

Scripta-Ingenia

On ternary computation systems



Fractal Science Kit, Borromean-Rings

Recent discoveries show the existence of three stable states which can be induced and measured in a new material called bigraphene. This suggests again an old idea, that of building ternary computer systems. The importance and relevance of constructing ternary computers, is justified by a list of facts showing that the number three is preferable to the number two:

- i) From the point of view of complexity and information theory, the best possible base for a number system is $e = 2.71\dots$, the Neper number, which is best approximated by the integer number three.
- ii) There is an alternative to the standard ASCII code. Indeed using a trit of trits, that is a sequence of three letters in the alphabet $\{0, 1, 2\}$, there are 27 possible such combinations which can be coded by the 26 letters of the standard English alphabet plus one extra space character or control key. With four trits we have 81 possibilities that can be used to code the small letters, the capital letters and the digits plus the most common symbols, but some studies are required to achieve a standard.
- iii) There is also the possibility of variable length coding for binary words, this is convenient for storing and transmitting information, and it is also related to cryptography.
- iv) Ternary logic: in order to compare two numbers in a binary system we always have to perform two steps; in ternary logic only one step is required.
- v) Menu systems: the most efficient way of organizing a menu system is in the form of a trinary tree.
- vi) The Cantor set: in mathematics, special in topology, the cantor set is an important example of a compact space; its points are the real numbers in the unit interval $[0,1]$ for which the ternary expansion does not use the digit.

A ternary computer (also called trinary computer) is a computer that uses ternary logic (three possible values) instead of the more common binary logic (two possible values) in its calculations. One of the earliest calculating machines, built by Thomas Fowler entirely from wood in 1840, was a ternary computer. The only modern, electronic ternary computer Setun was built in the late 1950s in the Soviet Union at the Moscow State University by Nikolay Brusentsov, and it had notable advantages over the binary computers which eventually replaced it (such as lower electricity consumption and lower production cost). In 1970 Brusentsov built an enhanced version of the computer, which he called Setun-70. Ternary computing is commonly implemented in terms of balanced ternary, which uses the three digits -1, 0, and +1.

— Wikipedia

A Scripta-Ingenia assume-se como uma revista de divulgação científica tratando temas da ciência e da tecnologia, cobrindo todas as áreas do saber no domínio das ciências exactas ou aplicadas. Interessa-se ainda por artigos de opinião, sobre tópicos científicos ou não, desde que escritos por autores na área das ciências e da engenharia, e que reflitam as suas opiniões enquanto membros dessa comunidade. A revista tem uma periodicidade semestral e os seus números são libertados na mesma data dos solstícios. Este é o segundo número e corresponde ao Solstício de Verão.

Chief Editor — Nelson Martins-Ferreira
CDRSP-ESTG, IPLEiria



Wikipedia:Picture of the day/September 10, 2005

Stonehenge is a Neolithic and Bronze Age monument located near Amesbury in the English county of Wiltshire. Constructed between 2500 BC and 2000 BC, it is composed of earthworks surrounding a circular setting of large standing stones, known as megaliths. Archaeoastronomers claim that Stonehenge represents an “ancient observatory”, with significant alignments for the sunrise on the solstice and equinox days.

Whatever you dream on this night will come to pass. Have great dreams on June 21! And if you are in the northern hemisphere, enjoy the longest day of the year.

William Shakespeare, *in A Midsummer’s Night Dream*

Processamento de Imagem Digital e Cancro de Pele:

uma abordagem interdisciplinar

by RUI FONSECA-PINTO

IPL – Instituto Politécnico de Leiria
IT – Instituto de Telecomunicações - Leiria

1 Introdução: Cancro e Epidemiologia

Anualmente a 4 de fevereiro, a marcar a aprovação da carta de Paris aprovada nesse mesmo dia no ano 2000 na cimeira mundial contra o cancro no novo milénio, celebra-se o Dia Mundial do Cancro. Esta carta apela à cooperação entre investigadores, profissionais de saúde, indústria e governos no âmbito da prevenção e tratamento do cancro. Os dados epidemiológicos apontam para números que em Portugal ascendem a 70 mortes por dia, o que em média significa 3 mortes a cada hora devido ao cancro. Por ano morrem no nosso país cerca de 25 000 pessoas. De entre os vários tipos de cancro, os que apre-

sentam maior incidência são o cancro coloretal, do pulmão, do estômago, da bexiga e os linfomas não-Hodgkin. Os que apresentam maior mortalidade são os cancros do pulmão, o cancro coloretal, do estômago, do pâncreas e do fígado. Os números relativos a novos casos/número de mortes em Portugal encontram-se na Figura 1, onde se podem verificar os números de novos casos e de incidência por género. No sexo masculino é o cancro da próstata, do pulmão e coloretal que apresentam números mais elevados, ao passo que nas mulheres, o cancro da mama e coloretal são os cancros que se destacam tanto em incidência quanto em mortalidade. De referir ainda que a nível de incidência, o cancro da próstata e da mama, respetivamente nos homens e mulheres, apresentam números muito elevados [1 e 2]. (ver Figura 1)

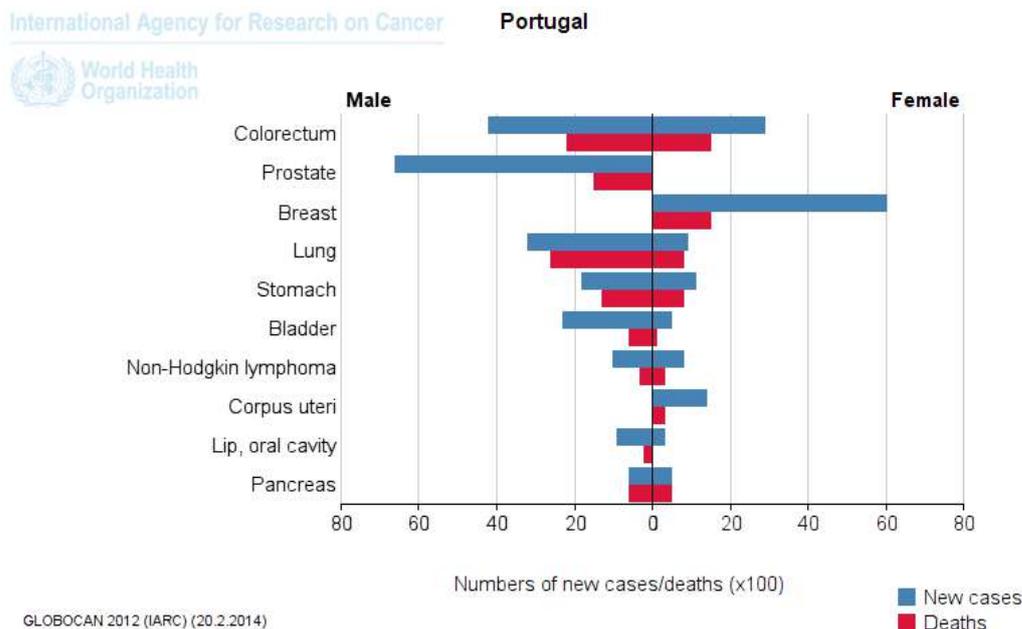


Fig. 1: Mortalidade/novos casos de cancro em Portugal (fonte: Globocan 2012)

Em todo o mundo o cancro é uma das principais causas de morte, estimando-se que entre 2005 e 2015, seja responsável por mais de 84 milhões de mortes. Estima-se ainda que o número de casos de cancro e mortes relacionadas a nível mundial venha a duplicar nos

próximos 20 a 40 anos. Anualmente os número de mortes devido a cancro é superior ao número de mortes por tuberculose, malária e VIH/SIDA. Em todo o mundo os cancros com maior incidência são o dos pulmões, em ambos os géneros, seguido do cancro da próstata e cancro co-

loretal nos homens, e de cancro da mama e coloretal nas mulheres [1 e 2].

Se olharmos para os números relativos à Europa, os dados por género apresentam um comportamento similar no que diz respeito ao nível dos tipos de cancro mais incidentes, mas neste caso aparecem agora na lista dos

10 tipos mais incidentes, e dos que provocam maior mortalidade os cancros do colo do útero, do pâncreas, do rim e o melanoma cutâneo. Segundo a OMS cerca de 40% do total dos cancros podem ser prevenidos ou detetados numa fase precoce do seu desenvolvimento.

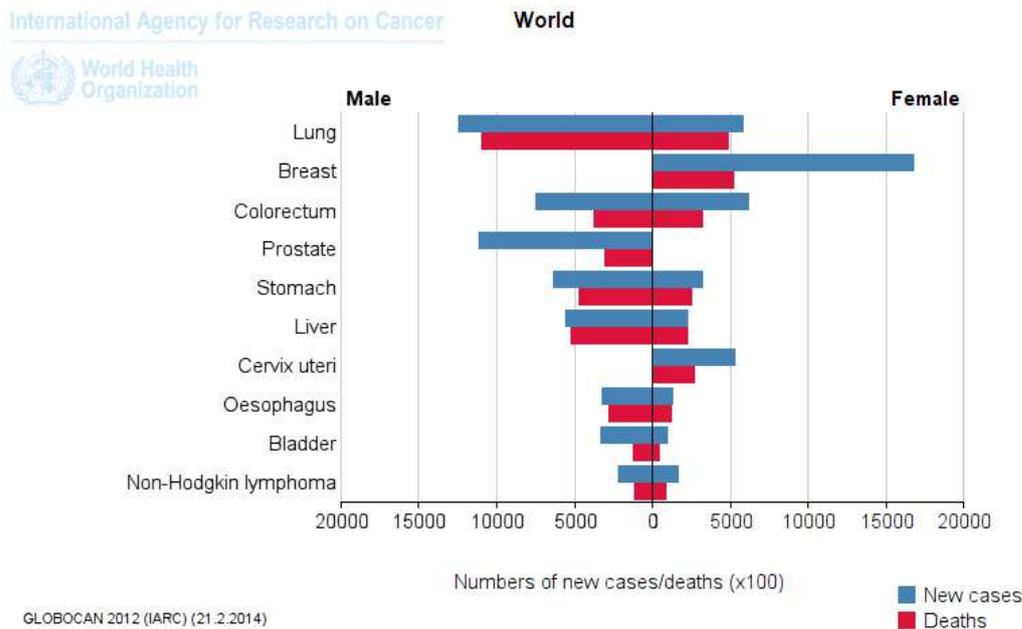


Fig. 2: Mortalidade/novos casos de cancro no mundo (fonte: Globocan 2012)

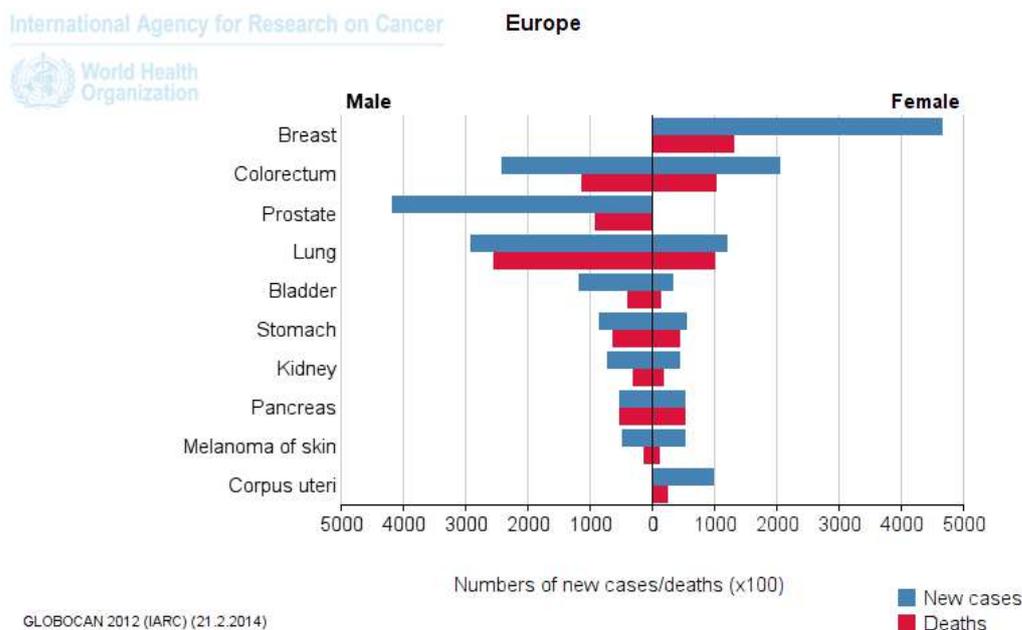


Fig. 3: Mortalidade/novos casos de cancro na Europa (fonte: Globocan 2012)

O cancro possui uma história natural de evolução que começa pelo desenvolvimento de células malignas (na grande maioria dos casos não se percebe bem qual o motivo pelo qual isto acontece) que não são destruídas pelo sistema imunitário, e que durante um período de tempo se mantêm em desenvolvimento sem dar sinais, até que comecem a manifestar-se através de sinais e sintomas num estado mais avançado. Há alguns tipos de cancro relativamente aos quais são conhecidos os mecanismos da carcinogénese (quer seja por exemplo genética ou através de um vírus) e por esse motivo é possível mapear a sua evolução e prever o seu desenvolvimento. Os casos do cancro da mama, colo do útero e do cólon são um bom exemplo de como hoje em dia é possível atuar antes da fase clínica com grandes sucessos. Nestes como noutros tipos de cancros a deteção precoce é mesmo o fa-

tor primordial. Infelizmente há casos em que as manifestações clínicas só aparecem numa fase muito avançada da doença. Ainda assim, temos hoje ao dispor toda uma possibilidade de quimioterapia antineoplásica e de métodos de irradiação ionizante (radioterapia) que são ferramentas usadas em complemento com a cirurgia cito-reductora sempre que possível. Tal como já referido, um dos tipos de cancro em que o diagnóstico precoce é um fator essencial para a cura é o melanoma cutâneo. Os números relativos ao melanoma no mundo destacam a Austrália (34.9 ASR ¹), Nova Zelândia (35.84 ASR), a América do Norte (14.3 ASR) com taxas muito elevadas. Em Portugal a taxa de incidência normalizada para a idade – ASR, é de 6.66 ASR [1 e 2]. Os dados relativos à incidência e mortalidade nos países do sul da Europa encontram-se no gráfico seguinte.

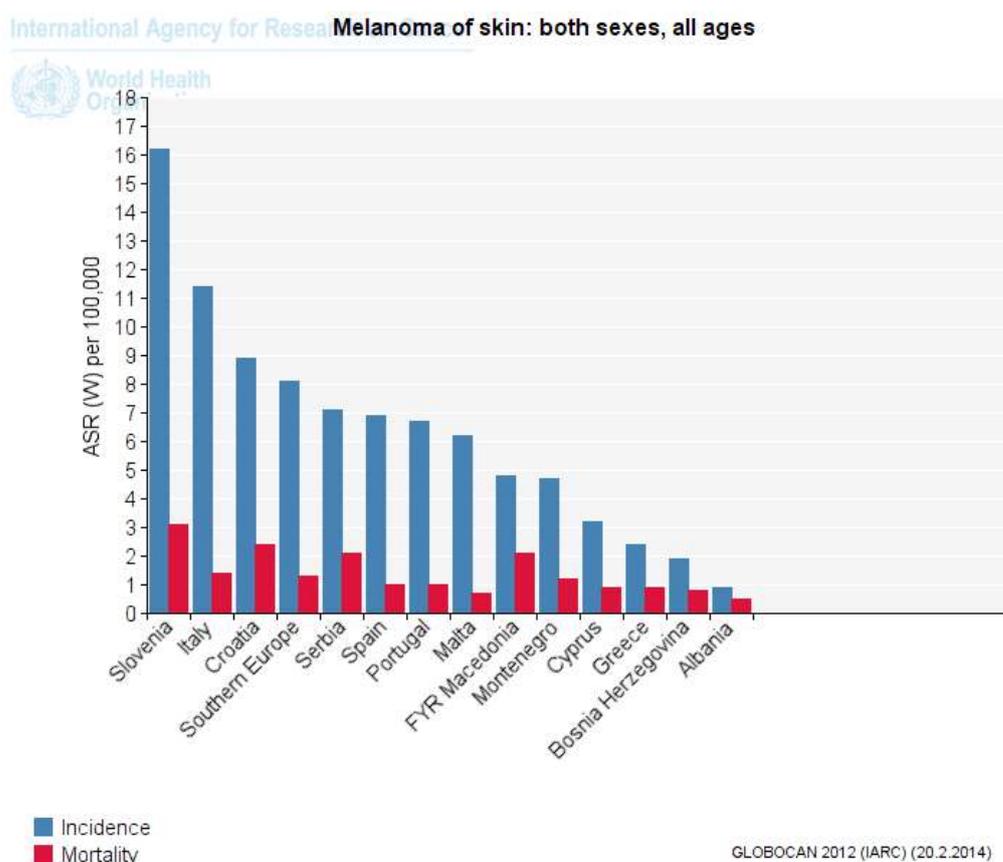


Fig. 4: Mortalidade/novos casos de melanoma em países da Europa (fonte: Globocan 2012)

2 O melanoma cutâneo: Aspectos clínicos e diagnóstico

Nos indivíduos caucasianos o cancro de pele é o tipo de cancro mais frequente e a sua incidência tem vindo a aumentar desde os anos 60. Se observarmos um corte histológico de pele podemos observar que este se divide

em 3 camadas principais ilustradas na Figura 5 (epiderme, derme e hipoderme). A camada mais superficial – epiderme, encontra-se subdividida ainda em outras 5 camadas (córnea, lúcida, granulosa, espinhosa e basal). (ver Figura 6) Queratinócitos, melanócitos, células de Langerhans e células de Merkel são as células que se encontram na epiderme, cada qual com uma função associ-

¹Incidência ASR (W)– Age Standardised Rate Incidence, número de novos casos num dado período, em termos de idade normalizada, e é expresso num determinado número por 100000

ada. No caso dos melanócitos estes são células especializadas, presentes sobretudo na camada basal, que podem sintetizar um pigmento escuro, a melanina, cuja função é a de proteger a pele contra os raios ultravioletas do sol, pelo que a exposição à luz solar provoca a produção de maiores quantidades de melanina.

O cancro de pele pode classificar-se em três tipos:

Carcinoma baso-celular (BCC), Carcinoma epino-celular (SCC) e o melanoma maligno. Na literatura usualmente também é referida uma tipologia binária distinguindo-se apenas os melanomas dos não-melanomas (BCC e SCC). O Carcinoma baso-celular é o tipo de cancro de pele mais vulgar e é assim designado porque tem origem nas células da camada basal da epiderme (ver Figuras 5 e 6).

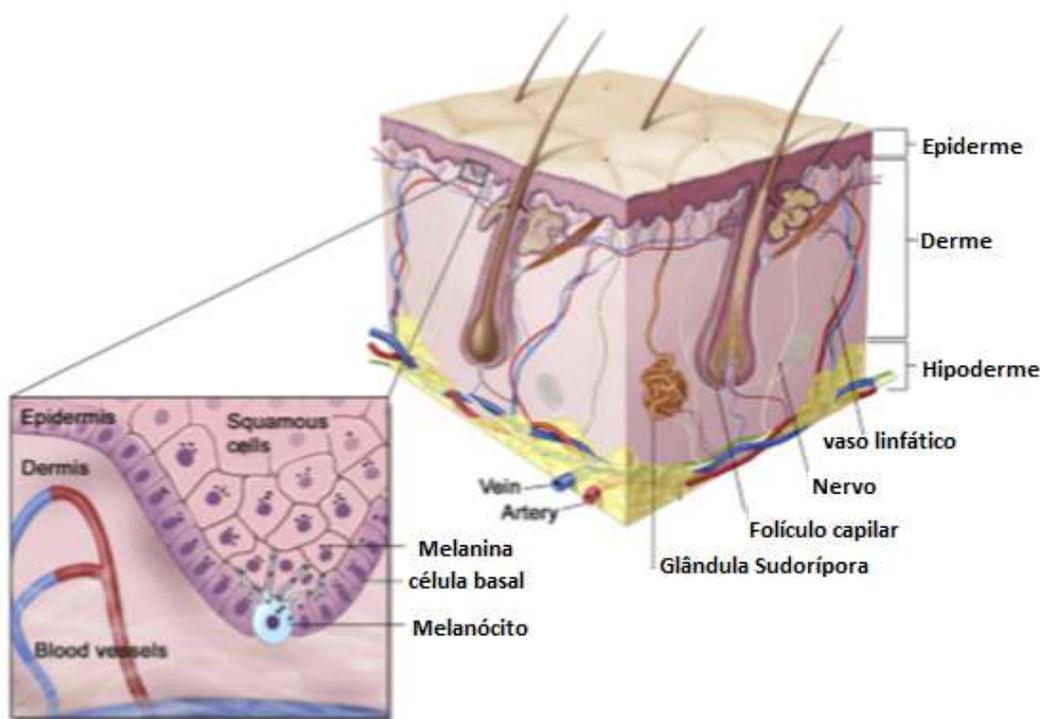


Fig. 5: Anatomia da pele e alguns dos seus constituintes funcionais (adaptado de [3])

Atinge preferencialmente caucasianos que estão expostos de forma crónica ao sol, de que são exemplo no nosso país os trabalhadores agrícolas, da construção civil e pescadores entre outros. Pode manifestar-se sob a forma de um nódulo rosado/brilhante ou fenda superficial de crescimento lento, que surge na face, pescoço ou região dorsal e que com o tempo não desaparece. Trata-se de um tumor que apenas é localmente invasivo pelo que não dá metástases à distância, sendo o tratamento na fase inicial feito através de cirurgia ou laser. Se não for tratado atempadamente pode tornar-se muito agressivo e invade os tecidos circundantes o que provoca grandes defeitos e mutilações por exemplo no nariz, pavilhões auriculares e pálpebras. Mesmo em fases avançadas é possível recorrer à cirurgia e radioterapia, apresentando estas abordagens uma elevada taxa de cura. O Carcinoma epino-celular tem origem nas células das camadas intermédias do epitélio pavimentoso estratificado da epiderme, também designada camada espinhosa. Trata-se do segundo tipo de cancro de pele mais frequente. Tal como no BCC, está associado a grupos profissionais com elevada exposição solar, mas neste caso a faixa etária

é mais elevada que no caso anterior, caso em que a pele já está mais envelhecida e em que se trata de uma evolução de uma lesão pré-existente. Manifesta-se habitualmente sob a forma de um nódulo, de crescimento rápido, com tendência para ulcerar e sangrar facilmente e é mais agressivo que o BCC. Neste caso é possível que desenvolva metástases à distância, mas quando diagnosticado e tratado a tempo, apresenta elevadas probabilidades de cura. No caso do melanoma, trata-se do tipo de cancro de pele mais perigoso e um dos mais agressivos no homem atingindo grupos etários mais jovens que os outros dois tipos de cancro de pele já referidos. Tem origem nos melanócitos da epiderme, células responsáveis pela produção de melanina (pigmento que dá a cor à pele). Pode também ocorrer nos olhos, meninges, aparelho digestivo gânglios linfáticos ou noutras zonas onde há melanócitos. Ao contrário dos tumores anteriores, que estão relacionados com a exposição crónica ao sol, o melanoma maligno parece estar mais associado à exposição solar intermitente, aguda e intempestiva, muitas vezes acompanhada de queimaduras solares, conhecidos “escaldões solares” a que as campanhas de

prevenção do cancro da pele fazem referência nos tempos de praia. O aspeto inicial do melanoma é variado mas habitualmente caracteriza-se pelo aparecimento de um pequeno nódulo ou mancha, de cor negra tipo alcatrão, sobre pele aparentemente sã ou sobre um sinal preexistente. Tem uma fase precoce de crescimento superficial (fase de crescimento horizontal) que pode durar meses, ou até anos, e uma fase mais tardia de crescimento

em profundidade (fase de crescimento vertical). O tratamento é quase sempre cirúrgico, e quando efetuado nas fases iniciais em que o tumor ainda é muito fino (espessura inferior a 0,5mm), acompanha-se de elevadas taxas de cura. Todavia quando o tumor já é muito espesso (espessura superior a 4mm) as probabilidades de cura ficam drasticamente reduzidas, existindo o risco eminente de metastização à distância.

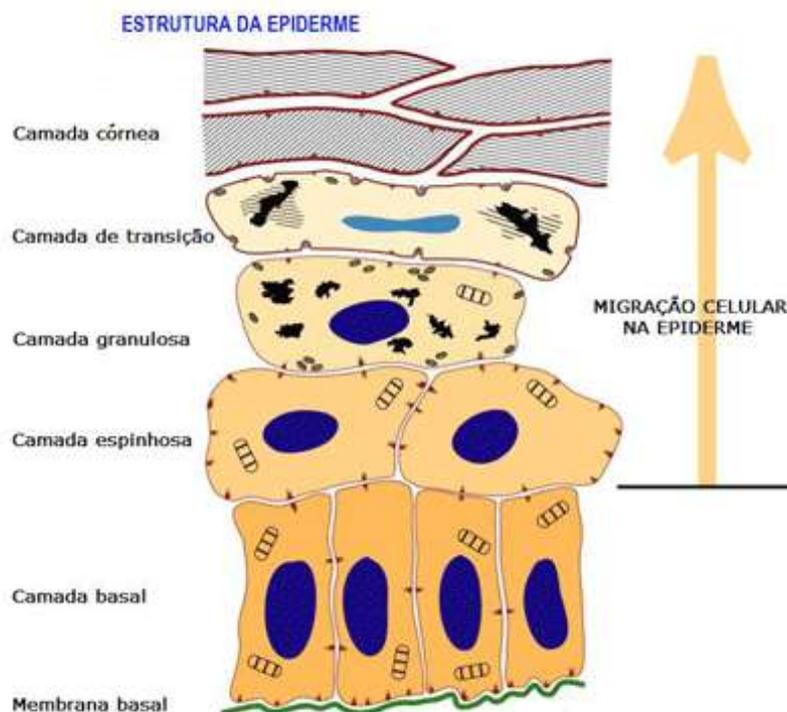


Fig. 6: Camadas da epiderme (adaptado de [3])

O diagnóstico do melanoma é feito após excisão da lesão e é portanto um diagnóstico histológico. Dado que as lesões malignas têm inicialmente um crescimento horizontal cujas características foram sendo descritas ao longo de anos de experiência de observação de lesões cutâneas, o desenvolvimento de instrumentação que possibilitasse este acompanhamento precocemente foi sendo inicialmente transferido da histologia para a observação microscópica cutânea. Remonta a 1663 o uso da microscopia de superfície por Kolhaus e mais tarde por Ernest Abbe em 1878 que usou a microscopia de superfície com imersão em óleo. O dermatologista Alemão Johann Saphier juntou ao instrumento anteriormente desenvolvido uma fonte de luz. Foi Goldman o primeiro dermatologista a usar a expressão dermoscopia e a usar um dermatoscópio para avaliar lesão pigmentadas da pele. Em 2001 o fabricante californiano 3Gen produziu o primeiro dermatoscópio que usa luz polarizada o Dermlite. O dermoscópico é pois um instrumento usado por dermatologistas que permite amplificar a imagem da superfície da

pele, e que através do uso de uma fonte de luz possibilita ainda perceber alterações ao nível da epiderme. Trata-se de uma técnica de diagnóstico não invasiva para observação in vivo da morfologia e de estruturas que de outra forma não eram observáveis a olho nu. Esta técnica envolve o uso de um instrumento ótico com aumento de 6 a 10 vezes. Cerca de 4% a 7% da luz incidente é refletida devido à superfície seca da pele, o que limita a visualização de estruturas mais profunda. Por forma a reduzir esta reflexão da luz um fluido é usado entre o dermoscópico e a pele. Mais recentemente são também usados dermoscópico com uma fonte de luz polarizada. A imagem aumentada da superfície cutânea é guardada em formato digital para posterior processamento e acompanhamento da sua evolução com o tempo. O diagnóstico diferencial das lesões de pele é uma tarefa desafiadora, até para dermatologistas experientes. O CNMD (Consensus Net Meeting on Dermoscopy) acordou uma estratégia de duas etapas para o diagnóstico em dermoscopia, onde os avanços da tecnologia e a imagem digital têm tido um papel crucial

para o diagnóstico precoce do cancro de pele. O objetivo da primeira etapa do diagnóstico diferencial numa lesão cutânea é a de perceber se a lesão é melanocítica ou não. A decisão é tomada com base na presença de características como redes pigmentadas, glóbulos, pigmentação não homogénea, padrões vasculares entre outros. Caso a lesão seja do tipo melanocítica, esta é posteriormente classificada de benigna, suspeita ou maligna. Existem alguns algoritmos para ajudar na diferenciação entre lesões benignas e malignas, sendo os três mais conhecidos a regra ABCD, a lista dos 7 pontos (7-point check-

list) ou o método de Menzies. Na figura 7 encontram-se 8 lesões melanocíticas, as 4 apresentadas na linha superior (a) – (d) são benignas e as 4 da linha inferior (e) – (h) são malignas.

A regra ABCD foi proposta em 1994 por Wilhelm Stolz. É hoje o padrão utilizado nas análises dermatoscópicas para a classificação das lesões melanocíticas. As características analisadas pela metodologia fundamentam-se em identificadores como: assimetria da lesão (A); irregularidade das fronteiras (B); variabilidade de cores existentes (C) e estruturas diferenciais (D).

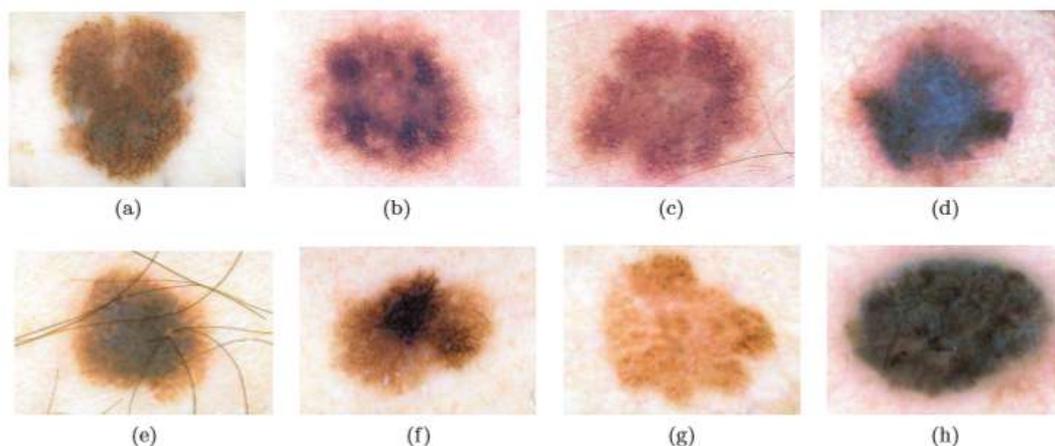


Fig. 7: Exemplos de imagens de dermoscopia. As quatro imagens da linha superior são benignas e as restantes 4 são malignas (adaptado de [3])

Como resultado da verificação destes atributos obtém-se o potencial melanocítico da lesão, denominado Dermatoscopy Point Value (DPV). O DPV tem por finalidade auxiliar o dermatologista durante o processo de diagnóstico de lesões. A classificação das lesões de acordo com o DPV é feita através do resultado do cálculo. Um valor de DPV inferior a 4.75 dá uma classificação benigna. Entre 4.75 e 5.45 é um classificador para lesão suspeita e entre 5.45 e 8.9 uma lesão potencialmente maligna. No primeiro caso sugere-se um acompanhamento clínico, no segundo acompanhamento clínico ou remoção cirúrgica e no terceiro caso, excisão cirúrgica e diagnóstico anátomo-patológico. Esta classificação segue as orientações publicadas em [4]. Mais recentemente foi também introduzida a letra E na regras ABCD, que significa evolução. Trata-se de verificar a evolução ou mudanças que podem ocorrer como por exemplo alterações sensitivas, sangramento, conformação, e novas lesões. A análise de padrões é a forma clássica para o diagnóstico diferencial, que é baseada em critérios qualitativos, em que numa primeira etapa é avaliado o aspeto geral e depois as características locais. Existem 4 tipos de padrões pré-definidos (reticular, globular, homogéneo e estrelar). O padrão reticular é o padrão global mais comum em lesões melanocíticas. É caracterizado por uma rede pigmentar que cobre a maior parte de uma lesão.

A rede pigmentar aparece como uma grade de segmentos alinhados (semelhantes a uma colmeia) em diferentes tonalidades de preto, castanho ou cinza. As modificações da rede pigmentar variam de acordo com as mudanças no comportamento biológico das lesões cutâneas melanocíticas, portanto, essas variações merecem atenção especial. No padrão globular, estruturas castanhas de tamanhos variados, redondas a ovais, preenchem as referidas lesões melanocíticas. Este padrão pode ser encontrado em nevos melanocíticos congénitos e adquiridos e de Clark (displásicos). O padrão homogéneo é caracterizado por uma cor sem estrutura, uniforme e difusa, que preenche a maior parte da lesão. As cores incluem preto, castanho, cinza, azul, branco ou vermelho. Uma cor predominantemente azulada é a marca morfológica do nevo azul. O Padrão estrelar (starburst, “explosão de estrelas”) é caracterizado pela presença de estrias pigmentadas e/ou pontos e glóbulos num arranjo radial na periferia de uma lesão melanocítica. Este padrão é o estereótipo morfológico dos nevos de Spitz. A lista dos 7 pontos (7-point checklist) é outra das formas usada para fazer a classificação, tratando-se de uma forma semiquantitativa. São usadas sete características da imagem divididas em critérios major (com peso de 2) e minor (com peso de 1) tendo sendo esta soma o valor máximo de 7. Se forem obtidos menos de 3 pontos a lesão é consider-

ada benigna, no caso de serem 3 ou mais pontos a lesão tem uma probabilidade elevada de ser um melanoma. [5, 6, 7]. Um exemplo desta classificação encontram-se na

Figura 8, na qual podemos observar a pontuação máxima.

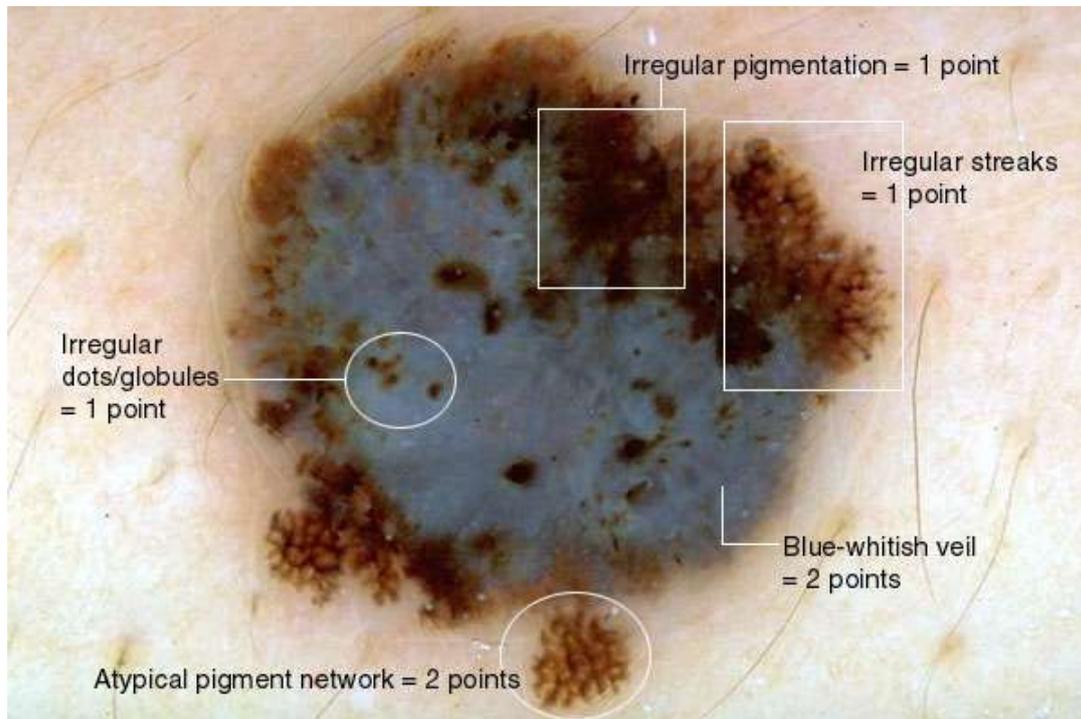


Fig. 8: Exemplo de aplicação da escala de 7 pontos. Adaptado de [2]

O método de Menzies é outra forma qualitativa de avaliação em que onze características, divididas em dois grupos (positivas e negativas) são avaliadas, sendo classificadas como presentes ou ausentes. As características negativas são a simetria de cor e a presença de cor única. Estas duas características definem as lesões como benignas. A simetria é definida como simetria de cores e/ou estruturas presentes na lesão segundo eixos que passem no centro de gravidade da lesão (o que não obriga à simetria da mesma). Estão ainda classificadas nove características positivas que incluem múltiplos pontos castanhos, zonas azuis-brancas, redes pigmentadas, etc. De acordo com esta classificação, uma lesão com pelo menos uma característica positiva e sem qualquer característica negativa é considerada melanoma.

3 Investigação Aplicada e Dermoscopia

Nos últimos anos tem-se assistido a um aumento significativo da investigação na área do cancro de pele e em particular nas técnicas de diagnóstico não invasivas que têm como ponto de partida a dermoscopia. Desde logo os primeiros avanços foram no próprio equipamento, que hoje já permite o registo e o envio de dados remotamente

sem necessidade de fios ou de cartões de memória. Além da evolução na instrumentação, a grande contribuição tem vindo do processamento da imagem digital. Se olharmos para as publicações que nos últimos anos se têm feito nesta área do cancro da pele e deteção precoce podemos verificar que os artigos se dividem em duas classes de acordo com a sua natureza (artigos clínicos e artigos de visão computacional) e de acordo com o número de lesões analisadas (lesão única ou múltiplas). Os artigos clínicos geralmente estão mais relacionados com questões de investigação de doenças dermatológicas com auxílio da imagem, estudos clínicos e estudos de caso, tendo como público-alvo médicos. Estas publicações constituem aproximadamente 24% das publicações da área entre 1984 e 2012.

Os artigos de visão computacional têm como público-alvo investigadores na área da visão computacional e peritos de processamento de imagem, e contêm assuntos mais técnicos do processamento de imagem e mais relacionados com os algoritmos, constituindo 76% do total das publicações da área relativamente ao mesmo período. Dos artigos que se têm dedicado às questões do processamento de lesão única, 28% são de deteção de fronteiras, 29% de identificação de estruturas da lesão, 19% de diagnóstico automático computacional, e 16% de classificação automática de estruturas, sendo os restantes classificados na classe outras investigações. A informação que

se pode obter das imagens obtidas por dermoscopia depende em larga escala de um bom algoritmo para se realizar a segmentação. Muitas das vezes, existe ainda uma etapa anterior às anteriormente referidas, que se trata de um pré-processamento que visa essencialmente remoção de artefactos (cabelos e bolhas de ar) e realce da imagem (correção de cor e de iluminação entre outras). Estas técnicas de pré-processamento aliadas a algoritmos de processamento de imagem permitem obter a segmentação da lesão com boa qualidade para as etapas seguintes.



Fig. 9: Imagem de dermoscopia após segmentação da lesão para seguimento da evolução da mesma

Em seguida apresentam-se dois exemplos de imagens de dermoscopia, com os limites da lesão identificados de

forma automática. Nos casos apresentados nas figuras 9 e 10 a presença de artefactos (cabelos e bolhas de ar) exigem um pré-processamento e técnicas de segmentação especificamente desenhadas para estas imagens.

Além dos marcadores clássicos já identificados, a interdisciplinaridade que tem vindo a fazer parte das equipas de investigação tem possibilitado que novos marcadores temporais característicos das lesões tumorais sejam mapeados. Neste sentido, as classificações das lesões estão a evoluir no sentido de se tornarem mais quantitativas. Nas mais recentes publicações têm sido reportados marcadores de evolução cuja génese estão nas teorias das séries temporais, dos sistemas dinâmicos e nos sistemas de equações diferenciais através da modelação matemática do crescimento e das características das lesões tumorais. Encontra-se a decorrer no Instituto de Telecomunicações, unidade de investigação da qual faz parte o Instituto Politécnico de Leiria, um projeto com financiamento europeu, dedicado ao processamento de imagem de dermoscopia de onde foram extraídas as imagens usadas neste artigo e cujos primeiros resultados podem ser encontrados em [9,10].



Fig. 10: Imagem de dermoscopia após segmentação da lesão para seguimento da evolução da mesma

Referências

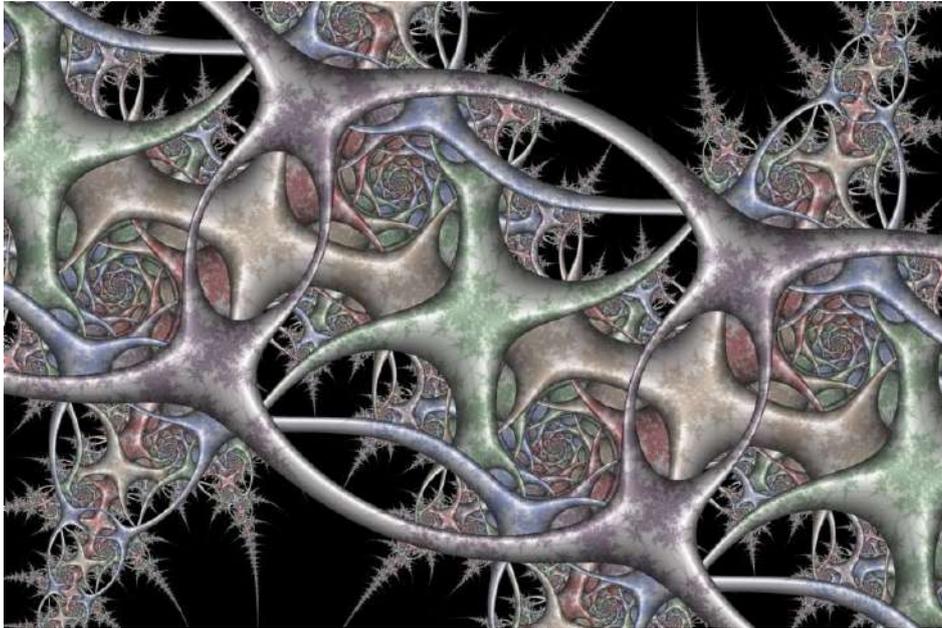
- [1] Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, “Cancer Incidence and Mortality Worldwide: IARC CancerBase” No. 11. Available from: <http://globocan.iarc.fr>, accessed on day/month/year.
- [2] Bray F, Ren JS, Masuyer E, Ferlay J. “Estimates of global cancer prevalence for 27 sites in the adult population in 2008”. *Int J Cancer*. 2013 Mar 1;132(5):1133-45, 2008.
- [3] G. Argenziano, H. Soyer, V. D. Giorgio, D. Piccolo, P. Carli, M. Delfino, A. Ferrari, R. Hofmann-Wellenhof, D. Massi, G. Mazzocchetti, M. Scalvenzi, and I. H. Wolf, “Dermoscopy, an interactive atlas.” EDRA Medical Publishing, 2000. [ONLINE]. Available at: <http://www.dermoscopy.org>.
- [4] BINDER. M, Schwarz M, Winkler A, Steiner A, Kaider A, Wolff K, Pehamberger H. “Epiluminescence microscopy, a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists”, *Arch Dermatol*,131:286-291, 1995.
- [5] G. C. do Carmo and M. R. e Silva, “Dermoscopy: basic concepts,” *International Journal of Dermatology*, vol. 47, pp. 712–719, July 2008.
- [6] R. P. Braun, H. S. Rabinovitz, M. Oliviero, A. W. Kopf, and J. H. Saurat, “Dermoscopy of pigmented skin lesions,” *American Academy of Dermatology*, vol. 55, no. 1, pp. 109–121, 2005.
- [7] R. H. Johr, “Dermoscopy: Alternative melanocytic algorithms - the abcd rule of dermatoscopy, menzies scoring

method, and 7-point checklist”, Clinics in Dermatology, vol. 20, no. 3, pp. 240–247, 2002.

[8] Fonseca-Pinto R, Caseiro P, A. Andrade, “Bi-dimensional Empirical Mode Decomposition (BEMD) in dermoscopic images: artefact removal and border lesion detection”, Proceedings of the 7th IASTED international Conference Signal Processing, Pattern recognition and Applications; 341-345, 2010.

[9] Pereira J, Mendes A, Nogueira C, Baptista D, Fonseca-pinto R, “An adaptive approach for skin lesion segmentation in dermoscopy images using a multiscale Local Normalization”, Proceedings of the CIM series in Mathematical Sciences, Springer-Verlang, 2014. (to appear)

[10] Mendes A, Nogueira C, Pereira J, Baptista D, Fonseca-Pinto R, “Growth model in melanocytic skin lesions: a geometric look to skin pigmented lesions”, Proceedings of 2nd International Congress Ipleiria Health: Challenges & Innovations in Health, Rev Saúde Publica, Vol 48, pp 288; USP 2014.



Fractal Science Kit, Fractal-Circle-Orbit-Trap-06



Fractal-Steiner-Chain-Orbit-Trap-18



Fractal-Mobius-Dragon-IFS-04

Some examples of images generated by The Fractal Science Kit. This fractal generator is a Windows program to generate a mathematical object called a fractal. The term fractal was coined by Benoit Mandelbrot in 1975 in his book *Fractals: Form, Chance, and Dimension*. In 1979, while studying the Julia set, Mandelbrot discovered what is now called the Mandelbrot set and inspired a generation of mathematicians and computer programmers in the study of fractals and fractal geometry.

DRUG DELIVERY SYSTEMS – A Brief Approach

by MARGARIDA CABRITA FRANCO

IPL – Instituto Politécnico de Leiria
CDRSP - Centro Para o Desenvolvimento Rápido e
Sustentado do Produto

1 Drug delivery systems

Although drug delivery may not attract the attention and press coverage of some of the other disciplines involved in drug discovery, it is an equally important issue in the development of a successful product. Besides the choice of the active drug substance (active pharmaceutical ingredient - API), a responsible decision regarding the route of administration and dosage form needs to be made. A wrong choice can cause failure of therapy. Dosage forms may be classified into:

- Conventional release;
- Modified release which can be delayed, controlled or sustained.

The conventional release dosage forms are designed to release the drug promptly after administration of a bo-

lus for immediate release. In contrast, modified release dosage forms are designed to modulate drug release, delaying or prolonging the dissolution. Various technologies may be employed to promote the gradual release of a drug from a pharmaceutical form.

In the conventional therapy, aliquot quantities of a drug are introduced into the system at specified time intervals resulting in a considerable fluctuation on drug concentration levels. However, an ideal dosage regimen would be the one in which the concentration of the drug is maintained between the limits of the therapeutic window at a constant level throughout all the treatment period. Both situations can be graphically represented in Figure 1.

The basic goal of this kind of therapy is to achieve a steady state blood level that is therapeutically effective and non-toxic for an extended period of time. The design of a proper dosage regimen is an important element when we want to accomplish this objective. An excellent drug delivery system provides the right amount of drug or its active metabolite to the target, at a time and rate that optimizes its effect, without compromising patient safety.

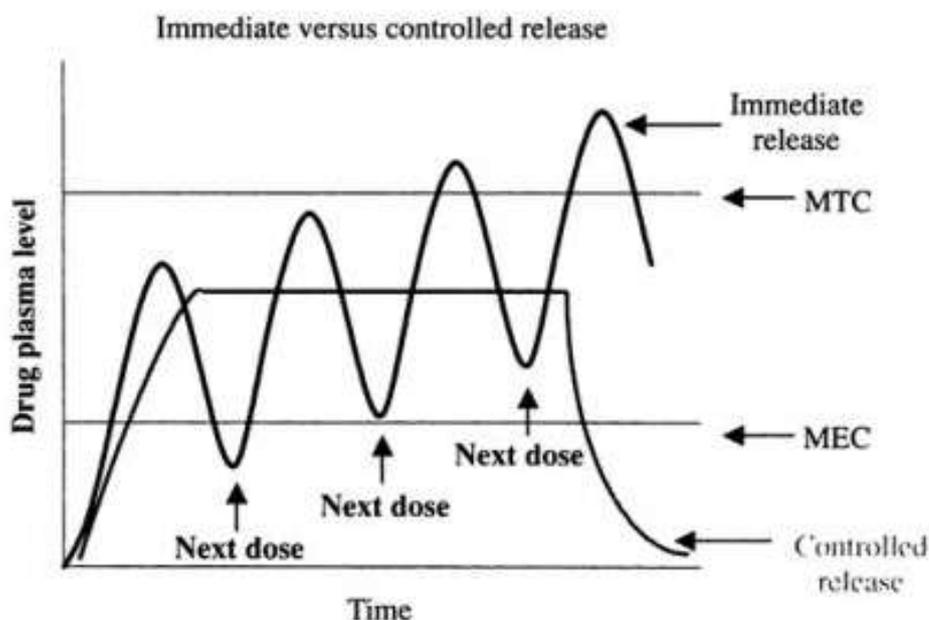


Fig. 1 - Immediate versus controlled release dosage forms.
(MEC-Minimum effective concentration; MTC-Minimum toxic concentration)

in Steve I. Shen, Bhaskara R. Jasti, and Xiaoling Li, Standard Handbook of Biomedical Engineering and Design, Myer Kutz (ed), Chapter 22: Design of Controlled-Release Drug Delivery Systems, Copyright The McGraw-Hill Companies, Inc. 2003

Controlling release to meet these requirements can involve preventing release or delivery in enzymatically or other hostile environments, targeting a specific location to enhance therapeutic index, or prolonging delivery to sustain an effect. Delivery on demand may also be required with biosensors linked to a delivery device like a drug reservoir, for example. Pros and cons of controlled Drug Delivery Systems (DDS) can be listed as follows:



- Improved patient convenience and compliance due to less frequent drug administration.
- Reduction in fluctuation achieving a steady-state level and therefore better control of disease condition.
- Increased safety margin of the drug due to better control of plasma levels.
- Maximum utilization of drug enabling reduction in total amount of dose administered.
- Reduction in health care cost through improved therapy and shorter treatment period.
- Less frequency of dosing and reduction in personnel time to dispense, administer and monitor patients.
- Better control of drug absorption can be obtained, since the high blood level peaks that may be observed after administration of a single dose corresponding to a high bioavailability drug can be reduced.



- Decreased systemic availability in comparison to immediate release conventional dosage forms; this may be due to incomplete release or insufficient residence time for complete release, for example.
- Poor in-vivo/in-vitro correlation.
- Possibility of dose dumping due to food interaction with formulation and thus increased risk of toxicity.
- Retrieval of drug is difficult in case of toxicity, poisoning or hypersensitivity reaction.
- The physician has less flexibility in adjusting dosage regimens. This is fixed by the dosage form design.
- Economic factors must also be assessed, since more costly processes and equipment are involved in manufacturing many of the sustained release forms.

Medication delivery to the systemic circulation through the skin is an alternative to oral and parenteral routes of administration. The first Transdermal Drug Delivery Systems (TDDS) or drug loaded patches entered the U.S. market in 1979 with the approval of the scopolamine patch by the Food and Drug Administration (FDA) for treatment of motion sickness. Today there are more than twenty different types of medication patches available, and researchers are constantly striving to develop additional and improved methods for transdermal delivery of medications. The great advantage lays on the fact that it avoids first pass metabolism reactions which occur in the liver and gut wall resulting in a decrease on drug bioavailability. TDDS (transdermal patches or transdermal drug delivery systems) are designed to show a perfect adhesion onto healthy and clean skin in order to assure controlled drug delivery into the systemic circulation. Problems linked with factors that affect percutaneous absorption must be studied and solved. This is the purpose of several research teams working in universities all over the world. Actually those problems are connected with the following questions:

- Drug physicochemical properties: solubility, molecular weight, partition coefficient;
- Stratum corneum characteristics like thickness and lipophilicity;
- Concentration of the drug;
- Area of the applied field;
- Drug affinity for the skin.

Some promoters of percutaneous absorption have been suggested. Chemical permeation enhancers, iontophoresis, sonophoresis and even micro or nano needles are good examples. Different types of delivery systems involve different technologies to sustain release and can be listed as shown below:

- Matrix systems: the matrix can be hydrophilic or hydrophobic and drug release processes can involve swelling of the polymer in use, drug diffusion and matrix erosion;
- Reservoir systems;
- Osmotic pumps;
- Other systems – micro or nano sized needle arrays.

The improvement and innovation in the development of controlled release systems strongly depend on select-

ing an appropriate agent capable of controlling drug release. The ideal agents for these preparations should ensure that the drug is released at the right place in the right dose and at (or during) the required time. In addition, it must be said that with the advances in the field of chronobiology, modern drug delivery approaches have been elevated to a new concept of chronopharmacology, which is the ability to deliver the therapeutic agent to a patient in a pulsatile profile. New technology for medications delivering in a time-modulated fashion, by bedside or ambulatory pumps, is being developed to manage human diseases, mainly chronic diseases.

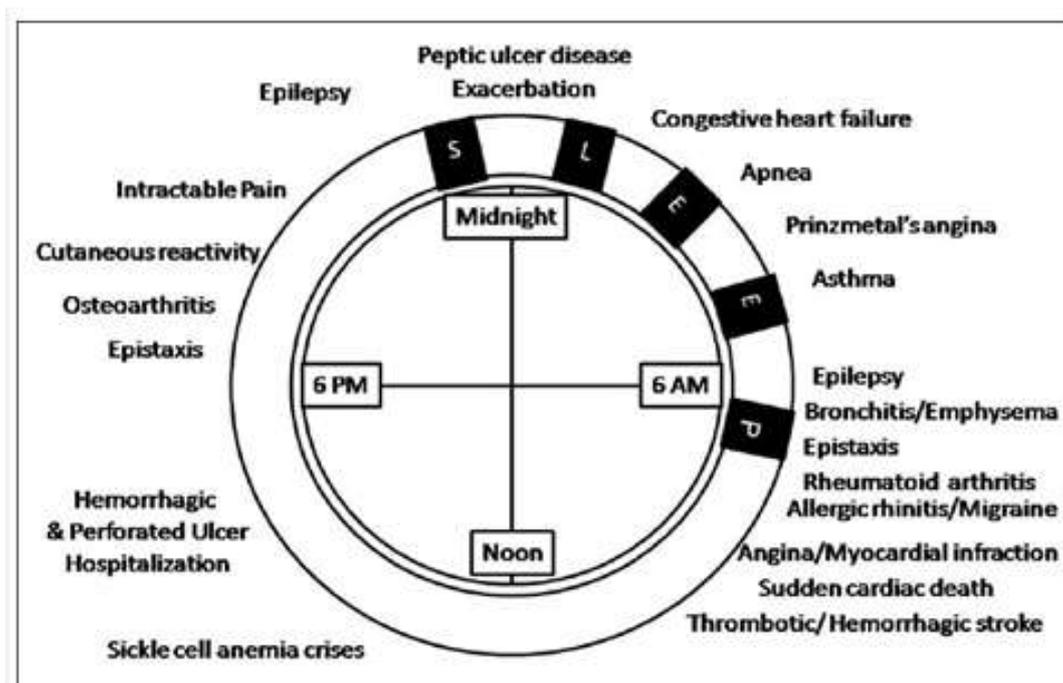


Fig 2 - Time cycle when the diseases show their maximum effect

in Bisht R., Chronomodulated drug delivery system: A comprehensive review on the recent advances in a new sub-discipline of 'chronopharmaceutics'. Asian J Pharm; 2011; 5:1-8

2 Controlled Drug Delivery at CDRSP

In our research center, we are trying to "build" a model in which levodopa (3,4-Dihydroxy-L-phenylalanine) is entrapped inside a sodium alginate membrane for skin administration with a controlled mechanism of delivery. It is expected that transdermal route would provide a supply of the drug to the blood stream without fluctuations

on therapeutic levels. This fact is very important in order to avoid the typical on-off effects related to the therapeutics of Parkinson Disease, the disorder in which levodopa is still the gold standard. In our labs, we produce and characterize our drug loaded membranes; then, we carry out the necessary dissolution assays required by the Pharmacopoeias and recommended by the drug national and international authorities. Finally, specific mathematical models are applied in order to understand the drug release mechanism from the polymeric matrix.

References

[1] Ale I, Lachapelle JM, Maibach HI. Skin tolerability associated with transdermal drug delivery systems: an overview. Adv Ther. 2009; 26:920-35

- [2] Bisht R. Chronomodulated drug delivery system: A comprehensive review on the recent advances in a new sub-discipline of 'chronopharmaceutics'. *Asian J Pharm* 2011; 5:1-8
- [3] Hong I, Gabay M, Ladolce A. Safety concerns involving transdermal patches and magnetic resonance imaging (MRI). *Hosp Pharm.* 2010; 45:771-8
- [4] Goole J., Amighi K., "Levodopa delivery systems for the treatment of Parkinson's disease: an overview", *International Journal of Pharmaceutics.* 2009; 380:1-15
- [5] Luigi G. Martini, Patrick J. Crowley, Sarah Ibrahim, *Controlling Release from Dosage Forms: The Future*, American Pharmaceutical Review, 2012, June 25
- [6] Scheindlin S. Transdermal drug delivery: past, present, future. *Mol Interv.* 2004; 4:308-12
- [7] Steve I. Shen, Bhaskara R. Jasti, and Xiaoling Li, *Standard Handbook of Biomedical Engineering and Design*, Myer Kutz (ed), Chapter 22: Design of Controlled-Release Drug Delivery Systems, Copyright The McGraw-Hill Companies, Inc. 2003
- [8] Tanner T., Marks R. Delivering drugs by the transdermal route: review and comment. *Skin Res Technol.* 2008; 14:249-60



PONTIAC DEBUTS 2007 SOLSTICE GXP AT LA AUTO SHOW
The GXP version of the Solstice debuted at the Los Angeles Auto Show in January 2006.

The Pontiac Solstice Coupes are considered to be quite rare: There were a total of 1,266 Solstice Coupes that were able to be manufactured before the production line in Wilmington, Delaware was shut down: 102 pre-production 2009 models, 1,152 sequentially-vin'd regular production 2009 models, and 12 pre-production 2010 models. This is in contrast to over 64,000 of the Pontiac Solstice Convertibles that were manufactured.

— Wikipedia

Sobre o cálculo dos ângulos de sopé e de convergência para veículos automóveis

by MIGUEL BELBUT GASPAR, NELSON MARTINS FERREIRA E ARTUR MATEUS

ESTG, CDRSP,
Instituto Politécnico de Leiria

1 Introdução

A inspeção periódica obrigatória de veículos automóveis verifica, entre outras coisas, os ângulos de convergência e de sopé para cada uma das quatro rodas de um veículo em inspeção. A medição desses valores é feita por máquinas que requerem elas próprias, com uma certa regularidade, a sua calibração de modo a garantir que os valores que estão a medir têm associado um erro que é significativamente pequeno quando comparado com a folga máxima permitida para a folga real dos veículos em inspeção.

Surge assim a necessidade de um software de cálculo que dadas as coordenadas medidas em locais específicos de cada uma das rodas de um veículo, em relação a um plano base de referência, devolva os valores dos ângulos de sopé e de convergência calculados apenas por deduções geométricas e implementadas em código, produzindo um algoritmo numérico, tal como o descrito em [1]. Este texto surge assim como uma breve descrição dos métodos e algoritmos utilizados na resolução do problema em análise.

O software desenvolvido é designado por *Automotive Wheels TCGC* sendo a sigla TCGC o acrónimo de Toe and Camber Geometric Calibration, e apresenta o seguinte comportamento:

1. Permite carregar um ficheiro com os dados de entrada do problema, os quais devem estar na forma de um ficheiro de texto, sem formatação adicional com os valores separados por vírgulas ou espaços, dispostos por linhas, cada linha com três valores (correspondentes às coordenadas xyz do espaço) e com um número de linhas múltiplo de 5 pois admite-se que cada plano de cada roda é dado por cinco pontos. O primeiro grupo de cinco pontos serve para definir o plano base e a partir daí se calculam os ângulos de sopé e convergência relativos a cada um dos grupos de cinco pontos lidos no ficheiro de entrada dos dados. De notar que ao carregar o ficheiro de dados é feita uma análise à sua estrutura, pelo que um erro será vertido caso a estrutura não esteja conforme a especificada. Na Figura 4 é apresentada uma imagem com um exemplo de um tal ficheiro.
2. Permite ver de forma resumida apenas os valores dos ângulos calculados, apresentando os resultados para cada um dos ângulos calculados, para cada um dos grupos de cinco pontos lidos no ficheiro de entrada (exceptuando o primeiro grupo que – tal como já referido, é utilizado para calcular o plano de referência) na forma de grau, minuto e segundo de arco. Por exemplo os dados de entrada, apresentados na Figura 4 resultam na visualização da matriz (1), apresentada a seguir.
3. Permite imprimir um relatório detalhado com todos os valores obtidos durante o processo de cálculo, tais como os coeficientes dos planos de suporte e

Fig. 4: Estrutura dos dados de entrada da aplicação.

de cada um dos grupos de cinco pontos correspondentes a cada uma das rodas em análise de convergência.

O relatório é gravado num ficheiro, cujo nome é solicitado ao utilizador, e apresenta-se com a seguinte informação, continuando com os dados de exemplo referidos na Figura 4.

```
Dados eixo traseiro
vector unit. [-0.289 0.957 0.00909]
largura 1.34
Dados do plano de referência
normal [0.00735 -0.00728 1]
eixo impulso [-0.289;0.957;0.00909]
Distância entre vias 1.52
Centros das rodas:
x y z
-0.686 +1.520 +0.005
+0.679 +1.525 -0.005
-0.671 -0.000 +0.000
+0.671 +0.000 -0.000
-----
Ensaio
conv. sopé
TD - 00°07'46" + 01°15'37"
TE - 00°13'20" + 01°36'51"
FD - 00°03'17" + 00°39'51"
FE - 00°11'23" + 01°19'09"
1 - 01°18'56" + 01°49'07"
2 + 01°10'17" + 01°50'46"
3 - 00°07'18" + 00°22'51"
4 - 00°21'20" + 03°25'30"
```

4. O software apresenta também a possibilidade de se aceder ao manual de utilizador.

$$\begin{pmatrix} 0 & 20.0 & 11.1941 & 2.0 & 4.0 & 30.1951 \\ 0 & 24.0 & 51.5217 & 0 & 47.0 & 50.2148 \\ 0 & 15.0 & 5.87463 & 0 & 18.0 & 50.0528 \\ 0 & 23.0 & 4.15565 & 0 & 31.0 & 29.2266 \\ 1.0 & 30.0 & 24.7812 & 1.0 & 2.0 & 44.2366 \\ 0 & -58.0 & 47.7826 & 1.0 & 2.0 & 0.828234 \\ 0 & 19.0 & 13.6543 & 0 & 29.0 & 10.3651 \\ 0 & 32.0 & 17.5752 & 2.0 & 36.0 & 37.5797 \end{pmatrix} \quad (1)$$

O presente artigo de divulgação tem como principal objectivo descrever a análise que foi feita para a implementação dos procedimentos utilizados na aplicação e é um resumo do Relatório Técnico [1]. Está organizado da seguinte maneira. Na secção *Análise do sistema* faz-se uma análise exhaustiva sobre a dedução de todas as equações e relações necessárias para implementar o método, mas apenas do ponto de vista teórico. Nesta parte assume-se que o leitor está familiarizado com os conceitos elementares do Cálculo Diferencial assim como da Geometria Analítica ou da Álgebra Linear. Estes conceitos estão presentes em qualquer livro ou manual sobre essas matérias, como por exemplo [2, 3, 4]. Na secção que conclui este texto são apresentados pequenos trechos com a listagem parcial do código utilizado na implementação do algoritmo descrito. Assume-se que o leitor está familiarizado com os rudimentos da programação em Matlab que podem ser encontrados por exemplo em [5, 6].

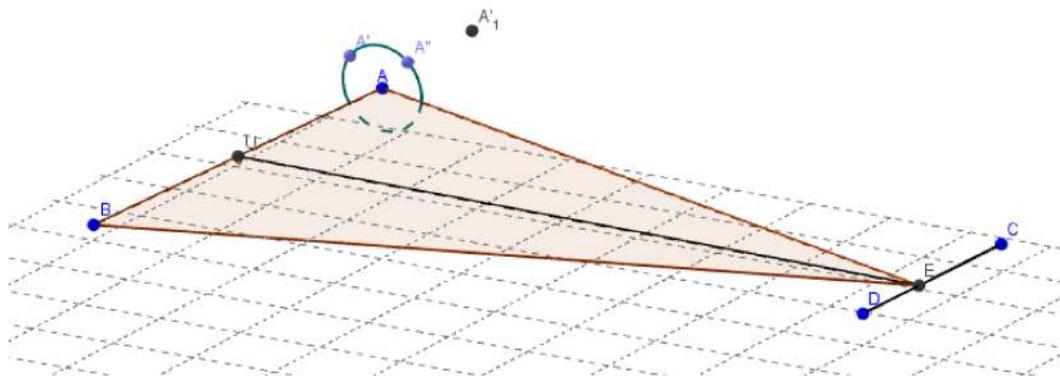


Fig. 1: Ilustração da definição do plano de referência. A, B, C e D são os centros dos planos-alvo.

2 Análise do sistema

O texto a seguir especifica o método e algoritmo para a determinação dos ângulos de sopé e convergência e respectivos erros. Para a determinação dos ângulos de sopé e convergência é necessário definir o plano de referência. Este corresponde ao plano que contém o eixo traseiro (definido pelos centros das duas rodas traseiras) assim

como o eixo longitudinal (definido pela normal ao eixo traseiro pelo seu centro, que intersecta o eixo dianteiro, cf. Figura 1 e Figura 3). Considerando que foram obtidas as coordenadas (esféricas) dos centros das rodas traseiras e dianteiras, (para roda dianteira esquerda: ϱ_{FL} , φ_{FL} , θ_{FL} ; em que o primeiro índice $\{F|R\}$ indica o eixo (*Front* ou *Rear*) e o segundo $\{L|R\}$ indica o lado (*Left* ou *Right*)). Estas doze coordenadas podem ser convertidas em coor-

denadas cartesianas: $(x, y, z)_{FL}$ ou de forma abreviada x_{FL}^i , com $i \in \{1, 2, 3\}$ correspondendo aos eixos coordenados x, y, z respectivamente. Designaremos por \vec{v}_{FR} o vector posicional do centro da roda dianteira esquerda, relativamente à origem O . Os ângulos de sopé e convergência de cada roda iremos designar por α e β respectivamente, com os índices apropriados.

Sejam $\vec{c}_F = \frac{\vec{v}_{FL} + \vec{v}_{FR}}{2}$ o centro do eixo dianteiro, e idênticamente \vec{c}_R ; e $\vec{u}_F = \frac{\vec{v}_{FL} - \vec{v}_{FR}}{2}$ o vector que representa o semi-eixo dianteiro; $d_F = \|\vec{u}_F\|$ a largura do semi-eixo; e ainda $\hat{u}_F = \frac{\vec{u}_F}{d_F}$ o vector unitário com a direcção do eixo dianteiro (e sentido da direita para a esquerda). Seja \vec{e} o ponto onde uma recta passando por \vec{c}_R perpendicular ao eixo traseiro, intersecta o eixo dianteiro:

$$\begin{cases} (\vec{e} - \vec{c}_R) \cdot \hat{u}_R = 0 \\ \vec{e} = \vec{c}_F + t \hat{u}_F \end{cases}$$

Deste sistema de equações podemos determinar o parâmetro t por:

$$t = \frac{\hat{u}_R \cdot (\vec{c}_R - \vec{c}_F)}{\hat{u}_F \cdot \hat{u}_R}$$

ou, após expansão,

$$t = \frac{\sum_i (x_{iFR} x_{iRR} - x_{iFR} x_{iRL} + \frac{1}{2}(x_{iRL}^2 - x_{iFL}^2))}{\sum_i ((x_{iFL} - x_{iFR}) \cdot (x_{iRL} - x_{iRR}))}$$

O plano de referência é definido pela sua normal, $\hat{w} = \hat{u}_R \otimes \hat{v}$, em que $\hat{v} = \frac{\vec{e} - \vec{c}_R}{\|\vec{e} - \vec{c}_R\|}$. A equação do plano é

$$\hat{w} \cdot (r - \vec{c}_R) = 0.$$

As componentes cartesianas de \hat{w} são assim proporcionais à expressão apresentada em nota de rodapé².

Passamos agora para a determinação do ângulo de sopé. Tendo definido o plano de referência, consideremos que além do ponto central de cada roda, foram medidos dois outros pontos no mesmo plano. Iremos considerar

por enquanto apenas uma das rodas, por exemplo a roda traseira esquerda. Os três pontos em questão iremos designar por A, A' e A'' (Figura 1).

Os vectores $A' - A$ e $A'' - A'$ definem o plano da roda, cuja orientação pode ser associada ao vector normal unitário

$$\hat{n}_{RL} = (\vec{v}_{A'} - \vec{v}_A) \otimes (\vec{v}_{A''} - \vec{v}_{A'}).$$

O ângulo de sopé corresponde ao ângulo do plano da roda com o plano de referência, e pode ser obtido por:

$$\varphi_{RL} = \arccos(\hat{n}_{RL} \cdot \hat{w}).$$

Já o ângulo de convergência terá de ser definido com referência ao ângulo de sopé; será o ângulo entre a projecção de \hat{n}_{RL} sobre o plano de referência e a recta que contém o eixo traseiro. Uma formulação equivalente é: o ângulo formado pela recta definida pela intersecção do plano da roda com o plano de referência e pelo eixo de impulso.

A projecção de \hat{n}_{RL} sobre o plano de referência é dada por

$$\hat{n}' = \frac{\hat{n} - (\hat{n} \cdot \hat{w}) \hat{w}}{\sin \varphi} = \frac{\hat{n}}{\sin \varphi} - \hat{w} \cdot \cot \varphi$$

e o ângulo de convergência será

$$\begin{aligned} \vartheta_{RL} &= \arccos(\hat{n}'_{RL} \cdot \hat{u}_R) \\ &= \arccos\left(\frac{\hat{n}_{RL} \cdot \hat{u}_R}{\sin \varphi} - \hat{w} \cdot \hat{u}_R \cdot \cot \varphi_{RL}\right) \\ &= \arccos\left(\frac{\hat{n}_{RL} \cdot \hat{u}_R}{\sin \varphi}\right) \end{aligned}$$

pois $\hat{w} \cdot \hat{u}_R = 0$.

Terminando assim a primeira análise do problema.

Uma outra forma de ver o problema consiste em procurar a transformação de coordenadas que nos leva para um sistema de eixos alinhado com o plano de referência. Ao aplicar essa transformação, os ângulos de sopé e convergência são apenas o ϑ e φ das coordenadas esféricas do vector normal a cada roda.

2

$$\hat{u}_R \otimes (\vec{c}_F - \vec{c}_R + t \hat{u}_F) = \frac{1}{2d_R}$$

$$\begin{bmatrix} 2(y_A - y_B)(z_A + z_B - z_C - z_D + k(z_C - z_D)) - 2(z_A - z_B)(y_A + y_B - y_C - y_D + k(y_C - y_D)) \dots \\ 2(z_A - z_B)(x_A + x_B - x_C - x_D + k(x_C - x_D)) - 2(x_A - x_B)(z_A + z_B - z_C - z_D + k(z_C - z_D)) \\ 2(x_A - x_B)(y_A + y_B - y_C - y_D + k(y_C - y_D)) - 2(y_A - y_B)(x_A + x_B - x_C - x_D + k(x_C - x_D)) \end{bmatrix}$$

com

$$k = \frac{x_A x_C + x_A x_D - x_B x_C - x_B x_D + y_A y_C + y_A y_D - y_B y_C - y_B y_D + \dots + z z_C + z z_D - z_B z_C - z_B z_D - x_A^2 + x_B^2 - y_A^2 + y_B^2 - z_A^2 + z_B^2}{x_A x_C - x_A x_D - x_B x_C + x_B x_D + y_A y_C - y_A y_D - \dots - y_B y_C + y_B y_D + z_A z_C - z_A z_D - z_B z_C + z_B z_D}$$

$$d_R = \|\vec{v}_R\|$$

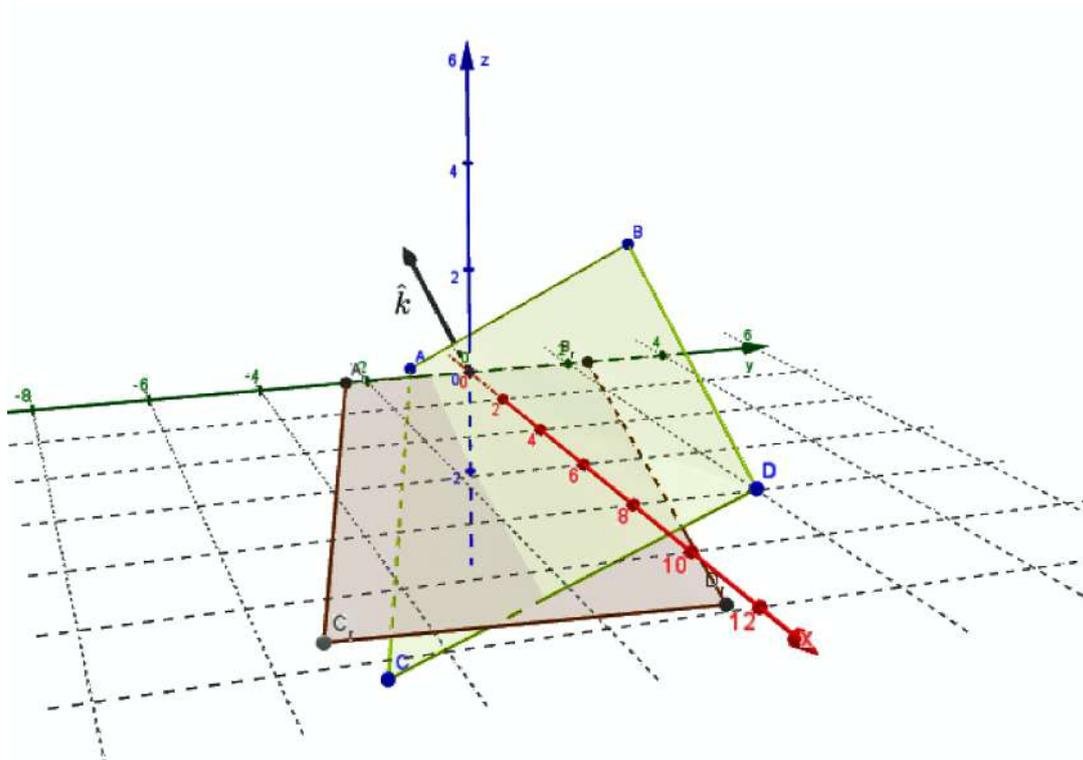


Fig. 2: Ilustração da transformação do sistema de coordenadas do plano-alvo para o sistema de coordenadas do plano de referência.

Para determinar essa transformação, vamos considerar sucessivamente: uma translacção para o centro do eixo traseiro, uma rotação em torno do eixo normal simultaneamente ao eixo traseiro e ao eixo dos xx , e por fim uma rotação em torno do (novo) eixo dos xx (Figura 2).

A primeira operação é trivial: $\vec{v}^f = \vec{v} - \vec{c}_R$.

Para a segunda, precisamos determinar o eixo de rotação $\hat{k} = \hat{i} \times \hat{u}_R$, assim como o ângulo,

$$\omega = \arccos(\hat{i} \cdot \hat{u}_R).$$

Com estes dados, a formula de rotação é a seguinte, também conhecida por formula de Rodrigues.

$$\vec{v}^f = \vec{v} \cos \omega + \hat{k} \times \vec{v} \sin \omega + \hat{k} (\hat{k} \cdot \vec{v}) (1 - \cos \omega)$$

A transformação composta é assim dada pela fórmula apresentada na nota de rodapé ³, que pode também ser

³

$$\vec{v}^f = (\vec{v} - \vec{c}_R) (\hat{i} \cdot \hat{u}_R) + (\hat{i} \times \hat{u}_R) \times (\vec{v} - \vec{c}_R) \sqrt{1 - (\hat{i} \cdot \hat{u}_R)^2} + \quad (2)$$

$$+ (\hat{i} \times \hat{u}_R) ((\hat{i} \times \hat{u}_R) \cdot (\vec{v} - \vec{c}_R)) (1 - (\hat{i} \cdot \hat{u}_R)) \quad (3)$$

ou na forma matricial

$$\vec{v}^f = \mathbf{M} (\vec{v} - \vec{c}_R) \quad (4)$$

$$\mathbf{M} = \begin{bmatrix} a^2 + b^2 - c^2 - d^2 & 2(bc - ad) & 2(bd + ac) \\ 2(bc + ad) & a^2 + c^2 - b^2 - d^2 & 2(cd - ab) \\ 2(bd - ac) & 2(cd + ab) & a^2 + d^2 - b^2 - c^2 \end{bmatrix} \quad (5)$$

escrita na forma matricial, considerando

$$\begin{aligned} a &= \sin(\omega/2) \\ b &= k_x \sin(\omega/2) \\ c &= k_y \sin(\omega/2) \\ d &= k_z \sin(\omega/2). \end{aligned}$$

A operação final irá alinhar o plano por forma que o eixo dos yy final intersecte o eixo dianteiro. Para tal precisamos determinar a cota z' onde o eixo dianteiro intersecta o plano $x = 0$.

Note-se que, sendo esta última rotação em torno do eixo final dos xx , e sendo tipicamente de amplitude reduzida, o seu efeito nos valores da convergência e sopé poderá ser na prática desprezável.

3 Conclusão

Uma estimativa para o erro cometido na dedução descrita na secção anterior é importante porque garante a fiabilidade numérica do sistema de medição.

Pela análise dos erros que foi efectuada pudemos concluir que a maior incerteza nos ângulos em altura e azimute do plano de referência, que se somam com as incertezas correspondentes a convergência e sopé, é da ordem de (em minutos de grau): 0.2410, 0.1990.

Estes resultados indicam assim que as incertezas associadas à determinação do plano de referência não são demasiado elevadas; combinadas com as incertezas já analisadas para o a orientação do plano de cada roda.

Em resumo foi implementada uma função que utiliza as funções auxiliares apresentadas em [1] para calcular os ângulos de sopé e de convergência relativamente a um plano de referência. Os dados de entrada consistem em

grupos de 5×3 valores numéricos obtidos a partir de um ficheiro em formato csv (comma separated values) tal como ilustrado na Figura 4. Os primeiros quatro blocos (ou grupos de coordenadas) correspondem, respectivamente, às rodas traseira direita, traseira esquerda, frente direita e frente esquerda. Uma das funções auxiliares apresentadas calcula o plano de referência de acordo com as fórmulas apresentadas na secção anterior. Esta função recebe os centros das quatro rodas determinadas pela médias dos cinco pontos medidos em cada plano, e devolve entre outros, a normal unitária ao plano e a direcção do eixo traseiro (vector unitário paralelo ao eixo traseiro). Uma outra função implementa a matriz de mudança de coordenadas deduzida em (5) que transporta os dados de entrada para o sistema de eixos do veículo conforme o plano de referência determinado. A Figura 3 ilustra os elementos geométricos relevantes representados no sistema de coordenadas do veículo.

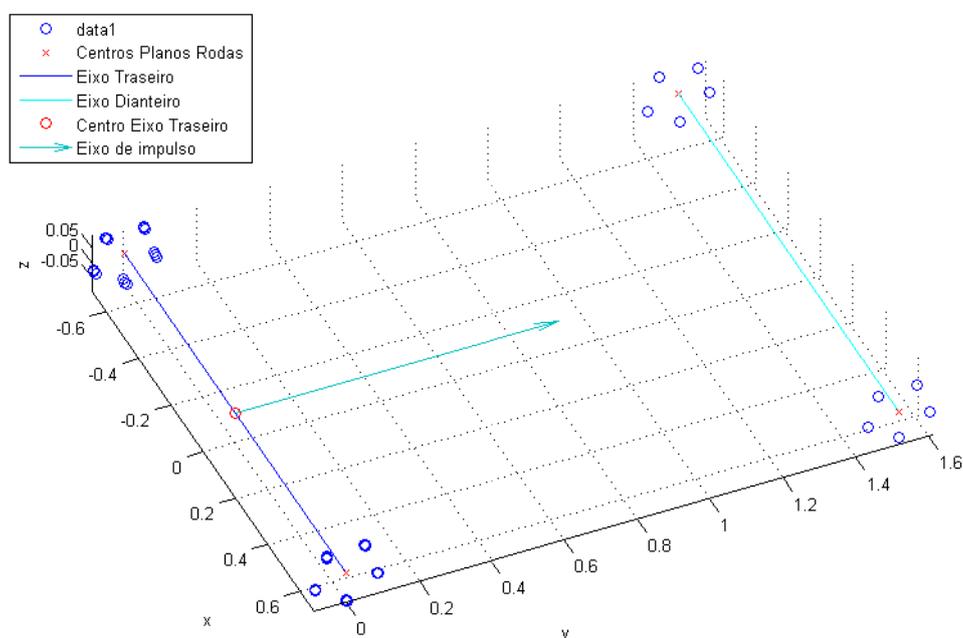


Fig. 3: Alguns elementos geométricos relevantes representados no sistema de coordenadas do veículo.

Notamos que esta operação pode ser efectuada de forma eficiente em termos numéricos recorrendo aos operadores de divisão de matrizes em Matlab, evitando assim o cálculo explícito desta matriz, tal como se pode observar na linha 26 do código da função principal:

```
Mt=M-repmat(c_R,[size(M,1),1])/[uxv,uyv,uzv]';
```

sendo que os vectores uxv , uyv e uzv são os vectores da base do sistema de eixos do veículo.

Para cada grupo de 5 coordenadas, que correspondem a cada um dos ensaios, é ajustado um plano, pelo método dos mínimos desvios quadrados, através da função `fit`

do Matlab. Este procedimento é executado pela função a seguir que devolve a normal unitária a cada um dos planos.

A última das funções auxiliares apresentadas em [1] converte um vector de coordenadas cartesianas para coordenadas esféricas. Pela sua simplicidade é também apresentada aqui.

```
function tpr=c2s(V)
% convert a vector to spherical coordinates
[tpr(1,:) tpr(2,:)]
tpr(3,:)=cart2sph(V(1,:),V(2,:),V(3,:));end
```

A aplicação desta última função devolve os ângulos θ e φ directamente relacionados com os ângulos de convergência e sopé. Em rigor, existe uma ligeira diferença relativamente à definição formal ISO8855, mas que se situa na ordem dos segundos de grau para valores típicos de convergência e sopé.

Por fim apresentamos um exemplo de aplicação do software e respectivos resultados. O primeiro passo para se poder utilizar a aplicação é ter um ficheiro de texto tal como se ilustra na Figura 4 apresentada a seguir.

Na figura podemos observar que a estrutura deve apresentar os valores com as coordenadas (x, y, z) de cada ponto por linha e separadas por vírgulas. Em Matlab pode ser utilizado o comando `csvwrite` para escrever num ficheiro uma matriz definida no ambiente de trabalho no referido formato.

A informação constante das primeiras 20 linhas serve para definir o plano de referência. As linhas 1:5 correspondem a pontos medidos sobre o plano da roda traseira direita equidistantes do eixo da roda, e da mesma

maneira as linhas 6:10 correspondem à roda traseira esquerda, as linhas 11:15 correspondem à roda dianteira direita, e as linhas 16:20 correspondem à roda dianteira esquerda. As linhas restantes organizam-se igualmente em blocos de cinco, e correspondem aos diferentes ensaios efectuados sobre cada uma das rodas indiferenciadamente.

Ao ser lido o ficheiro, os seus dados são seleccionados e carregados no sistema. Em caso de erro na estrutura do ficheiro, será apresentada uma mensagem de erro. Em caso de sucesso, será apresentado o número total de pontos lidos.

O cálculo é executado de seguida. Na tabela da aplicação será apresentado o resultado, consistindo nos valores do ângulo de convergência e sopé, apresentados na forma grau, minuto, segundo.

Por fim os dados são exportados na forma de um relatório detalhado para um ficheiro de texto seleccionado pelo utilizador.

References

- [1] Miguel Belbut Gaspar, Nelson Martins-Ferreira and Artur Mateus, *Automotive Wheels TCGC*, Relatório Técnico produzido por GTLab, The Geometry and Topology Laboratory, CDRSP, Ipleiria, September, 2013.
- [2] Kettner, Lutz, and Stefan Näher. *Handbook of Discrete and Computational Geometry*, Chapman & Hall/CRC, 2004.
- [3] Howard Anton, *Elementar Linear Algebra*, Wiley, 2006
- [4] Howard Anton, *Cálculo, um novo horizonte*, Bookman, 2007
- [5] A. Batel Anjo, R. Fernandes e A. Simões Carvalho, *Curso de Matlab*, Principia, 2003
- [6] Vagner Morais e Cláudio Vieira, *Matlab 7 & 6 - curso completo*, FCA, 2006



Pontiac Solstice GXP

Curvas Geodésicas em Superfícies

by J. P. FATELO AND N. MARTINS-FERREIRA

ESTG, CDRSP,
Instituto Politécnico de Leiria

1 Introdução

O nome *Geodésica* tem a sua origem na Grécia antiga onde já se sabia que a Terra ("Geo") não era plana. Como encontrar o caminho mais curto entre lugares da superfície terrestre era então uma questão importante. Daí em diante, as curvas de menor comprimento entre dois pontos numa determinada superfície ficaram conhecidas como geodésicas. A noção de geodésica foi formalizada com rigor a partir do século XVIII, com o surgimento do cálculo diferencial. No início do século XX, a teoria da relatividade geral, na qual as geodésicas têm um papel de destaque (ver por exemplo [3]), veio definitivamente reforçar a importância do estudo destas curvas.

No espaço euclidiano, todos aprendemos que o caminho mais curto entre dois pontos é o segmento de reta que os liga. Em geometria não euclidiana, encontrar uma geodésica é, em geral, uma tarefa árdua que, na maior parte dos casos, se consegue realizar apenas com a ajuda de aproximações numéricas. O objetivo deste artigo é introduzir, de forma concisa, as ferramentas necessárias para abordar este problema no caso de superfícies em \mathbb{R}^3 . Em particular, são apresentadas as equações diferenciais que a parametrização de uma curva geodésica deve verificar.

Vejamus um exemplo: considere a superfície cônica, parcialmente representada na figura seguinte, na qual os pontos **A** e **B** pertencem a um paralelo, ou seja, a uma circunferência contida na superfície e num plano perpendicular ao eixo de simetria da mesma.

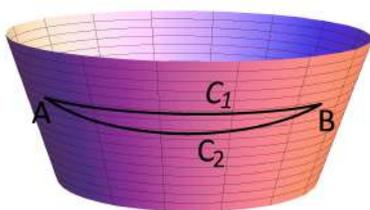


Fig. 1: Vista parcial de um cone com um paralelo e uma geodésica

Numa primeira análise, poderíamos pensar que a distância mais curta, sobre a superfície, entre **A** e **B** seria percorrida ao longo do paralelo que liga os dois pontos (curva C_1 assinalada na figura). Todavia, isso não corresponde à verdade, uma vez que girar em torno do cone sobre

paralelos de menor raio equivale a uma menor distância percorrida. Contudo, um desvio de C_1 também implica percorrer uma distância adicional. Assim, é necessário ponderar o que se ganha em percorrer um perímetro mais curto e o que se perde ao afastar-se da curva C_1 . A solução para este problema de otimização é naturalmente uma geodésica que, no caso da figura, é a curva C_2 .

A determinação de geodésicas numa superfície cônica é relativamente simples, principalmente porque o cone é localmente isométrico a um plano, o que significa, em particular, que as geodésicas no cone são as imagens de retas do plano. O mesmo acontece numa superfície cilíndrica. Mas, em geral, o problema de encontrar as geodésicas numa superfície qualquer é bastante mais complicado.

Uma superfície regular σ em \mathbb{R}^3 pode ser parametrizada por uma função $\vec{\sigma} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ da forma

$$\vec{\sigma}(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (1)$$

onde x , y e z são funções diferenciáveis de duas variáveis reais. Por exemplo, a função vetorial \vec{s} definida por

$$\vec{s}(r, \theta) = (r \cos \theta, r \sin \theta, r^2), \quad 0 \leq \theta < 2\pi, \quad r \geq 0$$

parametriza um parabolóide elíptico,

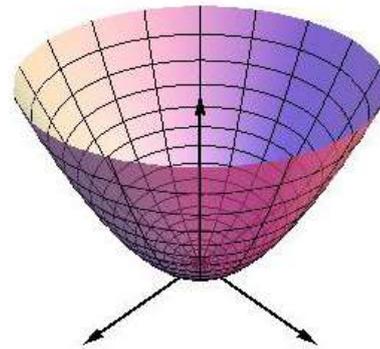


Fig. 2: Parabolóide elíptico

enquanto a função

$$\vec{w}(x, y) = (x, y, x^2 - y^2), \quad x, y \in \mathbb{R}$$

parametriza um parabolóide hiperbólico.

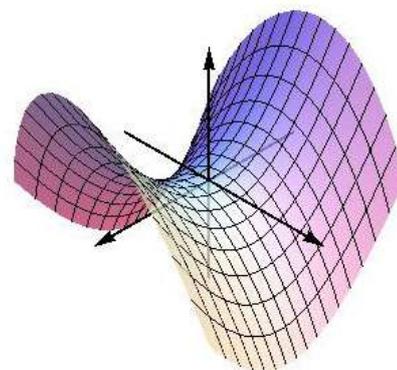


Fig. 3: Parabolóide hiperbólico

2 Equações das geodésicas

Se uma superfície é parametrizada conforme (1), então uma curva dessa superfície admite a parametrização

$$\vec{r}(t) = \vec{\sigma}(u(t), v(t)), \quad t \in I \subseteq \mathbb{R}, \quad (2)$$

onde u e v são funções reais de uma variável real. Por conseguinte, a uma curva de superfície com determinadas características estão associadas expressões específicas $u = u(t)$ e $v = v(t)$. Em particular, se a superfície for Riemanniana, ou seja, se estiver equipada de uma métrica, é possível dar um significado ao comprimento de uma curva e caracterizar curvas geodésicas.

Para determinar a métrica correspondente a uma superfície Riemanniana parametrizada por (1), consideram-se variações infinitesimais du e dv dos parâmetros u e v às quais corresponde um deslocamento infinitesimal $\vec{d\sigma}$ na superfície dado por

$$\vec{d\sigma} = \left(\frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv, \frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv, \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv \right).$$

O tensor métrico g é uma matriz simétrica ($g_{ij} = g_{ji}$) definida, através da norma $ds = \|\vec{d\sigma}\|$, por

$$(ds)^2 = (du \ dv) g \begin{pmatrix} du \\ dv \end{pmatrix}.$$

Por exemplo, os tensores métricos dos paraboloides elíptico e hiperbólico apresentados acima são, respetivamente, os seguintes:

$$g_s = \begin{pmatrix} 1 + 4r^2 & 0 \\ 0 & r^2 \end{pmatrix}, \quad g_w = \begin{pmatrix} 1 + 4x^2 & -4xy \\ -4xy & 1 + 4y^2 \end{pmatrix}.$$

Conhecendo o tensor métrico associado a uma superfície, podemos calcular o comprimento L de uma curva limitada da mesma, digamos γ , pela fórmula

$$L = \int_{\gamma} ds.$$

Dados dois pontos, um caminho geodésico entre eles é uma curva que, pelo menos localmente, minimiza L . Mostra-se (ver por exemplo [2]) que, numa superfície Riemanniana σ em \mathbb{R}^3 , o problema de encontrar uma curva geodésica parametrizada pela expressão (2) reduz-se à resolução do seguinte sistema de equações diferenciais de 2.^a ordem não lineares acopladas:

$$\ddot{u} + \Gamma_{11}^1 \dot{u}^2 + 2\Gamma_{12}^1 \dot{u}\dot{v} + \Gamma_{22}^1 \dot{v}^2 = 0 \quad (3)$$

$$\ddot{v} + \Gamma_{11}^2 \dot{u}^2 + 2\Gamma_{12}^2 \dot{u}\dot{v} + \Gamma_{22}^2 \dot{v}^2 = 0. \quad (4)$$

Uma variável com um "ponto" sobreposto representa a derivada da variável em ordem ao parâmetro t :

$$\dot{u} = \frac{du}{dt}, \quad \dot{v} = \frac{dv}{dt}, \quad \ddot{u} = \frac{d^2u}{dt^2}, \quad \ddot{v} = \frac{d^2v}{dt^2}.$$

Os oito coeficientes Γ_{ij}^k ($i, j, k = 1, 2$), chamados símbolos de Christoffel, podem ser obtidos diretamente a partir do tensor métrico g através da fórmula

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{l=1}^3 (g^{-1})^{kl} (g_{il,j} + g_{jl,i} - g_{ij,l}),$$

onde $(g^{-1})^{ij}$ ($i, j = 1, 2$) são os elementos do tensor inverso de g e as derivadas de g são denotadas por

$$g_{ij,1} = \frac{\partial g_{ij}}{\partial u}, \quad g_{ij,2} = \frac{\partial g_{ij}}{\partial v}.$$

Na resolução das equações diferenciais, é crucial notar que qualquer solução (u, v) do sistema de equações (3) e (4) verifica a condição

$$g_{11} \dot{u}^2 + 2g_{12} \dot{u}\dot{v} + g_{22} \dot{v}^2 = C. \quad (5)$$

Esta propriedade, que facilmente se demonstra, tem como interpretação física o facto de uma geodésica ser percorrida a velocidade constante $\|\frac{ds}{dt}\| = C \in \mathbb{R}$. Efectivamente, é possível deduzir as equações (3) e (4) impondo a condição que a aceleração de uma partícula que percorre uma geodésica deve ser perpendicular à superfície, não havendo nenhuma aceleração tangencial (e portanto nenhuma força tangencial) aplicada à partícula. Isto significa que uma geodésica é o caminho que uma partícula livre, em movimento, percorre naturalmente.

Segue a apresentação explícita das equações (3) e (4) em dois casos particulares importantes: superfícies de revolução e gráfico de uma função real de duas variáveis reais.

3 Superfícies de revolução

Uma superfície de revolução é obtida por rotação de uma curva plana simples de \mathbb{R}^3 em torno de uma reta contida no mesmo plano da curva e que não a intersecta. Esta reta é designada por eixo de revolução. Sem perda de generalidade, podemos escolher o eixo de revolução como sendo o eixo dos zz . Neste caso, a parametrização de uma superfície de revolução toma a seguinte forma

$$\vec{\sigma}(\alpha, \theta) = (r(\alpha) \cos \theta, r(\alpha) \sin \theta, z(\alpha)), \quad (6)$$

onde as funções r e z são funções diferenciáveis e $\theta \in [0, 2\pi[$. Por exemplo, com $r(\alpha) = \alpha$, $z(\alpha) = \alpha^2$ e $\alpha \geq 0$, obtemos a parametrização de um parabolóide elíptico; com $r(\alpha) = 2 + \cos \alpha$, $z(\alpha) = \sin \alpha$ e $\alpha \in [0, 2\pi[$, obtemos a parametrização de um *torus* [4].

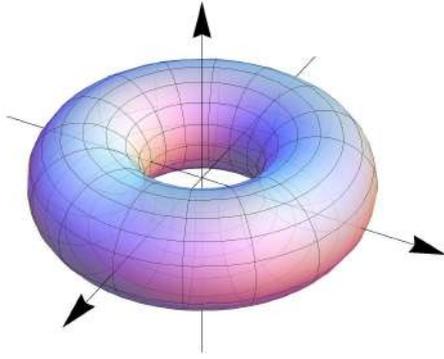


Fig. 4: Torus

Numa superfície de revolução parametrizada conforme (6), o sistema de equações (3)-(4) assume a seguinte forma:

$$\ddot{\alpha} = \frac{r r'}{r'^2 + z'^2} \dot{\theta}^2 - \frac{r' r'' + z' z''}{r'^2 + z'^2} \dot{\alpha}^2$$

$$r \ddot{\theta} = -2 r' \dot{\theta} \dot{\alpha}.$$

No exemplo do cone com parametrização

$$\vec{r}(\alpha, \theta) = (\alpha \cos \theta, \alpha \sin \theta, \alpha)$$

estas equações reduzem-se a

$$\ddot{\alpha} = \frac{\alpha}{2} \dot{\theta}^2$$

$$r \ddot{\theta} = -2 \dot{\theta} \dot{\alpha},$$

e a equação (5), com a escolha $C = 1$, escreve-se

$$2 \dot{\alpha}^2 + \alpha^2 \dot{\theta}^2 = 1.$$

Este sistema de equações admite a solução $\dot{\alpha}(t) = \pm \frac{1}{\sqrt{2}}$ e $\dot{\theta}(t) = 0$, o que significa que qualquer geratriz do cone é um caminho geodésico. As outras soluções têm a seguinte forma:

$$\alpha(t) = \sqrt{C_1^2 + \frac{(t + C_2)^2}{2}}$$

$$\theta(t) = \sqrt{2} \operatorname{arctg} \left(\frac{t + C_2}{\sqrt{2} C_1} \right) + C_3.$$

Os valores das constantes C_1 , C_2 e C_3 distinguem as várias geodésicas existentes na mesma superfície. Por exemplo, a indicação de um ponto por onde passa uma determinada geodésica bem como de um vetor unitário tangente à geodésica nesse ponto determina essas constantes. Vale a pena observar que a escolha $C = 1$ pode ser substituída pela escolha de outro valor constante de C , mudando assim apenas a parametrização da curva geodésica mas não a curva em si. Na figura seguinte estão representados alguns caminhos geodésicos no cone exemplificado.

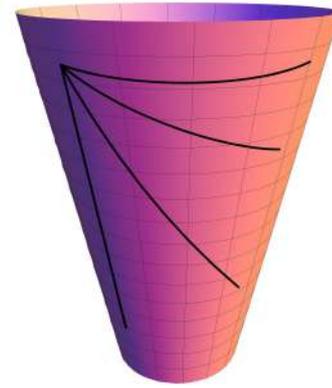


Fig. 5: Exemplo de caminhos geodésicos numa superfície de revolução

4 Gráfico de uma função

O gráfico de uma função f , real de duas variáveis reais, é convenientemente parametrizado por

$$\vec{r}(x, y) = (x, y, f(x, y)). \quad (7)$$

Neste caso, as geodésicas $\vec{r}(t) = (x(t), y(t), f(x(t), y(t)))$ são obtidas após resolver o seguinte sistema de equações diferenciais

$$\ddot{x} = \frac{-f_x}{1 + f_x^2 + f_y^2} (f_{xx} \dot{x}^2 + 2f_{xy} \dot{x} \dot{y} + f_{yy} \dot{y}^2)$$

$$\ddot{y} = \frac{-f_y}{1 + f_x^2 + f_y^2} (f_{xx} \dot{x}^2 + 2f_{xy} \dot{x} \dot{y} + f_{yy} \dot{y}^2).$$

É usada a notação $f_x = \frac{\partial f}{\partial x}$, $f_y = \frac{\partial f}{\partial y}$ e similarmente para as derivadas parciais de segunda ordem f_{xx} , f_{xy} e f_{yy} . No exemplo do parabolóide hiperbólico apresentado no início, as equações reduzem-se a

$$\ddot{x} = \frac{4x}{1 + 4x^2 + 4y^2} (\dot{y}^2 - \dot{x}^2)$$

$$\ddot{y} = \frac{4y}{1 + 4x^2 + 4y^2} (\dot{x}^2 - \dot{y}^2).$$

Na próxima figura estão representadas algumas curvas geodésicas no parabolóide hiperbólico obtidas por aproximação numérica de soluções das equações anteriores.

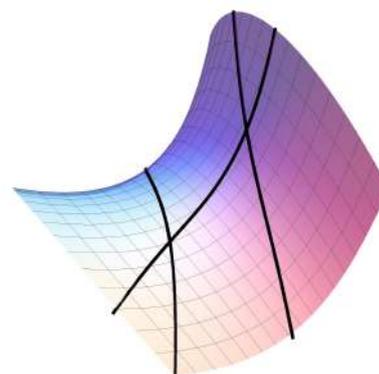


Fig. 6: Exemplo de caminhos geodésicos numa superfície

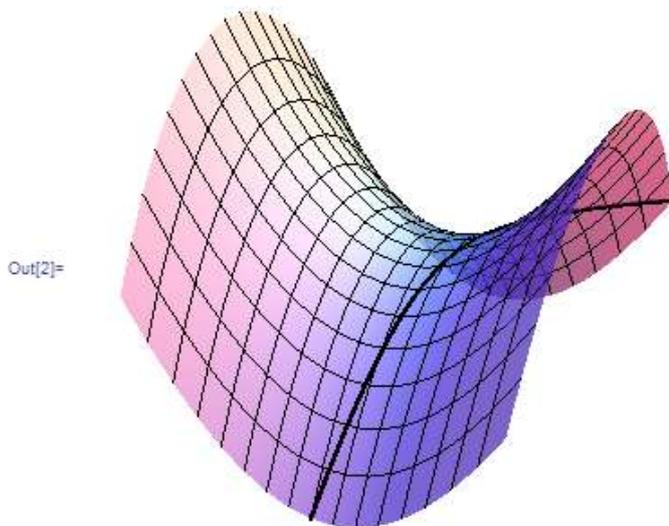
Na maioria dos casos, a determinação de geodésicas é possível apenas com o apoio de aplicações computacionais. As soluções apresentadas geometricamente neste artigo, bem como as figuras expostas foram geradas us-

ando a aplicação Mathematica [1]. Segue um exemplo com duas linhas de comando no Mathematica para determinar e representar uma geodésica e a superfície considerada.

```
In[1]:= sol = NDSolve[{x''[t] =  $\frac{-4 x[t]}{1 + 4 (x[t])^2 + 4 (y[t])^2} (x'[t]^2 - y'[t]^2)$ ,
  y''[t] =  $\frac{4 y[t]}{1 + 4 x[t]^2 + 4 y[t]^2} (x'[t]^2 - y'[t]^2)$ , x[0] = 0,
  y[0] = -1, x[10] = 1, y[10] = 1}, {x, y}, {t, 0, 10}]
```

```
Out[1]= {{x -> InterpolatingFunction[{{0., 10.}}, <>], y -> InterpolatingFunction[{{0., 10.}}, <>]}}
```

```
In[2]:= Show[ParametricPlot3D[{x, y, x^2 - y^2}, {x, -1, 1}, {y, -1, 1},
  Boxed -> False, Axes -> False, PlotStyle -> Directive[Opacity[0.75]]],
  ParametricPlot3D[Evaluate[{x[t], y[t], x[t]^2 - y[t]^2} /. sol, {t, 0, 10}],
  PlotRange -> All, PlotStyle -> {Thickness[0.0075], Black}], ImageSize -> 350]
```



References

- [1] Wolfram Research, Inc., Mathematica, Version 9.0, Champaign, IL (2012).
- [2] A. Pressley, *Elementary Differential Geometry, 2nd Edition*, Springer Undergraduate Mathematics Series, London, 2010.
- [3] L. Ryder, *Introduction to general relativity*, Cambridge University Press, New York, 2009.
- [4] R. T. Jantzen, *Geodesics on the Torus and other Surfaces of Revolution Clarified Using Undergraduate Physics Tricks with Bonus: Nonrelativistic and Relativistic Kepler Problems*, <http://arxiv.org/abs/1212.6206> [math.DG], 2012.