editado pela MAA, com duas lições do próprio Morse sobre a sua teoria, a primeira das quais a um nível bastante elementar.

REFERÊNCIAS

- [1] M. Morse, *The Foundations of a Theory of the Calculus of Variations in the Large*, Trans. Am. Math. Soc., (1929)
- [2] G. Szpiro, *A Conjectura de Poincaré, cem anos para resolver um dos maiores problemas da matemática*, Gradiva (2007).

[3] J. Milnor, “Morse Theory”, *Annals of Mathematic Studies* AM-51, (1963), Princeton University Press.

[4] R. Bott, “Lectures on Morse Theory, Old and New”, *Bull. Amer. Math. Soc.* (N.S.) Volume 7, Number 2 (1982), 331-358.

[5] M. Morse, *Pits, Peaks and Passes*, Mathematical Association of America, vídeo (1967).

[6] S. Smale, “Differentiable Dynamical System”, *Bull. Am. Math. Soc.* 73 (1967), 747- 817.

¹ Este último termo significa que um caminheiro que percorra qualquer rota que inicie e acabe no mesmo ponto da superfície, não volta ao ponto de partida de cabeça para baixo. Exemplos de superfícies não orientáveis são a fita de Möbius e a garrafa de Klein.

SOBRE OS AUTORES

Pedro M. Duarte licenciou-se em Matemática pela Faculdade de Ciências da Universidade de Lisboa, em 1985, e obteve o seu doutoramento em 1993, na área dos Sistemas Dinâmicos, pelo Instituto de Matemática Pura e Aplicada no Rio de Janeiro. Atualmente é professor auxiliar no DM da FCUL.

Alessandro Margheri licenciou-se em Matemática pela Facoltà di Scienze dell'Università degli Studi di Firenze (Itália) em 1988. Desde o ano 2000 é professor auxiliar do DM da FCUL. A sua área de especialização é a de Equações Diferenciais Ordinárias.