

Faculdade de Ciência e Tecnologias  
da Universidade de Coimbra

Departamento  
de Física

Departamento  
de Engenharia Informática

# **Tópicos de Física Moderna**

Apontamentos da cadeira

Vitaly Chepel

Coimbra, 2006



# Conteúdo

<b>Capítulo 1. Medições e erros</b>	<b>.....5</b>
1. Medições e seus limites	.....5
2. Incertezas	.....8
3. Propagação de erros	.....20
<b>Capítulo 2. Revisão da mecânica newtoniana</b>	<b>.....23</b>
1. Cinemática	.....23
2. Dinâmica	.....26
2.1 Força	26
2.2 A lei fundamental da dinâmica	27
2.3 O círculo vicioso da mecânica	29
3. Trabalho e Energia	.....31
3.1 Trabalho de uma força. Energia cinética	31
3.2 Trabalho das forças conservativas. Energia potencial	40
3.3 Sistemas com forças conservativas e não conservativas	42
<b>Capítulo 3. Interacções</b>	<b>.....45</b>
1. As forças fundamentais da Natureza	.....45
2. Campos	.....47
3. O campo gravítico e o campo electrostático	.....50
3.1 O campo gravítico	50
3.2 Comparação do campo gravítico com o campo electrostático. Potencial atractivo e potencial repulsivo	53
4. Movimento num potencial	.....56
4.1 Potencial repulsivo	56
4.2 Potencial atractivo	57

4.3 Potencial realista – uma molécula diatômica	58
4.4 Equilíbrio	62
<b>Capítulo 4. Oscilações</b>	<b>.....65</b>
1. Oscilações harmônicas simples	.....65
2. Oscilações num sistema arbitrário	.....73
3. Oscilações amortecidas	.....75
4. Oscilações forçadas. Ressonância.	.....79
<b>Capítulo 5. Ondas</b>	<b>.....85</b>
1. Propriedades gerais das ondas	.....86
2. Ondas harmônicas	.....89
2.1 Função da onda harmônica	89
2.2 Exemplos das ondas harmônicas	92
2.3 Polarização	94
2.4 Energia da onda	94
3. A equação da onda	.....95
4. Sobreposição das ondas harmônicas	...100
4.1 Interferência construtiva e destrutiva: <i>Sobreposição de duas ondas planas com fases diferentes que se propagam no mesmo sentido.</i>	100
4.2 Onda estacionária: <i>Sobreposição de duas ondas com a mesma frequência que se propagam em sentidos opostos.</i>	101
4.3 Batimentos: <i>Sobreposição de duas ondas planas que se propagam no mesmo sentido com frequências ligeiramente diferentes.</i>	105
5. Efeito de Doppler	...112
6. Princípio de Huygens. Reflexão e Refracção das ondas.	...117
7. Interferência e Difraccção	...124
7.1 Interferência em duas e três dimensões	124

7.2 A experiência com fenda dupla	127
7.3 Difracção	130
<b>Capítulo 6. Mecânica Quântica</b>	<b>.....135</b>
1. Como tudo começou	...136
2. O princípio da incerteza	...138
3. A equação de Schrödinger	...141
3.1 Análise da equação de Schrödinger	143
3.2 A equação de Schrödinger independente do tempo	145
3.3 Soluções da equação de Schrödinger com $U=const$	146
3.4 Uma partícula numa caixa infinita	149
3.5 Uma partícula numa caixa finita	151
3.6 Efeito túnel	153
<b>Capítulo 7. Informação Quântica</b>	<b>.....157</b>
1. O gato de Schrödinger	...157
2. Bits e qubits	...162
3. Criptografia quântica	...166



# Capítulo 1

## Medições e erros

*“Nothing exists until it is measured”  
(Niels Bohr)*

### 1. Medições e seus limites

As medições fazem parte inerente do processo de conhecimento do nosso mundo material. Nós adquirimos os conhecimentos através dos nossos sentidos que, no fundo, são instrumentos de medição generosamente oferecidos pela Natureza. A partir dessas medições e ainda outras, realizadas com instrumentos mais sofisticados e mais sensíveis feitos pelo homem, construímos as teorias que tentam explicar as relações entre os efeitos e as suas causas e, assim, criamos uma visão coerente do mundo. As teorias podem ser simples ou complexas, baseadas fortemente num formalismo matemático ou apenas nas conclusões lógicas, podem ser auto-consistentes ou algo controversas. Mas há pelo menos uma coisa que todas elas têm em comum: todas as teorias se baseiam num conjunto de suposições iniciais, os postulados (i.e. as afirmações sem qualquer prova) sobre os quais se constrói o prédio inteiro e, portanto, por muito êxito que uma teoria tenha na explicação dos fenómenos observados existe sempre alguma incerteza na validade dessa teoria. Existirá sempre um pequeno demónio de dúvida a cochichar: será que não há outra explicação? E tem toda a razão, temos de admitir. Não podemos ter provas absolutas de que uma teoria seja verdadeira, mas podemos acumular prova após prova tentando atenuar as dúvidas e, assim, ganhar mais confiança na nossa construção teórica. A única maneira de o fazer é comparar as conclusões e previsões da teoria com a experiência. Isto significa mais medições.

“Uma medição é uma pergunta que a ciência faz à Natureza, e uma experiência é a gravação da resposta dela” disse Max Plank, um físico proeminente. “Nada existe até ser medido” – proferiu Niels Bohr, outro físico. E embora o tivesse dito num outro contexto, relacionado com diferentes interpretações da teoria quântica (voltaremos a

falar deste assunto no último capítulo que diz respeito à possibilidade da computação quântica), mas a sua frase ajusta-se muitíssimo bem ao nosso tema, exprimindo a extrema importância das medições.

A medição é sempre imprecisa. Ao medir o comprimento de uma mesa com uma fita métrica não podemos esperar que este seja determinado com uma precisão de  $1\text{ }\mu\text{m}$  ( $10^{-6}\text{ m}$ ), simplesmente porque a fita só tem divisões um em um milímetro ( $10^{-3}\text{ m}$ ) e mesmo que tivesse divisões mais pequenas os nossos olhos não conseguiriam distingui-los. Ao medir a distância entre duas cidades entendemos que esta não pode ser exacta até porque as cidades são objectos extensos e não pontos matemáticos. Se medíssemos várias vezes, uma a seguir a outra, o número de carros que passam por minuto num dado ponto da estrada seria altamente improvável que os resultados das medições coincidissem. Portanto, ao apresentar o resultado de uma medição não basta referir apenas o valor medido, mas é necessário mencionar também a incerteza associada a este (muitas vezes também se diz “o erro experimental”). A forma matemática habitual para reflectir esta imprecisão no resultado da medida é  $l \pm \Delta l$ , o que significa que o resultado está algures dentro do intervalo  $(l - \Delta l, l + \Delta l)$ . Mais adiante, vamos discutir o significado de  $l$ , de  $\Delta l$  e da própria expressão  $l \pm \Delta l$ , em termos mais rigorosos. Para já, ao falar destes, vamos confiar na nossa intuição.

Em muitos casos, não há necessidade de medir um ou outro parâmetro com uma precisão muito boa. Quando se trata de uma peça de mobília, é difícil de imaginar porque seria preciso conhecer o seu comprimento com uma precisão de  $1\text{ }\mu\text{m}$ . Mas se se tratar de uma instalação para uma experiência científica, já é diferente. Em princípio, é sempre desejável melhorar a precisão com que se fazem as medições nas experiências científicas porque o objectivo destas, no fim das contas, é comprovar (ou rejeitar) uma teoria ou descobrir um novo fenómeno ainda não conhecido. Por exemplo, uma experiência que pretendesse provar (ou desacreditar) a teoria da relatividade geral de Einstein através da medição do efeito do campo gravítico da Terra sobre as ondas electromagnéticas, que esta teoria prevê (e que nada tem a ver com a gravidade propriamente dita porque os fotões têm a massa gravítica igual a zero), devia procurar uma alteração da frequência das ondas, a propagar-se à superfície da Terra, na ordem de  $10^{-13}\%$ . Se a imprecisão com que medimos a frequência for superior a esta, a experiência será inconclusiva, i.e. não poderá nem provar nem desacreditar a teoria. (De facto, esta experiência foi realizada em 1959 em



Harvard. A fonte de fótons de uma dada frequência foi colocada a uma altura de cerca de 23 m e mediu-se a frequência dos fótons que chegavam à superfície da Terra com uma precisão de ordem de  $10^{-15} \%$ . Esta experiência confirmou a previsão da teoria de relatividade.)

Então, tudo o que temos de fazer é melhorar os instrumentos e os métodos experimentais e, mais cedo ou mais tarde, reduzir os erros até praticamente zero? E, no limite chegaremos ao conhecimento absolutamente exacto e exaustivo do mundo material? Sem entrar no campo da filosofia, a física diz-nos que existem limites para as incertezas com que é possível medir as grandezas físicas. Ou melhor, há sempre um preço a pagar para quem quer melhorar a precisão na medição de uma dada grandeza física. E este preço é ... piorar a precisão com que pode ser medida uma outra grandeza! É isto que declara o princípio de incerteza de Heisenberg: o produto da incerteza na posição de uma partícula,  $\Delta x$ , pela incerteza com que pode ser medido o momento linear desta,  $\Delta p$  ( $p=mv$ , ou seja a massa da partícula vezes a velocidade dela), é uma constante  $h$  (constante de Plank,  $h = 6.626 \cdot 10^{-34} \text{ J}\cdot\text{s}$ )

$$\Delta x \cdot \Delta p \sim h$$

(o símbolo  $\sim$  significa “igual em ordem de grandeza”; utilizamos  $\sim$  em vez do habitual símbolo de igualdade porque não definimos ainda qual o significado, em termos rigorosos, das incertezas  $\Delta x$  e  $\Delta p$ ). Como o produto das duas incertezas é constante, a diminuição de uma delas resulta no aumento da outra. Por outras palavras, de acordo com o princípio de incerteza não existe nenhuma experiência real ou imaginária em que ambas as grandezas  $x$  e  $p$  (ou a velocidade) pudessem ser medidas simultaneamente com exactidão. E se uma delas, a posição por exemplo, for medida com exactidão (i.e.  $\Delta x=0$ ), a incerteza no momento será infinita (porque  $\Delta p \sim h/\Delta x$ ), e a incerteza na velocidade também (porque  $\Delta p = \Delta(mv) = m\Delta v$ ). Saberíamos exactamente a posição do corpo, mas não poderíamos concluir absolutamente nada acerca da velocidade deste. A velocidade pode ser igual a 0 ou a 100 km/h, ou à da luz, ou até infinita. E, portanto, não teríamos hipótese nenhuma de prever onde estaria o corpo no instante imediatamente a seguir.

O valor da constante de Plank é muito pequeno (mas não é nulo!) e, por isso, no mundo macroscópico ( $m$  relativamente grande) o princípio de incerteza praticamente não tem impacto nos fenómenos habitualmente observados. Mas para os objectos microscópicos, por exemplo para um electrão com a massa  $m_e = 9.1 \cdot 10^{-31} \text{ kg}$ , os

valores de incertezas  $\Delta x$  e  $\Delta v$  são apreciáveis. É fácil de calcular, a partir da relação de Heisenberg, que se se pretende medir a velocidade do electrão com uma precisão de  $1 \text{ m/s}$ , a posição deste não pode ser determinada com uma incerteza menor que  $\approx 0.1 \text{ mm}$ , mesmo que os instrumentos utilizados permitissem fazê-lo com uma precisão muito superior. Contrariamente, quem consegue localizar o electrão com uma incerteza de  $1 \text{ }\mu\text{m}$ , não tem hipótese nenhuma de medir a velocidade da partícula com um erro menor que  $\approx 100 \text{ m/s}$ .

É importante notar que esta limitação não está relacionada com a precisão dos nossos instrumentos de medição ou com a habilidade do experimentador. É um limite fundamental e incontornável, inerente à própria natureza das coisas – à dualidade onda-partícula da matéria. Voltaremos a discutir este assunto no Capítulo 6, dedicado à Mecânica Quântica, mas antes disso temos que falar das incertezas relacionadas com o método experimental e os meios utilizados, que na grande maioria das situações são muito superiores aos limites impostos pelo princípio de incerteza.

## 2. Incertezas

Imaginemos uma experiência simples: vamos medir a aceleração da gravidade através da medição do tempo de queda de uma bola de ténis de uma dada altura e sem velocidade inicial. Sabemos da cinemática que, para um movimento uniformemente acelerado, a distância percorrida pode ser escrita em função do tempo como  $s = v_0 t + \frac{at^2}{2}$ . No nosso caso,  $a = g$  é a aceleração de gravidade, que se pretende determinar,  $v_0 = 0$  (velocidade inicial) e  $s = h$  é a altura da queda. Temos, portanto,  $h = \frac{gt^2}{2}$  ou, invertendo a equação,  $g = \frac{2h}{t^2}$ . Daqui se vê, que se soubermos a altura e o tempo de queda podemos determinar  $g$ . Assim, colocamos uma bola a uma altura, digamos, de  $2 \text{ m}$ , que medimos com uma fita métrica, e largamo-la carregando simultaneamente no botão “start” do cronómetro. Quando ouvimos a bola bater no chão, paramos o cronómetro e efectuamos a leitura deste. Imaginemos que o valor lido é  $0.69 \text{ s}$ , por hipótese. Repare-se, que o valor registado tem duas casas decimais, exactamente como se lê no mostrador do cronómetro. O facto de o mostrador do

cronómetro ter apenas duas casas decimais já limita a precisão com que poderemos medir o tempo da queda. Isto é, o valor da incerteza no tempo é igual ou maior de 0.005 s (metade de 0.01 s de acordo com a convenção de representar o valor obtido na forma  $t \pm \Delta t$ , i.e. metade de 0.01 s para cima e metade para baixo). Este valor caracteriza a precisão do instrumento utilizado para a medição. Vamos chamá-lo *imprecisão instrumental*. (De facto, deveríamos ir ver o manual do cronómetro porque a precisão deste pode ser pior, mas melhor não pode ser porque mesmo se interiormente o relógio for muito preciso, nós não temos acesso a esta informação).

Identificámos uma fonte de incerteza que afecta o valor do tempo medido – a precisão do instrumento. Será que existem outras fontes? Vamos repetir a experiência. Colocamos a bola à mesma altura e medimos o tempo da queda. Quem alguma vez tinha feito uma experiência destas ou semelhante, sabe que é improvável obtermos exactamente o mesmo valor para o tempo. Se efectuarmos mais medições, observamos que os valores que o cronómetro regista são quase sempre diferentes. Será que o próprio tempo da queda varia? Para averiguar isto, modifiquemos a experiência: distribuímos uma dúzia de cronómetros entre as pessoas pedindo a todos que meçam o tempo da queda simultaneamente. Já se pode adivinhar que o resultado será o mesmo. Fizemos essas medições numa sala de aula e obtivemos os resultados indicados na tabela. Vê-se claramente que praticamente todos os valores são diferentes representando uma dispersão apreciável. A razão para isto é bastante clara. Existe um intermediário entre o fenómeno em estudo e o instrumento – o experimentador. É ele que carrega no botão “start/stop” do cronómetro e, portanto, o tempo de reacção dele tem de ser tomado em conta. Este tempo não é constante: por vezes carrega no botão com alguma antecedência, outras distrai-se e carrega com algum atraso. Se forem vários experimentadores a medir o tempo, cada um terá os seus atrasos e antecipações, de modo que as leituras do tempo de queda apresentarão flutuações imprevisíveis, aleatórias. Daqui vem o termo – *erros aleatórios* (também se diz *erros estatísticos*).

Número da medida	1	2	3	4	5	6	7	8	9	10	11
Valor (s)	0.69	0.84	0.55	0.66	0.68	0.59	0.62	0.75	0.63	0.64	0.66
Número da medida	12	13	14	15	16	17	18	19	20	21	
Valor (s)	0.69	0.59	0.58	0.63	0.71	0.52	0.62	0.69	0.49	0.71	

Nesta experiência, o erro aleatório domina claramente o erro instrumental. Podemos, no entanto, pensar em melhorar a experiência eliminando o intermediário: colocando, por exemplo dois sensores, um em cima e outro no chão, e ligando-nos a uma placa de aquisição de dados de um computador. Neste caso concreto é, realmente, possível, diminuir os erros aleatórios e aproximar a precisão na medição do tempo à precisão do cronómetro. Mas, no fundo, a própria Natureza tem carácter aleatório (falaremos nisso nos Capítulos 6 e 7) e, portanto, os erros aleatórios estarão sempre presentes. É verdade que em algumas situações os erros aleatórios não se notam, mas isto é apenas porque outros tipos de erros estão a dominar a incerteza. Um simples exemplo de uma experiência em que o próprio processo em estudo tem carácter aleatório é o seguinte. Imaginemos que se pretende determinar o número de gotas de chuva por  $1 \text{ m}^2$  que caem no solo por unidade de tempo. Espalhamos uma dúzia de folhas A4 no chão, expomo-las à chuva durante um certo intervalo de tempo e depois contamos quantas gotas apanhou cada folha (claro que este método não funciona para qualquer chuva). Será que todas as folhas terão o mesmo número de marcas? O senso comum diz-nos que não. Como cada gota é independente das outras, as distâncias entre elas no plano horizontal são distribuídas aleatoriamente e, portanto, não existe razão nenhuma para que o número de gotas seja igual em todas as folhas. É, assim, um bom divertimento para um fim de semana chuvoso.

Os erros aleatórios, embora sejam inevitáveis, seguem as leis da teoria estatística o que permite reduzir significativamente a sua contribuição para a precisão do resultado. Olhando outra vez para a tabela com os resultados das medições do tempo da queda surge uma pergunta: qual destes valores teremos que apresentar como resultado? A média aritmética, - já se ouve a resposta. É verdade, a teoria estatística confirma que a média aritmética, i.e.  $\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$ , é uma boa escolha (aqui  $x_i$  são os resultados das medições – tempo da queda, no nosso exemplo, e  $N$  é o número de medições efectuadas).

Antes de prosseguirmos, temos de ressaltar o seguinte. Quando estávamos a medir o tempo da queda, subentendia-se que este tinha um certo valor. Isto é, mesmo que ninguém em qualquer parte do mundo esteja a tentar medi-lo, a queda dura um certo tempo ( $X$  segundos), que depende, obviamente, das condições experimentais e das leis da Natureza, mas não depende da vontade do experimentador. Podemos nem

sequer saber que a bola está em queda (ou estava, ou vai estar). O intervalo de tempo que este processo demora é uma realidade objectiva e é independente daquilo que nós pensamos dele e, portanto, temos que admitir que o *valor verdadeiro* do tempo da queda existe. São tais  $X$  segundos que a Natureza ditou que sejam e que nós podemos tentar determinar através das medições.

Provavelmente, os filósofos vão se opor a esta conclusão dizendo que tudo depende de se acreditamos na existência da realidade objectiva (i.e. independente da consciência) ou não. No entanto, do ponto de vista prático não temos muita escolha. Se não acreditarmos na existência do valor verdadeiro de uma grandeza, não faz sentido nenhum tentar medi-lo! (Para sermos absolutamente rigorosos, temos de mencionar que, de facto, na teoria quântica surgem algumas dúvidas quanto ao significado que costumamos depositar na palavra “existe”. Vamos falar um pouco disso no Capítulo 7. Mas, no mundo clássico não parece que isto tenha algum impacto. Pelo menos, toda a nossa experiência adquirida, aponta para isso.)

Concluindo: vamos continuar a acreditar no postulado “o valor verdadeiro de uma grandeza física existe”. O objectivo de uma medição, então, é aceder a este valor. Será possível? Qual dos valores medidos relativamente ao tempo da queda é verdadeiro? A média? Mas não é verdade que a média também vai variar se efectuarmos mais medições e adicionarmos mais valores à tabela? A única resposta, que parece razoável, é que nós não sabemos qual é “valor” do valor verdadeiro e não podemos saber, mas podemos, isso sim, estimar este valor recorrendo às medições. Também podemos tentar reduzir a incerteza da nossa estimativa aproximando-a ao valor verdadeiro, mas nunca poderemos concluir a partir dos dados experimentais que “agora este é que é o valor verdadeiro” pela simples razão que não temos nenhum critério para aplicar: não é possível comparar as medições com o valor verdadeiro sem o saber, primeiro.

Agora podemos rephrasear a afirmação vaga que refere que a média aritmética é “uma boa escolha” substituindo-a por uma outra, mais precisa: *a melhor estimativa para o valor verdadeiro* é a média aritmética dos resultados das medições. Vamos designar a melhor estimativa por  $\tilde{x}$ . Então,  $\tilde{x} = \langle x \rangle$ .

A média dos valores da tabela é  $\langle x \rangle = 0.644761905$  (tal como é apresentado por uma máquina de calcular). Este valor em si faz pouco sentido porque não diz nada sobre a incerteza nele (nem o número de casas decimais a deixar podemos determinar sem termos alguma ideia da incerteza). Esta incerteza depende, obviamente, da

dispersão dos dados experimentais. Para visualizar a dispersão dos dados da tabela, construímos um histograma (um diagrama de colunas) que vamos chamar *histograma de ocorrências*. Para isso, substituímos a escala do tempo, contínua, por uma escala discreta, dividida em intervalos (*bins*), e

contamos o número de ocorrências dentro de cada intervalo, i.e. quantas vezes surge na experiência um valor compreendido entre os limites inferior e superior de cada intervalo.

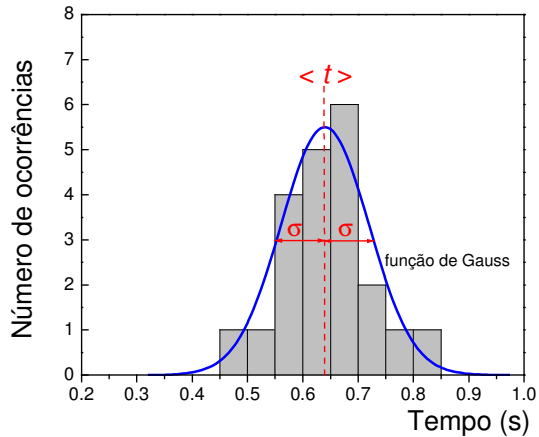
A figura mostra o histograma de ocorrências construído com os dados da tabela. A largura de todos os intervalos foi escolhida, neste caso, de 0.05 s. A interpretação do histograma é a seguinte:

os valores compreendidos no intervalo  $0.45\text{ s} \leq t < 0.5\text{ s}$  foram obtidos em 1 medida;  
os valores compreendidos no intervalo  $0.5\text{ s} \leq t < 0.55\text{ s}$  foram obtidos em 1 medida;  
os valores compreendidos no intervalo  $0.55\text{ s} \leq t < 0.6\text{ s}$  foram obtidos em 4 medidas;  
os valores compreendidos no intervalo  $0.6\text{ s} \leq t < 0.65\text{ s}$  foram obtidos em 5 medidas  
etc.

É óbvio que a soma das ocorrências em todos os intervalos deve dar o número total de medidas efectuadas (21 no nosso exemplo):

$$\sum_i^{\text{número de intervalos}} n_i = N.$$

Um histograma de ocorrências bem construído mostra, pelo menos, duas coisas. Primeiro, indica a região na escala onde se concentram os dados obtidos e, conseqüentemente, permite estimar a média visualmente. Em segundo lugar, dá uma noção sobre a dispersão dos resultados que, como intuitivamente se entende, está relacionada com a incerteza associada à experiência. Isto consegue-se com uma escolha adequada da largura dos intervalos que não devem ser nem demasiado grandes (deste modo todos os dados poderiam cair num intervalo e em vez da distribuição teríamos apenas uma coluna), nem demasiado pequenos (neste caso as contagens espalhar-se-iam sobre toda a escala). Obviamente, a largura deve ser igual para todos os intervalos.



No nosso histograma vê-se claramente, mesmo sem fazer cálculos, que o tempo médio da queda é, aproximadamente, 0.65 s. Também podemos ter uma grosseira ideia a cerca da largura da distribuição definindo a, por exemplo, como largura a meia altura: esta é de ordem de 4 intervalos ou, seja, 0.2 s.

Com mais rigor, a dispersão dos dados experimentais pode ser estimada através da seguinte fórmula

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\langle x \rangle - x_i)^2}$$

Vamos chamar a esta grandeza *desvio médio quadrático* (muitas vezes também o chamam *desvio padrão da amostra*). Como se vê na expressão,  $\sigma_x$  caracteriza o desvio dos pontos experimentais  $x_i$  em relação à média  $\langle x \rangle$ . Porquê os desvios se somam em quadrado e não linearmente pode-se compreender no exemplo de uma distribuição simétrica. Nesta distribuição, para cada ponto desviado em relação à média de um certo valor à esquerda, existe um ponto respectivo do lado direito, de modo que ao somar os dois desvios estes se anulam  $(\langle x \rangle - x_1) + (\langle x \rangle - x_2) = 0$ . Em resultado, temos  $\sigma_x = 0$  mesmo para uma distribuição larga. Este exemplo mostra que uma simples média dos desvios não pode ser utilizada para caracterizar a largura da distribuição.

Com os dados da tabela calculamos  $\sigma_x = 0.079348538$  s. Como este valor é uma característica da largura da distribuição dos valores experimentais, também caracteriza a incerteza no resultado das medições. Para podermos dizer algo mais preciso, temos de esclarecer, antes de mais, o significado do  $\sigma_x$ .

Quando mais medidas se efectuam, o histograma de ocorrências fica cada vez mais simétrico aproximando-se à forma de um sino e, como se prova na teoria estatística, tende para a *distribuição normal*. A distribuição normal descreve-se com a função de Gauss definida como

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

É esta função que está sobreposta no histograma de ocorrências. A função de Gauss depende de dois parâmetros  $\mu$  e  $\sigma$ . É fácil de verificar (por exemplo, construindo um gráfico desta função no Excel e

Distribuição normal  
(função de Gauss)

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

fazendo variar os parâmetros) que o parâmetro  $\mu$  determina o  $x$  em que a função atinge o seu máximo, enquanto  $\sigma$  caracteriza a largura do “sino”: quanto maior for  $\sigma$  mais larga fica a distribuição. O parâmetro  $\sigma$  chama-se *desvio padrão* e o parâmetro  $\mu$  é a média da distribuição.

O gráfico da função  $g(x) = A \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ , sobreposto no histograma, foi construído com  $\mu = 0.644761905 \text{ s}$  (igual a média dos valores experimentais,  $\langle x \rangle$ ) e  $\sigma = 0.079348538 \text{ s}$  (igual a  $\sigma_x$  também calculado a partir dos valores medidos). O factor multiplicativo  $A$  é apenas o factor de escala e determina a altura do “sino”. Este foi ajustado “à mão” para melhor sobreposição com o histograma.

O facto de curva de Gauss com estes parâmetros descrever bem a distribuição dos nossos dados experimentais sugere que existe uma ligação entre as duas grandezas  $\sigma_x$  (experimental) e o parâmetro  $\sigma$  na função  $g(x)$ . O mesmo pode ser dito em relação de  $\langle x \rangle$  e  $\mu$ . A ligação, realmente, existe e isto contribui para a compreensão do significado do desvio médio quadrático  $\sigma_x$ .

As propriedades da função de Gauss estão bem estudadas. Em particular, podemos notar que a função  $g(x)$  nunca se torna igual a zero a não ser para  $x = \pm\infty$ .

No entanto, a área sob a curva, i.e.  $\int_{-\infty}^{+\infty} g(x)dx$ , é um valor finito. Se calcularmos a área

sob a curva no intervalo  $(\mu - \sigma, \mu + \sigma)$  descobrimos que esta é igual a 68.3% da área total (ver a figura). Para o intervalo  $(\mu - 2\sigma, \mu + 2\sigma)$  a área é igual a 95.5% e para o

intervalo  $(\mu - 3\sigma, \mu + 3\sigma)$  a área

é 99.7% da área total. A área

interessa-nos pela simples razão

que, se aproximássemos o

histograma de ocorrências com a

função de Gauss, a área sob a

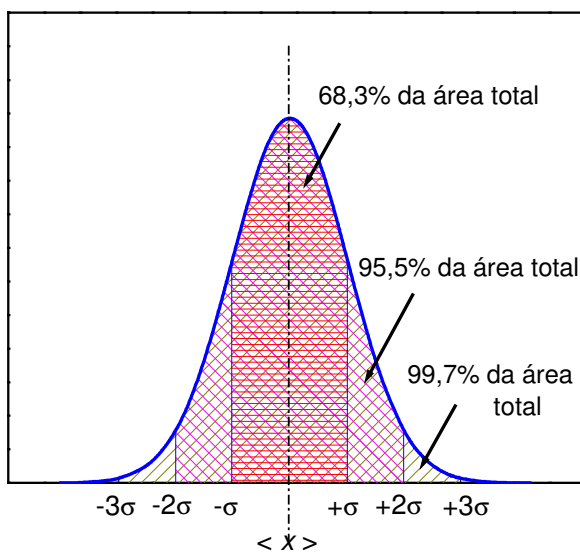
curva seria equivalente ao número

total de medidas efectuadas (a

soma dos números de ocorrências

em todos os intervalos do

histograma). Deste modo,





podemos concluir que para um grande número de medições, 68.3% dos valores obtidos serão contidos no intervalo  $(\langle x \rangle - \sigma_x, \langle x \rangle + \sigma_x)$ ; 95.5% dos valores caberão no intervalo  $(\langle x \rangle - 2\sigma_x, \langle x \rangle + 2\sigma_x)$  e 99.7% dos valores estarão dentro do intervalo  $(\langle x \rangle - 3\sigma_x, \langle x \rangle + 3\sigma_x)$ . Repare-se que os parâmetros da função  $\mu$  e  $\sigma$  (que são desconhecidos) foram aqui substituídos pelos valores  $\langle x \rangle$  e  $\sigma_x$  calculados a partir dos dados experimentais.

A importante conclusão (e a função de Gauss para isso contribuiu) é que a incerteza, que se atribui ao resultado obtido através das medições, está fortemente ligada com a probabilidade de ter (ou não ter) razão ou, como se costuma dizer, com o *nível de confiança* no resultado. Não é possível, de acordo com as leis da estatística, apresentar um intervalo de incerteza que garanta a 100% que o valor verdadeiro da grandeza medida se encontra dentro deste intervalo (simplesmente porque a curva de Gauss não tem limites). É possível afirmar apenas que o valor verdadeiro de uma grandeza encontra-se dentro de um certo intervalo de valores com um certo nível de confiança. Quanto mais elevado o nível de confiança, maior é o intervalo da incerteza.

É importante sublinhar a diferença entre  $\sigma$  e  $\sigma_x$ . Ambas as grandezas caracterizam a largura da distribuição dos valores aleatórios, e até, por vezes, a ambas se atribui o mesmo nome – desvio padrão. No entanto, há uma diferença significativa entre elas. A primeira,  $\sigma$ , é um dos dois parâmetros na função de Gauss, cujo significado é bem claro (ver a figura que mostra as áreas nas diferentes partes da curva), mas o valor exacto é desconhecido (refere-se, obviamente, à função que descreve os dados experimentais). Do outro lado está  $\sigma_x$  que é calculado a partir dos

dados experimentais, como  $\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\langle x \rangle - x_i)^2}$ , mas o valor deste

depende da amostra concreta dos valores em que se baseia o cálculo. Por exemplo, o valor de  $\sigma_x$  referido em cima foi obtido com 21 valores experimentais que estão incluídos na tabela. Mas se escolhêssemos outro número de valores da mesma tabela, o valor de  $\sigma_x$  seria diferente. Semelhantemente, se em vez de usar os 21 valores tabelados fizéssemos uma outra série de medidas, também 21, e calculássemos  $\sigma_x$  com estes valores, o  $\sigma_x$  não coincidiria com os valores anteriores. Portanto,  $\sigma_x$  não tem um valor exacto, é apenas uma estimativa, mas é a única característica da largura da distribuição a que temos acesso. Como já tivemos oportunidade de notar, com o

aumento do número de medições a distribuição dos valores experimentais tende para a distribuição de Gauss e, como na teoria estatística se mostra,  $\sigma_x \rightarrow \sigma$ .

Falta-nos um último passo para chegarmos ao nosso objectivo – a incerteza estatística do valor medido do tempo de queda. A largura da distribuição está realmente relacionada com esta incerteza mas não coincide com ela. De facto, o que nos interessa não é propriamente a largura, mas a precisão com que podemos determinar o valor médio  $\langle x \rangle$ , i.e. a posição do máximo da função de Gauss. Basta olhar para o histograma com a gaussiana sobreposta para dizer que a precisão na determinação do máximo é muito melhor que  $\sigma_x$  (temos certeza absoluta que o pico não está nem em  $\langle x \rangle - \sigma_x$  nem em  $\langle x \rangle + \sigma_x$ ). Também claro está, que quanto mais se aproxima o histograma à curva suave da gaussiana (i.e. quando mais medições se efectuam), melhor a definição do pico. A teoria estatística confirma esta observação e mostra que a incerteza na definição do pico diminui com o número de medições  $N$  como  $\frac{\sigma_x}{\sqrt{N}}$ . Vamos designar a incerteza no valor médio por  $\sigma_{\langle x \rangle}$ :  $\sigma_{\langle x \rangle} = \frac{\sigma_x}{\sqrt{N}}$ .

Deste modo, podemos representar o resultado de uma experiência em que se efectuam  $N$  medidas de uma grandeza  $x$  como

$$\tilde{x} = \langle x \rangle \pm \sigma_{\langle x \rangle}$$

em que  $\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$ ,  $\sigma_{\langle x \rangle} = \frac{\sigma_x}{\sqrt{N}}$ , e  $\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\langle x \rangle - x_i)^2}$ ,

entendendo que o significado da expressão  $\tilde{x} = \langle x \rangle \pm \sigma_{\langle x \rangle}$  é: *a melhor estimativa para o valor verdadeiro do  $x$  encontra-se no intervalo entre  $(\langle x \rangle - \sigma_{\langle x \rangle})$  e  $(\langle x \rangle + \sigma_{\langle x \rangle})$  com uma probabilidade de cerca de 68,3% (também se diz com o nível de confiança de 68,3%).*

No caso de o resultado ser apresentado em forma  $\tilde{x} = \langle x \rangle \pm 2\sigma_{\langle x \rangle}$  pode-se dizer que *a melhor estimativa para o valor verdadeiro do  $x$  encontra-se no respectivo intervalo com uma probabilidade de cerca de 95,4%*, e, no caso de  $\tilde{x} = \langle x \rangle \pm 3\sigma_{\langle x \rangle}$ , - *a melhor estimativa para o valor verdadeiro do  $x$  encontra-se no respectivo intervalo com uma probabilidade de cerca de 99,7%.*

Para a nossa experiência, em que foi medido o tempo da queda da bola de ténis, isto traduz-se no seguinte resultado:  $\sigma_{\langle x \rangle} = \frac{0.079348538}{\sqrt{21}} = 0.01731527 \text{ s}$  e, então

$$\tilde{t} = (0.645 \pm 0.017) \text{ s, com o nível de confiança de } 68,3\%;$$

$$\tilde{t} = (0.645 \pm 0.035) \text{ s, com o nível de confiança de } 95,5\%;$$

$$\tilde{t} = (0.645 \pm 0.052) \text{ s, com o nível de confiança de } 99,7\%.$$

Repare no número de casas significativas com que o resultado foi representado. Como o valor da incerteza é apenas uma estimativa que depende da amostra, não faz sentido escrevê-la com muita precisão. Normalmente considera-se que uma ou duas casas significativas são suficientes. Ao representar a incerteza com 3 casas decimais, como fizemos, obriga-nos a escrever, também, a estimativa para o tempo da queda com 3 casas decimais (como a segunda casa decimal já é incerta, não faz sentido acrescentar mais; por outro lado, representar o resultado com menor precisão é pouco razoável, também, porque assim pioramos nós próprios o resultado da experiência).

Ao comparar com a precisão do cronómetro concluímos que a incerteza devido à dispersão estatística das medidas é superior à contribuição do instrumento de medida ( $\pm 0,005 \text{ s}$ ). Portanto, se pretendermos melhorar o resultado, não vale a pena (para já, pelo menos) procurar um aparelho melhor para medir o tempo. Antes de mais, tem que ser melhorado o procedimento que, como sabemos, é a causa da dispersão aleatória dos valores medidos.

O resultado que em cima apresentamos não seria totalmente honesto se não incluíssemos na incerteza o erro instrumental, embora este seja pequeno. A maneira mais correcta de o fazer é somar as incertezas em quadrado

$$\Delta t_{\text{tot}}^2 = \sigma_{\text{aleatória}}^2 + \Delta t_{\text{instrumental}}^2.$$

Há duas observações que devem ser feitas a cerca desta fórmula. A primeira é que o conceito de desvio padrão só é aplicável a um processo aleatório. A imprecisão instrumental (pelo menos ao nível do display do cronómetro, que só tem duas casas decimais) não revela este tipo de comportamento. Esta é a razão porque não designamos a contribuição do instrumento por um  $\sigma$  mas por um  $\Delta t$ , que não tem o significado do desvio padrão. Pela mesma razão, a fórmula em cima não está absolutamente correcta e utiliza-se apenas por falta de uma melhor...

No nosso caso concreto,  $\Delta t_{tot} \approx 0.18 \text{ s}$ , de modo que  $\tilde{t} = (0.645 \pm 0.018) \text{ s}$ , se usássemos o resultado para o tempo em forma  $\langle t \rangle \pm \sigma_{\langle t \rangle}$ . Todavia, temos de ter a consciência que depois de adicionar o  $\Delta t_{instrumental}$  o significado do nível de confiança deixa de ser muito preciso. Por isso, não é suficiente mencionar apenas a incerteza. Tem que se especificar o que é que esta inclui.

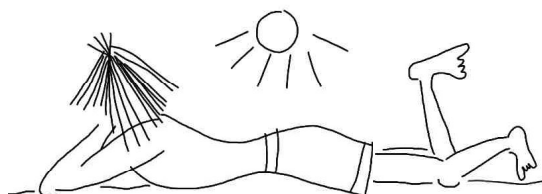
Será que existem outros factores que influenciam os resultados das medidas? Se pretendêssemos determinar a aceleração da gravidade a partir da medição do tempo de queda seria importante assegurar que o próprio processo de queda não seria afectado por factores exteriores que não fazem parte dos objectivos do estudo. A resistência do ar é um desses factores que não se pode incluir, nem na imprecisão instrumental, nem no erro aleatório. O efeito da resistência do ar não faz com que os resultados das medidas sofram uma dispersão ora para um lado ora para o outro, como no caso de um processo aleatório, mas resulta num sistemático aumento do tempo em todas as medidas. Os erros deste tipo chamam-se *erros sistemáticos*. Em princípio, as fontes dos erros sistemáticos devem ser identificadas e eliminadas, tal como no nosso exemplo, podíamos eliminar o efeito da resistência do ar ao efectuar a experiência no vácuo. No entanto, isto nem sempre se consegue. Muitas vezes é difícil eliminar completamente os erros sistemáticos e mesmo identificá-los pode não ser possível.

$$\tilde{x} = \langle x \rangle \pm \sigma_{\langle x \rangle}$$

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\sigma_{\langle x \rangle} = \frac{\sigma_x}{\sqrt{N}}$$

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\langle x \rangle - x_i)^2}$$



Para o nível de confiança de 68,3%  $\tilde{x} = \langle x \rangle \pm \sigma_{\langle x \rangle}$

Para o nível de confiança de 95,4%  $\tilde{x} = \langle x \rangle \pm 2\sigma_{\langle x \rangle}$

Para o nível de confiança de 99,7%  $\tilde{x} = \langle x \rangle \pm 3\sigma_{\langle x \rangle}$

#### Nomenclatura:

$N$  – número de medidas

$\tilde{x}$  – melhor estimativa para o valor verdadeiro do  $x$

$\langle x \rangle$  – média aritmética dos valores experimentais

$\sigma_{\langle x \rangle}$  – desvio padrão da média

$\sigma_x$  – desvio médio quadrático dos valores experimentais da sua média

$\mu$  – parâmetro da função de Gauss que caracteriza a posição do máximo (também a média) da distribuição normal

$\sigma$  – desvio padrão na função de Gauss (caracteriza a largura da distribuição normal)

Em qualquer experiência há sempre uma possibilidade de existir um ou vários factores desconhecidos que podem afectar o resultado de uma forma significativa. Por isso, nunca podemos confiar totalmente no resultado de uma experiência isolada. É absolutamente necessário que o mesmo resultado seja confirmado noutras condições experimentais, com outros instrumentos, utilizando outros métodos etc.

Uma última observação que é necessário fazer é que em algumas situações não é fácil determinar a incerteza com todo o rigor. Por vezes não temos outra hipótese a não ser invocar a intuição e o senso comum. Por exemplo, no caso da queda da bola, precisamos também de medir a altura  $h$  para poder determinar a aceleração da gravidade. É fácil medir uma distância de cerca de  $2\text{ m}$  entre o chão e um prego na parede com uma precisão de ordem de  $1\text{ mm}$ , utilizando uma fita métrica. A questão é definir entre que pontos temos de medir. A bola de tennis tem cerca de  $6\text{ cm}$  de diâmetro. Poder-se-ia sugerir medir a distância do chão até ao centro da bola. Esta podia ser medida com bastante precisão, mas, neste caso, não serve. O método que utilizamos para detectar o instante em que a bola colide com o chão é através do sinal sonoro produzido quando a colisão ocorre. Isto leva-nos a concluir que o ponto de referência devia ser a face inferior da bola e não o centro. Mais – em que instante o som é produzido? Quando a bola toca no chão com o seu ponto extremo ou um pouco mais tarde? O som, como se sabe, é uma compressão (ou decompressão) do meio que, uma vez produzida, se propaga neste meio. No nosso caso, a deformação da bola ao bater no chão produz esta compressão. Portanto, também o ponto mais extremo do lado inferior da bola não serve como referência. Esta deve estar algures dentro da bola. Uma simples observação da bola quando esta bate no chão permite verificar que a deformação da bola é de ordem de  $0,5\text{ cm}$  a  $1\text{ cm}$ . Assim, podemos atribuir ao valor de  $h$  uma incerteza estimada por  $\Delta h \approx 1\text{ cm}$ . É uma estimativa, de certo modo subjectiva, mas pelo menos dá-nos uma ideia sobre o valor da incerteza. Como vamos ver na secção seguinte, a contribuição do  $\Delta h \approx 1\text{ cm}$  para a incerteza total na melhor estimativa para a aceleração  $g$  é pequena em comparação com a incerteza devido a dispersão dos valores do  $t$  e, portanto, esta avaliação subjectiva para  $\Delta h$  é suficiente.

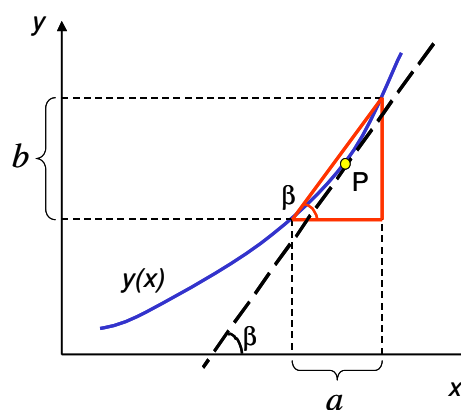
### 3. Propagação de erros

Medir uma grandeza física directamente raramente se consegue. Mais frequentemente efectua-se a medição de uma ou várias outras grandezas e a partir destas calcula-se a que interessa. No nosso exemplo de queda da bola, o objectivo final era determinar a aceleração da gravidade. Para isso, medimos o tempo da queda  $t$  e a altura  $h$ , cada uma com a sua incerteza. Sabendo essas duas grandezas podemos chegar à aceleração da gravidade utilizando a equação  $g = \frac{2h}{t^2}$ . A questão que surge é qual será a incerteza no  $g$ ?

Suponhamos para simplificar que a grandeza que se procura,  $y$ , é uma função de uma única variável  $y = y(x)$ . Suponhamos, também, que o  $x$  foi medido com uma incerteza  $\Delta x$ . O problema em causa é, então, a seguinte: se a melhor estimativa para o valor verdadeiro do  $x$  for  $\tilde{x} = \langle x \rangle \pm \Delta x$ , qual a melhor estimativa para o valor verdadeiro do  $y$ ? É mais ou menos intuitivo que podemos escrever uma coisa semelhante para  $\tilde{y}$ :  $\tilde{y} = \langle y \rangle \pm \Delta y$  em que  $\langle y \rangle$  calcula-se a partir do  $\langle x \rangle$ . Mas qual a incerteza  $\Delta y$ ?

A resposta vem automaticamente se representarmos a função  $y(x)$  graficamente.

Como se vê na figura, um intervalo  $a$  no eixo  $x$  traduz-se num intervalo  $b$  no eixo  $y$  de modo que  $b = a \cdot \tan \beta$ . A tangente do ângulo  $\beta$  pode ser determinada lembrando a interpretação geométrica da derivada de uma função. Traçamos uma linha recta (a tracejado no gráfico) paralela à hipotenusa



do triângulo a vermelho de tal maneira que esta seja tangente à curva num ponto qualquer dentro do intervalo  $a$  (ponto  $P$  no gráfico). Isto é sempre possível desde que a função seja monótona neste intervalo. O ângulo que a recta a tracejado forma com o eixo  $x$  é igual a  $\beta$  e  $\tan \beta = \frac{dy}{dx}$  no ponto  $P$  (esta é a interpretação geométrica da

derivada no ponto  $P$ ). Assim,  $b = a \cdot \left. \frac{dy}{dx} \right|_P$ .

Se  $a$  for o intervalo que caracteriza a incerteza em  $x$ , a incerteza no  $y(x)$ , então pode ser determinada com a fórmula  $\Delta y = \Delta x \cdot \left. \frac{dy}{dx} \right|_P$ . O problema desta fórmula é que não sabemos onde está o ponto  $P$ . Claro, que ao sabermos a forma exacta da função  $y(x)$ , em princípio, podemos encontrá-lo. No entanto, para muitos efeitos é suficiente supor que a abcissa do ponto  $P$  coincide com o meio do intervalo, i.e com  $\langle x \rangle$ . Isto vai dar-nos um valor aproximado do  $\Delta y$  mas não esqueçamos que a incerteza em  $x$  também é apenas uma estimativa:

$$\Delta y \approx \Delta x \cdot \left. \frac{dy}{dx} \right|_{x=\langle x \rangle}.$$

(À mesma conclusão podemos chegar utilizando a definição da derivada:  $\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} \Rightarrow \frac{dy}{dx} \approx \frac{\Delta y}{\Delta x} \Rightarrow \Delta y \approx \Delta x \cdot \frac{dy}{dx}$ ). Como o resultado será escrito em forma  $\langle y \rangle \pm \Delta y$ , o sinal de  $\Delta y$  não nos interessa e, portanto, usamos a derivada em módulo.

Na nossa experiência, supondo por um instante que a altura é conhecida com exactidão, o  $x$  será o tempo da queda  $t$  e a aceleração da gravidade  $g$  será o  $y$ . Então, a incerteza em  $g$  será igual a  $\Delta g \approx \Delta t \cdot \left. \frac{dg(t)}{dt} \right|_{t=\langle t \rangle} = \Delta t \cdot \left. -2 \cdot \frac{2h}{t^3} \right|_{t=\langle t \rangle} = 4h \frac{\Delta t}{\langle t \rangle^3}$  (ou, numa forma mais “simpática”,  $\frac{\Delta g}{g} \approx 2 \frac{\Delta t}{t}$  - o erro relativo em  $g$  é o dobro do erro relativo em  $t$ ).

O passo seguinte é generalizar a fórmula para  $\Delta y$  para o caso de  $y$  ser uma função de várias variáveis  $y = y(x_1, x_2, \dots)$  (tal como no nosso exemplo:  $g$  é uma função do  $t$  mas, também, do  $h$ ). A contribuição das flutuações em cada variável pode ser calculada da mesma maneira, mas com um pormenor – quando se calcula a derivada em  $x_i$ , todas as outras variáveis se consideram constantes:

$$\Delta y_i \approx \Delta x_i \cdot \left. \frac{\partial y}{\partial x_i} \right|_{x_i=\langle x_i \rangle} \quad (\text{matematicamente, esta condição é reflectida através da}$$

substituição da derivada completa pela *derivada parcial*, cujo significado é precisamente este: calcular a derivada da função  $y = y(x_1, x_2, \dots)$  em ordem do  $x_i$  considerando todas as outras variáveis como constantes).

Como as contribuições dos diferentes factores para a incerteza total se somam em quadrados, temos

$$\Delta y^2 \approx \left| \frac{\partial y}{\partial x_1} \right|^2 \Delta x_1^2 + \left| \frac{\partial y}{\partial x_2} \right|^2 \Delta x_2^2 + \dots$$

No caso da nossa experiência, isto traduz-se no seguinte

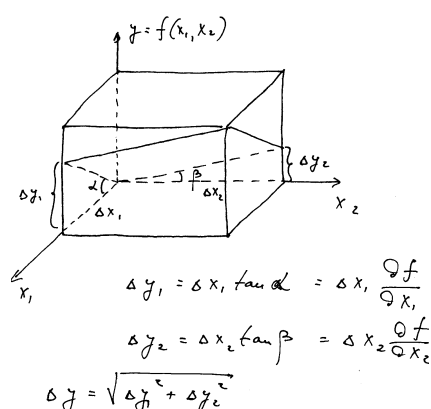
$$\Delta g^2 \approx \left| \frac{\partial g(t, h)}{\partial t} \right|^2 \Delta t^2 + \left| \frac{\partial g(t, h)}{\partial h} \right|^2 \Delta h^2 = \left( \frac{4h}{t^3} \Delta t \right)^2 + \left( \frac{2}{t^2} \Delta h \right)^2 \text{ ou, em termos dos erros}$$

relativos,  $\left( \frac{\Delta g}{g} \right)^2 = \left( 2 \frac{\Delta t}{t} \right)^2 + \left( \frac{\Delta h}{h} \right)^2$ . Nestas expressões, os valores de  $t$  e de  $h$  são os

que representam as melhores estimativas para as respectivas grandezas. Substituindo os valores  $t = 0.645 \text{ s}$ ,  $\Delta t = 0.018 \text{ s}$ ,  $h = 2.00 \text{ m}$  e  $\Delta h = 0.01 \text{ m}$ , temos  $2 \frac{\Delta t}{t} \approx 0.0558$  e

$\frac{\Delta h}{h} = 0.005$  de modo que  $\frac{\Delta g}{g} \approx 0.0560$ . Daqui se

vê que a contribuição da incerteza na altura da queda é insignificante para o resultado final. Isto justifica a nossa abordagem para o cálculo das incertezas: na secção anterior calculamos  $\Delta t$  com bastante rigor, mas para  $\Delta h$  utilizamos uma grosseira estimativa subjectiva. Mesmo se  $\Delta h$  fosse por nós subestimada e na realidade fosse 2 ou 3 vezes maior, não podia afectar o resultado significativamente.



Finalmente, temos a melhor estimativa do valor verdadeiro da aceleração da gravidade:  $g = (9.61 \pm 0.54) \text{ m/s}^2$ , o que está de acordo, dentro dos erros experimentais, com o valor esperado de  $9.81 \text{ m/s}^2$ .



## Capítulo 2

# Revisão da mecânica Newtoniana

## 1. Cinemática

O objectivo da cinemática é a descrição completa do movimento (posição em função do tempo) de um corpo em função do tempo. A pergunta que se coloca aqui é *como* se move um corpo, sem preocupação pelas causas desse movimento (i.e. *porque* o corpo se move dessa maneira e não de outra; a resposta a esta pergunta procura-se na *dinâmica*). Assume-se que a forma e as dimensões do corpo não são importantes e, portanto, este é considerado como sendo um ponto material. A possível rotação do corpo durante o movimento ignora-se também (não faz sentido falarmos de rotação de um ponto).

*Como ?* → Cinemática  
*Porque ?* → Dinâmica

“Descrever a trajectória” significa saber, em cada instante de tempo, a posição exacta do corpo (muitos vezes fala-se de uma partícula para sublinhar a suposição de ausência das dimensões do corpo), i.e.

$x(t)$  – no caso de movimento unidimensional

ou, no caso mais geral,

$\mathbf{r}(t)$  - se a partícula se move num plano ou num espaço de três dimensões ( $\mathbf{r}(t)$  chama-se vector posição).

Com a informação completa sobre a posição de uma partícula em função do tempo,  $\mathbf{r}(t)$ , podemos facilmente saber a sua velocidade,  $\mathbf{v}(t)$ , e a aceleração,  $\mathbf{a}(t)$ , em qualquer instante de tempo. Por definição temos

A escrita  $\mathbf{r}$  significa um vector com dois (no plano) ou três (no espaço) componentes:  
 $\mathbf{r}(t) = \{x(t), y(t), z(t)\}$ .

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt}$$

$$\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt} \Rightarrow \mathbf{a}(t) = \frac{d}{dt} \left( \frac{d\mathbf{r}(t)}{dt} \right) = \frac{d^2\mathbf{r}(t)}{dt^2}, \quad \text{numa dimensão}$$

e

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt}$$

$$\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt} \Rightarrow \mathbf{a}(t) = \frac{d}{dt} \left( \frac{d\mathbf{r}(t)}{dt} \right) = \frac{d^2\mathbf{r}(t)}{dt^2}, \text{ no caso de duas ou três}$$

dimensões.

Então, se sabemos o vector posição  $\mathbf{r}(t)$  num certo intervalo de tempo  $(t_1, t_2)$ , podemos, através da derivação, obter a velocidade e a aceleração do corpo neste intervalo de tempo. Por exemplo, efectuando as medições da posição da Lua em relação à Terra durante uma semana podemos calcular com este método  $\mathbf{v}(t)$  e  $\mathbf{a}(t)$  do nosso satélite, mas só durante essa semana.

$$\mathbf{r}(t) \Rightarrow \mathbf{v}(t) \Rightarrow \mathbf{a}(t)$$

A informação a cerca do  $\mathbf{r}(t)$  neste intervalo de tempo não nos permite, em geral, fazer uma previsão para a trajectória da Lua no futuro. Porquê? Porque a cinemática só se preocupa com a descrição do próprio movimento em si e não se interessa pelas factores que o afectam, i.e., pelas forças que actuam sobre o corpo. No nosso exemplo, é a força gravítica exercida sobre a Lua, principalmente pela Terra, que influencia a trajectória da Lua.

Mas, força não faz parte da linguagem da cinemática... Este já é assunto da dinâmica de que vamos falar, com mais pormenor, na secção seguinte. Para já lembremo-nos apenas que a lei fundamental da dinâmica (2ª lei de Newton) relaciona a força que actua sobre o corpo com a aceleração deste através da equação  $\mathbf{F} = m\mathbf{a}$ ,



A derivada de um vector, por exemplo  $\frac{d\mathbf{r}(t)}{dt}$ , é um outro vector cujas componentes são as derivadas das respectivas componentes do vector original:

$$\frac{d\mathbf{r}(t)}{dt} = \left\{ \frac{dx(t)}{dt}, \frac{dy(t)}{dt}, \frac{dz(t)}{dt} \right\} \text{ e como } \mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} \text{ e } \mathbf{v}(t) = \{v_x(t), v_y(t), v_z(t)\}$$

$$\text{concluimos que } v_x(t) = \frac{dx(t)}{dt}, v_y(t) = \frac{dy(t)}{dt} \text{ e } v_z(t) = \frac{dz(t)}{dt}.$$

Da mesma maneira para o vector aceleração,  $\mathbf{a}(t) = \{a_x(t), a_y(t), a_z(t)\}$ , escrevemos

$$\mathbf{a}(t) = \left\{ \frac{dv_x(t)}{dt}, \frac{dv_y(t)}{dt}, \frac{dv_z(t)}{dt} \right\} \text{ ou } \mathbf{a}(t) = \left\{ \frac{d^2x(t)}{dt^2}, \frac{d^2y(t)}{dt^2}, \frac{d^2z(t)}{dt^2} \right\}.$$

em que  $m$  é a massa do corpo. No exemplo com a Lua, esta força é bem conhecida e é a força gravítica. Ao saber a força em função do tempo, pode-se calcular a aceleração,

$$\mathbf{a}(t) = \frac{\mathbf{F}(t)}{m}.$$

Sabendo a aceleração podemos chegar à velocidade utilizando a operação inversa a derivação, i.e. calculando a primitiva

$$\mathbf{v}(t) = \int \mathbf{a}(t) dt$$

e, assim continuando, ao vector posição

$$\mathbf{r}(t) = \int \mathbf{v}(t) dt.$$

$$\mathbf{F}(t) \Rightarrow \mathbf{a}(t) \Rightarrow \mathbf{v}(t) \Rightarrow \mathbf{r}(t)$$

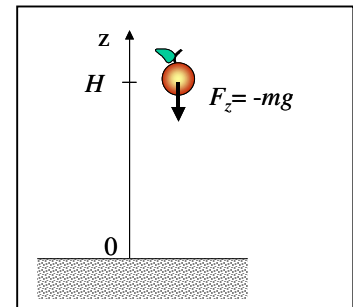
### ***Exemplo. Queda de uma maçã.***

#### Problema.

Uma maçã cai de um árvore de uma altura  $H$  sem velocidade inicial. Determine a posição da maçã em função do tempo.

#### Solução.

A única força, que actua sobre a maçã, é a força gravítica. Esta força é constante em módulo, actua na vertical e tem sentido para baixo. Escolhemos os eixos de tal maneira que o plano  $(x, y)$  corresponde ao solo e o eixo  $z$  aponta para cima com origem no chão. Como a velocidade inicial da maçã é nula e não existe nenhuma componente da força a actuar no plano



$(x, y)$ , estamos perante um movimento unidimensional ao longo do eixo  $z$ .

Os componentes da força são  $\vec{F} = \{0, 0, F_z\}$  em que  $F_z < 0$  porque o vector força aponta para baixo, i.e. no sentido contrário ao eixo  $z$ . De acordo com a 2ª lei de

Newton  $\mathbf{a}(t) = \frac{\mathbf{F}(t)}{m}$  e, então, o vector aceleração  $\mathbf{a} = \{0, 0, a_z\}$  com  $a_z = \frac{F_z}{m}$  ( $m$  é a

massa da maçã). A força gravítica perto da superfície da Terra, como se sabe, é igual a

$\mathbf{F} = m\mathbf{g}$  e, então,  $F_z = mg_z = -mg$ . Para a componente  $z$  da aceleração temos

$a_z = -g$ , i.e. a aceleração é constante como era de esperar.

A velocidade da maçã calcula-se através da primitivação da aceleração

$$v_z(t) = \int a_z(t) dt = \int (-g) dt = -g \int dt = -gt + C_1 ,$$

onde  $C_1$  é a constante de integração. Para determinar a constante, precisamos de mais uma equação. Neste caso, sabemos que a velocidade inicial é nula, i.e.

$$\text{em } t = 0: v_z(0) = -g \cdot 0 + C_1 = C_1 \Rightarrow C_1 = 0$$

$$\text{Assim, } v_z(t) = -gt .$$

Para a posição  $z$  da maçã em função do tempo temos

$$z(t) = \int v_z(t) dt = \int (-gt) dt = -g \int t dt = -g \frac{t^2}{2} + C_2$$

com a constante de integração que se determina pela posição inicial da maçã,  $z = H$ :

$$\text{em } t = 0: z(0) = -g \cdot \frac{0}{2} + C_2 = C_2 \Rightarrow C_2 = H .$$

Temos, finalmente, a posição da maçã em função do tempo

$$z(t) = H - \frac{gt^2}{2} .$$

Se fizermos o mesmo cálculo para as componentes  $x$  e  $y$  da velocidade, verificamos que estas são nulas. Por exemplo,  $v_x(t) = \int a_x(t) dt = \int 0 \cdot dt = C_1$  e, como  $v_x(0) = 0$ ,  $C_1 = 0$  o que significa que  $v_x(t) = 0$  em qualquer instante de tempo. Da mesma maneira mostra-se que  $x(t) = 0$  e, também,  $v_y(0) = 0$  e  $y(t) = 0$ .

---

## 2. Dinâmica

### 2.1 Força

Na secção anterior referimos duas abordagens diferentes para determinar as características do movimento, baseadas na informação inicial que temos ao nosso dispor. No primeiro caso, determinamos as características do movimento (a velocidade e a aceleração) a partir do conhecimento da posição da partícula, medida num dado intervalo de tempo  $(t_1, t_2)$ . Na segunda situação, fazemos o inverso, tendo a força (ou a aceleração, porque  $\mathbf{F} = m\mathbf{a}$ ) como o ponto de partida. Se no primeiro exemplo não nos foi possível obter  $v(t)$  e  $a(t)$  fora do intervalo de tempo  $(t_1, t_2)$ , no segundo caso a previsão do movimento no futuro (e no passado remoto) torna-se

possível porque as propriedades da força gravítica são bem conhecidas e pode-se saber, em princípio, como a força varia ao longo da trajectória da partícula, i.e. em qualquer instante de tempo.

$$\mathbf{F}(t) \Rightarrow \mathbf{a}(t) \Rightarrow \mathbf{v}(t) \Rightarrow \mathbf{r}(t) \text{ versus } \mathbf{r}(t) \Rightarrow \mathbf{v}(t) \Rightarrow \mathbf{a}(t)$$

Força é um conceito fundamental da mecânica clássica. E, por isso mesmo, por ser um conceito primário em que se baseia a teoria, não é fácil dar uma definição exacta do que é *força*, embora todos nós tenhamos uma noção intuitiva do que é (uma expressão do tipo “puxa com mais força!” toda a gente entende). O mais que se pode dizer, provavelmente, é que *força é uma medida de acção de um corpo sobre o outro*. Uma força caracteriza-se pela intensidade, sentido (isto quer dizer que é um vector) e pelo ponto em que se aplica.

## 2.2 A lei fundamental da dinâmica

A lei fundamental da dinâmica, também conhecida como segunda lei de Newton, relaciona a aceleração de um corpo com a força exercida sobre ele. A forma mais comum desta lei é

$$\mathbf{F} = m\mathbf{a}.$$

A lei afirma que a aceleração do corpo é directamente proporcional à força. Se não existisse nenhuma força a actuar sobre o corpo (i.e.  $\mathbf{F} = 0$ ), o corpo movia-se sem aceleração (i.e.  $\mathbf{a} = 0$ ). Como  $\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt}$  isso significa que  $\mathbf{v}(t) = \text{const}$  (mas não necessariamente igual a 0). Esta conclusão é muito importante. A segunda lei de Newton diz-nos que um carrinho pode mover-se sem ser puxado por ninguém. A existência de uma força não é uma condição necessária para existir o movimento. A força somente faz com que o vector velocidade varie.

### **1ª Lei de Newton (lei de inércia):**

Na ausência de forças, um corpo permanece em repouso ou em movimento rectilíneo uniforme (i.e.,  $\mathbf{v}(t) = \text{const}$ , mas não necessariamente igual a 0).

### **3ª Lei de Newton (acção = contra-acção):**

A força que um corpo exerce sobre outro corpo é igual em módulo e tem sentido oposto à força com que o segundo corpo actua sobre o primeiro.

Na altura em que Newton fez essa observação, dominava o ponto de vista de Galileu que afirmava que a força era necessária para que um corpo se movesse. Esta visão do Galileu está, de facto, de acordo com a nossa experiência do dia a dia. Precisamos de empurrar constantemente o carrinho de compras para chegar com ele a um lado qualquer. Mas também é verdade que se, uma vez em movimento, deixarmos de o empurrar, o carrinho continuará mover-se por uns instantes. A explicação correcta dessa situação é a do Newton: precisamos, sim, de exercer a força para mover o carrinho com uma velocidade constante, mas somente por causa da força de atrito que se opõe ao movimento. Por isso é que é necessária uma força - a força dos nossos músculos - a actuar no sentido do movimento e igual em módulo à força de atrito. Deste modo, a força resultante ( $F_{res} = F_{atrito} + F_{músculos}$ ) é igual a zero e, de acordo com Newton, a aceleração é nula e o vector velocidade é constante.

Deste exemplo com um carrinho, torna-se claro o que se deve escrever do lado esquerdo da 2ª lei de Newton – temos que lá pôr a força resultante, i.e. a soma vectorial de todas as forças a actuar sobre o corpo

$$\sum_i F_i = ma .$$

Mais uma observação importante que se pode fazer é acerca da massa. Intuitivamente todos sabemos o que é a massa, mas defini-la com rigor já é mais complicado. Na segunda lei de Newton a massa surge como coeficiente de proporcionalidade,  $m$ , entre a força e a aceleração e essa é a sua definição exacta. A mesma força resulta em menor aceleração quando a massa é maior e *vice-versa*. Costuma-se dizer que um corpo com uma massa maior tem uma inércia maior. A massa, portanto, é uma medida de *inércia* do corpo.

Na grande maioria das situações práticas a massa do corpo em movimento mantém-se constante ou, pelo menos, a variação da massa em função do tempo é desprezável (como, por exemplo, a massa do carro em viagem, embora a massa de combustível no depósito diminua). Nestas circunstâncias, a 2ª lei de Newton na forma  $F = ma$  aplica-se directamente. Existem, porém, casos em que a massa não pode ser considerada constante. Um lançador espacial seria um bom exemplo. A massa do combustível, dos depósitos e dos propulsores, que se separam da cápsula durante o lançamento, atingem 99% da massa inicial do lançador. A forma da 2ª lei de Newton que tem em conta as situações em que  $m \neq const$  é a seguinte

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}$$

onde  $\mathbf{p}$  é o momento linear. Substituindo  $\mathbf{p}$  por sua definição,  $\mathbf{p} = m\mathbf{v}$ , e utilizando as regras de derivação temos

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d(m\mathbf{v})}{dt} = \frac{dm}{dt}\mathbf{v} + m\frac{d\mathbf{v}}{dt} = \frac{dm}{dt}\mathbf{v} + m\mathbf{a}.$$

Vê-se, com esta expressão, que para além do componente habitual  $m\mathbf{a}$ , aparece mais um termo do mesmo lado da equação que diz respeito à variação da massa em função do tempo,  $\frac{dm}{dt}\mathbf{v}$ . Se a massa for constante,  $\frac{dm}{dt} = 0$  e a 2ª lei de Newton toma a sua forma mais bem conhecida,  $\mathbf{F} = m\mathbf{a}$ , o que era de esperar.

A lei fundamental da dinâmica na sua forma mais geral, portanto, é  $\sum_i \mathbf{F}_i = \frac{d\mathbf{p}}{dt}$ .

**A forma geral da 2ª lei de Newton:**  $\sum_i \mathbf{F}_i = \frac{d\mathbf{p}}{dt}$ .

## 2.3 O círculo vicioso da mecânica

Na grande parte dos problemas de mecânica procura-se saber a trajectória de um corpo, i.e. a posição deste em função do tempo,  $\mathbf{r}(t)$ . Já apontámos o caminho conveniente:  $\mathbf{F}(t) \Rightarrow \mathbf{a}(t) \Rightarrow \mathbf{v}(t) \Rightarrow \mathbf{r}(t)$ . Neste

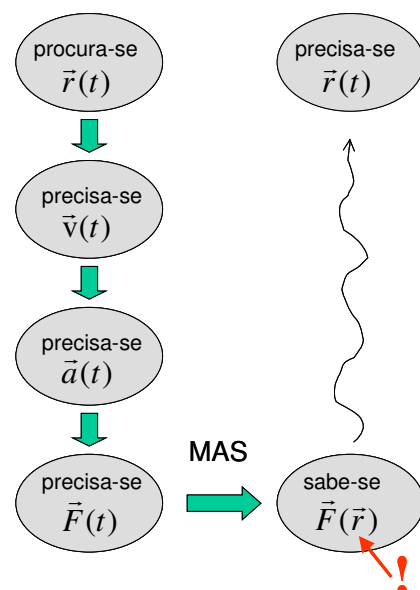
caminho utilizamos, simplesmente, i) as definições da velocidade e da aceleração como derivadas do  $\mathbf{r}(t)$  em ordem ao tempo,

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} \quad \text{e} \quad \mathbf{a}(t) = \frac{d^2\mathbf{r}(t)}{dt^2}; \quad \text{ii) a lei}$$

fundamental da dinâmica que permite estabelecer a ligação entre os parâmetros cinemáticos e a

força,  $\mathbf{a}(t) = \frac{\mathbf{F}(t)}{m}$  e iii) o facto de em muitas

situações as forças a actuar no sistema serem conhecidas. Parece estar resolvido o problema !



A desilusão vem quando nos apercebemos que precisamos de primitivar em ordem ao tempo, mas para isso necessitamos de saber, também, a força em função do tempo. No entanto, não é isso que se sabe... Sabemos geralmente a força, não em função do tempo, mas sim em função da posição. Como, por exemplo, no caso da força gravítica que determina o movimento dos corpos celestes,  $\mathbf{F}(\mathbf{r})$ . Ao percorrermos o caminho todo voltamos ao ponto de partida: para determinar  $\mathbf{r}(t)$  precisamos de saber  $\mathbf{r}(t)$ .

Há duas maneiras de quebrar este círculo vicioso. A primeira é compor uma equação diferencial usando a 2ª lei de Newton e tentar resolvê-la:

$m\mathbf{a} = \mathbf{F} \Rightarrow m \frac{d^2 \mathbf{r}}{dt^2} = \mathbf{F}(\mathbf{r})$ , sendo  $\mathbf{r}(t)$  a função incógnita (no caso de movimento

unidimensional temos  $m \frac{d^2 x}{dt^2} = F(x)$  com incógnita  $x = x(t)$ ). Em algumas situações

esta equação diferencial da segunda ordem pode ser resolvida, dependendo muito da forma da função  $\mathbf{F}(\mathbf{r})$ . Como vamos ver no capítulo seguinte, este é um bom caminho para descrever o movimento oscilatório, por exemplo. No caso em que não existe atrito e a única força a actuar no sistema é a força elástica  $F(x) = -kx$  (um bloco fixo numa das extremidades de uma mola a oscilar no plano horizontal), a

equação toma uma forma bastante simples:  $m \frac{d^2 x}{dt^2} = -kx$ , sendo  $k$  a constante elástica

da mola. Esta equação pode ser resolvida analiticamente resultando numa solução

$x(t) = A \cdot \cos\left(\sqrt{\frac{k}{m}} \cdot t + \delta\right)$  em que  $A$  e  $\delta$  são constantes de integração que só podem

ser determinadas com base numa informação adicional (o estado do sistema num determinado instante de tempo).

Existe, no entanto, um outro caminho que, embora não resolva o problema na totalidade, leva à lei de conservação da energia mecânica – uma das mais poderosas leis de conservação que conhecemos. Este caminho, que vamos explorar na secção seguinte, passa pelo conceito de trabalho de uma força.



### 3. Trabalho e Energia

#### 3.1 Trabalho de uma força. Energia cinética.

Vejamos se se consegue aproveitar de alguma maneira o facto de se conhecer  $\mathbf{F}(\vec{r})$  (e não  $\mathbf{F}(t)$ ). Definimos uma grandeza física, chamada *trabalho de uma força*, do seguinte modo. O trabalho realizado por uma força  $\mathbf{F}$  ao longo de uma trajectória, que vamos designar por  $L$ , com início no ponto  $A$  e fim no ponto  $B$  define-se como

$$W_{AB}(L) = \int_L \mathbf{F} d\mathbf{r}.$$

A expressão  $\mathbf{F} d\mathbf{r}$  significa produto escalar (também se chama produto interno) dos vectores  $\mathbf{F}$  e  $d\mathbf{r}$  e define-se como  $\mathbf{F} d\mathbf{r} = \|\mathbf{F}\| \cdot \|d\mathbf{r}\| \cdot \cos \alpha$ , onde  $\alpha$  é o ângulo entre os dois vectores (ver a caixa). O símbolo  $L$  debaixo do símbolo de integral significa que a integração (pode ser substituída por um somatório – ver a caixa) deve ser feita ao longo da curva que descreve a trajectória da partícula. Afirma-se, por esse meio, que o resultado da integração, em geral, depende da forma particular da trajectória que a partícula segue no caminho do ponto  $A$  para o ponto  $B$  e não somente das coordenadas destes dois pontos.

Repare, que a integração no cálculo de  $W$  faz-se em ordem de  $\mathbf{r}$  e não em ordem ao tempo como precisávamos fazer na secção anterior. Esta é uma grande vantagem do conceito de trabalho de força porque, como já observamos, a força é normalmente conhecida em função da posição e não do tempo.

Também se pode escrever a definição do trabalho em forma diferencial, i.e. para um deslocamento infinitamente pequeno,  $d\mathbf{r}$ :

$$dW = \mathbf{F}(\mathbf{r}) d\mathbf{r}$$

(se ainda é confusa esta forma de escrita, poder-se-ia pensar em termos de deslocamentos finitos,  $\Delta\mathbf{r}$  em vez de  $d\mathbf{r}$ ; assim, a expressão em cima transforma-se em  $\Delta W = \mathbf{F}(\mathbf{r}) \Delta\mathbf{r}$ ).

Aplicando a 2ª lei de Newton  $\mathbf{F} = m \frac{d\mathbf{v}}{dt}$  (vamos considerar  $m = \text{const}$ ) e usando a definição do trabalho, temos

$$W_{AB}(L) = \int_L \mathbf{F} d\mathbf{r} = \int_L m \frac{d\mathbf{v}}{dt} d\mathbf{r} = m \int_L \frac{d\mathbf{v}}{dt} d\mathbf{r} = m \int_L \frac{d\mathbf{v} \cdot d\mathbf{r}}{dt} = m \int_L d\mathbf{v} \frac{d\mathbf{r}}{dt} = m \int_L d\mathbf{v} \cdot \mathbf{v} = m \int_L \mathbf{v} d\mathbf{v}$$

(outra vez: tudo pode tornar-se muito mais claro se pensarmos em  $d\mathbf{r}$ ,  $d\mathbf{v}$  e  $dt$  como se fossem  $\Delta\mathbf{r}$ ,  $\Delta\mathbf{v}$  e  $\Delta t$  – ver caixa).

Antes de prosseguirmos, calcularemos primeiro a derivada do produto escalar  $\mathbf{v}\mathbf{v}$  (já vamos ver para quê). De acordo com a regra de derivação do produto de duas funções temos

$$\frac{d}{dt}(\mathbf{v}\mathbf{v}) = \frac{d\mathbf{v}}{dt} \mathbf{v} + \mathbf{v} \frac{d\mathbf{v}}{dt} = 2\mathbf{v} \frac{d\mathbf{v}}{dt}.$$

Se  $\frac{d(\mathbf{v}\mathbf{v})}{dt} = 2\mathbf{v} \frac{d\mathbf{v}}{dt}$ , também é verdade que  $d(\mathbf{v}\mathbf{v}) = 2\mathbf{v}d\mathbf{v}$  (pensando em “ $d$ -s” como se fossem “ $\Delta$ -as”, podemos cortar  $\Delta t$ ). O produto escalar  $\mathbf{v}\mathbf{v}$  é igual a  $v^2$ . Então,  $d(v^2) = 2\mathbf{v}d\mathbf{v}$  ou  $\mathbf{v}d\mathbf{v} = \frac{1}{2}d(v^2)$ . Podemos agora voltar ao nosso integral e substituir

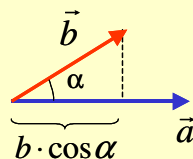
$\mathbf{v}d\mathbf{v}$  por  $\frac{1}{2}d(v^2)$ :

$$W_{AB}(L) = \frac{m}{2} \int_L d(v^2) = \frac{m}{2} v^2 \Big|_A^B = \frac{m}{2} (v_B^2 - v_A^2) = \frac{mv_B^2}{2} - \frac{mv_A^2}{2} = T_B - T_A$$

em que designamos  $\frac{mv^2}{2}$  por  $T$  que, para já, é apenas uma função - uma “estranha”



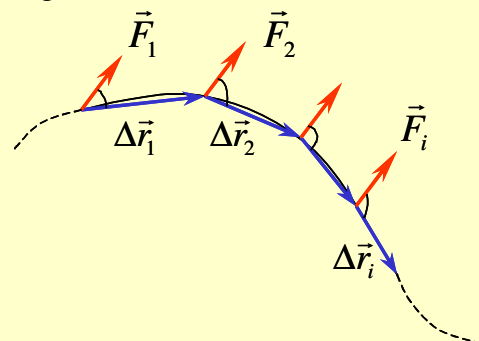
O **produto escalar** de dois vectores  $\mathbf{a} = \{a_x, a_y, a_z\}$  e  $\mathbf{b} = \{b_x, b_y, b_z\}$  é definido como  $\mathbf{a}\mathbf{b} = a_x b_x + a_y b_y + a_z b_z$ . Pode-se mostrar que  $\mathbf{a}\mathbf{b}$  é igual ao produto de módulos dos dois vectores vezes o co-seno do ângulo entre eles,  $\alpha$ :  $\mathbf{a}\mathbf{b} = ab \cdot \cos \alpha$  (para isso basta escolher os eixos de tal maneira que um dos dois vectores coincida com um eixo).



$$\begin{aligned} \alpha = 0 &\Rightarrow \mathbf{a}\mathbf{b} = ab \\ \alpha = \pi/2 &\Rightarrow \mathbf{a}\mathbf{b} = 0 \\ \alpha = \pi &\Rightarrow \mathbf{a}\mathbf{b} = -ab \end{aligned}$$



A interpretação do integral  $\int_L \mathbf{F} d\mathbf{r}$  torna-se mais clara se lembrarmos que este pode ser substituído pela soma  $\sum_i \mathbf{F}_i \cdot \Delta\mathbf{r}_i$  ao longo de toda a curva  $L$ :



combinação de variáveis dinâmicas.

Daqui se vê claramente qual o efeito das forças sobre um corpo em movimento: as forças, efectuando trabalho, obrigam o corpo a variar a sua velocidade. Se o trabalho for positivo,  $W_{AB}(L) > 0$ , a velocidade aumenta; se for negativo – diminui. Embora isto não deva surpreender-nos (já a segunda lei de Newton  $\mathbf{F} = m\mathbf{a}$  dizia que se existe uma força a actuar sobre um corpo, o corpo se move com aceleração, logo a velocidade varia) temos aqui uma nova relação entre a força e o seu efeito em que não aparece nenhuma variável temporal. A equação  $W_{AB}(L) = \frac{mv_B^2}{2} - \frac{mv_A^2}{2}$  não é “sensível” ao desenvolvimento do processo no tempo e inclui só dois valores da velocidade: no ponto inicial e no ponto final. Esta é uma grande vantagem porque não nos obriga a saber como variam as forças em função do tempo, como seria necessário se aplicássemos ao processo a segunda lei de Newton (ver Secção 2.3 deste Capítulo). Empregando o conceito de trabalho, só precisamos de saber como as forças dependem da posição, o que é normalmente, mais acessível. Este resultado confere uma importância muito especial à função  $T = \frac{mv^2}{2}$  que se costuma chamar *energia cinética* – uma grandeza já por nós conhecida desde os primeiros passos na ciência, mas que só aqui, i.e. no contexto do trabalho das forças, adquire o seu significado físico. E é precisamente aqui, da sua relação com o trabalho, que a energia cinética se materializa matematicamente como  $\frac{1}{2}$  do produto da massa do corpo e do quadrado da velocidade.

$$W_{AB} = T_B - T_A$$

A energia cinética é uma grandeza que caracteriza o movimento e, deste modo, a equação  $W_{AB} = T_B - T_A$  diz nos como uma força afecta



A “liberdade” com que manipulamos os diferenciais ( $dt$ ,  $d\mathbf{r}$ ,  $d\mathbf{v}$ ) na derivação da equação  $W_{AB} = T_B - T_A$  pode ser justificada se, mais uma vez, pensarmos neles como nas pequenas variações ( $\Delta t$ ,  $\Delta \mathbf{r}$ ,  $\Delta \mathbf{v}$ ) e lembrarmos a definição da derivada como sendo o limite da razão entre essas variações. Por exemplo, por definição  $\frac{d\mathbf{r}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{r}}{\Delta t}$ . Isso explica a passagem  $\frac{d\mathbf{v}}{dt} d\mathbf{r} = d\mathbf{v} \frac{d\mathbf{r}}{dt}$ : pensando em “ $\Delta$ -s” em vez dos “ $d$ -s” temos  $\frac{\Delta \mathbf{v}}{\Delta t} \Delta \mathbf{r} = \frac{\Delta \mathbf{v} \cdot \Delta \mathbf{r}}{\Delta t} = \Delta \mathbf{v} \frac{\Delta \mathbf{r}}{\Delta t}$ .

o movimento. Se uma força efectua um trabalho positivo, faz aumentar a energia cinética do corpo, ou seja, a velocidade deste. Um exemplo desta situação é um veículo movido a motor: o trabalho da força motriz do motor é positiva em qualquer parte da trajectória porque o vector força  $\mathbf{F}$  é paralelo ao vector deslocamento  $d\mathbf{r}$  (ou  $\Delta\mathbf{r}$ ) e ambos têm o mesmo sentido e, assim, deve resultar (de acordo com a equação) que a energia cinética do veículo aumenta. E é isto que acontece, como todos nós sabemos – o veículo acelera (desprezamos o atrito neste exemplo). Um exemplo oposto – empurramos um carrinho de compras e largamos de seguida. Sabemos muitíssimo bem que ao percorrer uma certa distância, o carrinho pára e também sabemos que a razão disso é a existência da força de atrito. Em termos do trabalho e da energia cinética este processo explica-se da seguinte maneira. A força de atrito efectua um trabalho negativo (porque  $\mathbf{F}_a$  tem sempre sentido oposto ao movimento, i.e.  $\mathbf{F}(\mathbf{r})d\mathbf{r} < 0$  em qualquer parte da trajectória) e isto resulta que a energia cinética do corpo diminui.

E se o vector força for perpendicular ao vector deslocamento  $\Delta\mathbf{r}$ ? A equação diz que neste caso o trabalho da força é igual a zero (porque o produto escalar  $\mathbf{F}(\mathbf{r})d\mathbf{r} = 0$ ) e, então, a energia cinética deve manter-se constante. É isto que acontece no caso de um movimento circular: a força centrípeta (tensão do fio, por exemplo) não efectua trabalho e o módulo da velocidade do corpo mantém-se constante. (O efeito da força centrípeta é a alteração da direcção do vector da velocidade, mas este efeito não está incluído na energia cinética – a energia cinética é uma grandeza escalar e é uma função do quadrado da velocidade  $v^2 = \mathbf{v} \cdot \mathbf{v}$ .)

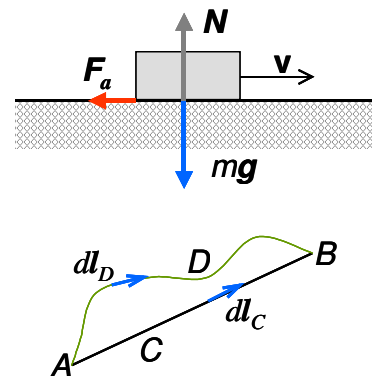
Apresentaremos em baixo os cálculos mais detalhados do trabalho para estes três exemplos.



### Exemplo 1 Trabalho da força de atrito.

Começaremos por calcular o trabalho que efectua a força de atrito. Um corpo de massa  $m$  desloca-se ao longo de uma superfície plana do ponto A até o ponto B com força de atrito dada pela equação  $\mathbf{F}_a = -kN\hat{\mathbf{e}}_v$

onde  $k$  é uma constante (coeficiente de atrito),  $N$  é o módulo da força da reacção normal da superfície,  $N = mg$ , e  $\hat{\mathbf{e}}_v$  representa o versor (vector unitário) do vector



velocidade  $\mathbf{v}$ ,  $\hat{\mathbf{e}}_v = \frac{\mathbf{v}}{|\mathbf{v}|}$ . O sinal *menos* na equação para a força de atrito significa que o vector da força tem sentido contrário ao do vector velocidade.

Vamos comparar o trabalho efectuado pela força de atrito em duas trajectórias diferentes,  $C$  e  $D$  (ver a figura). Para a trajectória  $C$  (recta) calculamos

$$W_{AB}(C) = \int_C \mathbf{F}_a d\mathbf{l}_C = \int_C (-kN\hat{\mathbf{e}}_v) \cdot d\mathbf{l}_C = -kN \int_C \hat{\mathbf{e}}_v \cdot d\mathbf{l}_C.$$

Como o sentido do vector deslocamento,  $d\mathbf{l}_C$  coincide com o sentido do vector velocidade em qualquer parte da trajectória,  $\hat{\mathbf{e}}_v \cdot d\mathbf{l}_C = \|\hat{\mathbf{e}}_v\| \cdot \|d\mathbf{l}_C\| \cdot \cos 0^\circ = dl_C$  (repare, que  $dl_C$  é um escalar e representa o comprimento do vector  $d\mathbf{l}_C$ ). Então,

$$W_{AB}(C) = -kN \int_C \hat{\mathbf{e}}_v \cdot d\mathbf{l}_C = -kN \int_C dl_C = -kN \cdot L_C$$

em que  $L_C$  corresponde ao comprimento da linha  $C$ .

Para a trajectória curvilínea  $D$ , escrevemos

$$W_{AB}(D) = \int_D \mathbf{F}_a d\mathbf{l}_D = \int_D (-kN\hat{\mathbf{e}}_v) \cdot d\mathbf{l}_D = -kN \int_D \hat{\mathbf{e}}_v \cdot d\mathbf{l}_D = -kN \int_D dl_D = -kN \cdot L_D$$

(para o produto  $\hat{\mathbf{e}}_v \cdot d\mathbf{l}_D$  são válidas as mesmas considerações, que para o caso da trajectória  $C$ , porque  $\hat{\mathbf{e}}_v \parallel d\mathbf{l}_D$  em qualquer parte da trajectória e, portanto, o  $\hat{\mathbf{e}}_v \cdot d\mathbf{l}_D = \|\hat{\mathbf{e}}_v\| \cdot \|d\mathbf{l}_D\| \cdot \cos 0^\circ = dl_D$ ).

Como já era de esperar, o trabalho efectuado pela força de atrito é negativo qualquer que seja a trajectória do corpo. De facto, isto é válido para qualquer força que resista ao movimento, i.e. que actue no sentido contrário ao movimento.

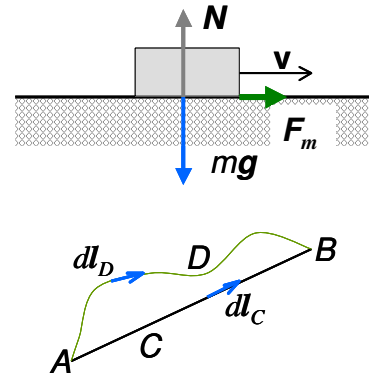
Outra observação importante é que o trabalho da força de atrito depende da forma concreta da trajectória que liga os dois pontos. Como  $L_C < L_D$ , obviamente

$$|W_{AB}(C)| < |W_{AB}(D)|.$$

### **Exemplo 2 Trabalho da força motriz.**

Consideremos um carro de massa  $m$  a deslocar-se ao longo de uma superfície plana e horizontal do ponto  $A$  até o ponto  $B$ . Suponhamos que não existe atrito no sistema e a intensidade da força motriz é constante durante todo o processo. O diagrama das forças a actuar no sistema está representado na figura. A força da reacção normal da superfície devido ao peso do corpo anula-se com a força gravítica,

de modo que a força motriz se torna a única força relevante para a descrição do processo. A direcção e o sentido desta força coincidem com o vector velocidade em qualquer parte da trajectória e, portanto, podemos escrever para ela  $\mathbf{F}_m = F_m \cdot \hat{\mathbf{e}}_v$  onde  $F_m = \|\mathbf{F}_m\|$  e  $\hat{\mathbf{e}}_v$  representa, como antes, o versor do vector  $\mathbf{v}$ ,  $\hat{\mathbf{e}}_v = \frac{\mathbf{v}}{|\mathbf{v}|}$ .



Como já fizemos para a força de atrito, vamos comparar o trabalho efectuado pela força motriz em duas trajectórias diferentes, C e D (ver a figura). Para a trajectória C (recta) temos

$$W_{AB}(C) = \int_C \mathbf{F}_m d\mathbf{l}_C = \int_C F_m \hat{\mathbf{e}}_v d\mathbf{l}_C = F_m \int_C \hat{\mathbf{e}}_v d\mathbf{l}_C.$$

Como o sentido do vector deslocamento,  $d\mathbf{l}_C$  coincide com o sentido do vector velocidade em qualquer parte da trajectória,  $\hat{\mathbf{e}}_v d\mathbf{l}_C = \|\hat{\mathbf{e}}_v\| \cdot \|d\mathbf{l}_C\| \cdot \cos 0^\circ = d\mathbf{l}_C$  (repare-se, que  $d\mathbf{l}_C$  é um escalar). Então,

$$W_{AB}(C) = F_m \int_C \hat{\mathbf{e}}_v d\mathbf{l}_C = F_m \int_C d\mathbf{l}_C = F_m \cdot L_C$$

em que  $L_C$  corresponde ao comprimento da linha C.

Para a trajectória curvilínea D, escrevemos

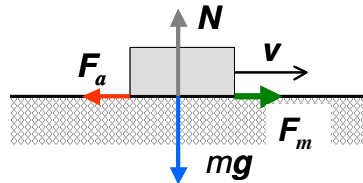
$$W_{AB}(D) = \int_D \mathbf{F}_m d\mathbf{l}_D = \int_D F_m \hat{\mathbf{e}}_v d\mathbf{l}_D = F_m \int_D \hat{\mathbf{e}}_v d\mathbf{l}_D = F_m \int_D d\mathbf{l}_D = F_m \cdot L_D$$

(para o produto  $\hat{\mathbf{e}}_v d\mathbf{l}_D$  são válidas as mesmas considerações, que para o caso da trajectória C, porque  $\hat{\mathbf{e}}_v \parallel d\mathbf{l}_D$  em qualquer parte da trajectória e, portanto, o  $\hat{\mathbf{e}}_v d\mathbf{l}_D = \|\hat{\mathbf{e}}_v\| \cdot \|d\mathbf{l}_D\| \cdot \cos 0^\circ = d\mathbf{l}_D$ ).

Vê-se claramente que o trabalho da força motriz é positivo qualquer que seja a trajectória e, deste modo, contribui para aumento da energia cinética do corpo.

Como  $L_C < L_D$ ,  $W_{AB}(C) < W_{AB}(D)$ , daqui concluímos que o trabalho da força motriz depende da forma concreta da trajectória que liga o ponto de partida ao ponto de chegada e, mesmo se no fim da viagem o carro voltar ao mesmo ponto, a força vai efectuar um trabalho não nulo.

A nossa experiência do dia a dia diz-nos que nas situações reais gastamos gasolina mesmo que o carro ande com uma velocidade constante. Em termos físicos isto significa que apesar de a força motriz efectuar trabalho positivo, a energia cinética (e a velocidade) do veículo não aumenta. Isto está em aparente contradição com as equações que escrevemos em cima. A solução desta contradição é simples e óbvia: numa situação real existe resistência ao movimento. Se adicionarmos uma força de atrito (ver a figura), torna-se claro que a razão para o não aumento da energia cinética é que a força de atrito anula a força motriz e a força resultante, então, será nula. Em termos de trabalho, podemos dizer que o trabalho da força motriz (positivo) anula-se pelo trabalho da força de atrito (negativo) de modo que o resultado é igual a zero. Matematicamente, isto é possível porque trabalho é uma função linear em relação à força e, então, se  $\mathbf{F}_{resul\ tan\ te} = \mathbf{F}_m + \mathbf{F}_a$ , é verdade que



$$W_{AB}(L) = \int_L \mathbf{F}_{resul\ tan\ te} d\mathbf{r} = \int_L (\mathbf{F}_m + \mathbf{F}_a) d\mathbf{r} = \int_L \mathbf{F}_m d\mathbf{r} + \int_L \mathbf{F}_a d\mathbf{r} = W_{AB}^{(motriz)}(L) + W_{AB}^{(atrito)}(L).$$

### **Exemplo 3 Trabalho da força de reacção normal.**

Um corpo move-se numa superfície plana (ver a figura do exemplo anterior). Determinemos o trabalho efectuado pela força da reacção normal.

A força de reacção normal é sempre perpendicular ao vector velocidade. Portanto

$$W_{AB}(C) = \int_C \mathbf{N} d\mathbf{r} = \int_C \mathbf{N} \cdot \mathbf{v} dt = 0$$

Utilizamos aqui a definição da velocidade  $\mathbf{v} = \frac{d\mathbf{r}}{dt} \Rightarrow d\mathbf{r} = \mathbf{v} dt$  e, também, o facto de

$$\mathbf{N} \cdot \mathbf{v} = \|\mathbf{N}\| \cdot \|\mathbf{v}\| \cdot \cos \frac{\pi}{2} = 0.$$

### **Exemplo 4 Trabalho da força de tensão do fio.**

Um corpo fixo na extremidade de um fio faz movimento circular. A tensão do fio é a força que obriga o corpo a seguir uma trajectória circular, i.e. alterar a direcção do vector velocidade e, portanto, mover-se com uma aceleração. Mas como esta força é

perpendicular ao vector velocidade,  $\mathbf{F}_T d\mathbf{r} = \mathbf{F}_T \mathbf{v} dt = \|\mathbf{F}_T\| \cdot \|\mathbf{v}\| \cdot \cos \frac{\pi}{2} \cdot dt = 0$ , o trabalho da força é igual a 0.

É importante sublinhar que na derivação da equação  $W_{AB} = T_B - T_A$  não utilizamos nenhuma restrição acerca das propriedades da força. Portanto, a equação é válida para qualquer tipo de força.

$$W_{AB} = T_B - T_A$$

para qualquer força

### 3.2 Forças conservativas. Energia potencial.

Se se conseguir calcular o trabalho efectuado por uma dada força numa dada trajectória  $L$ , que liga o ponto de partida  $A$  ao ponto de chegada  $B$ , podemos saber como varia a energia cinética do corpo. O cálculo do trabalho nem sempre é uma tarefa fácil. Em geral, o trabalho depende da forma concreta da trajectória  $L$ , como já tivemos possibilidade de notar quando o calculámos para as forças de motriz e de atrito. Felizmente que nem todas as forças são assim. Existe um grupo de forças cujo trabalho não depende da forma concreta do caminho percorrido mas é uma função das coordenadas de apenas dois pontos – ponto inicial e ponto final:

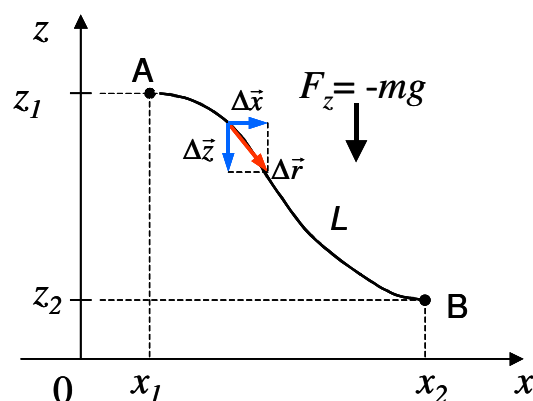
$$W_{AB} = \int_L \mathbf{F} d\mathbf{r} = u(\mathbf{r}_B) - u(\mathbf{r}_A),$$

onde  $u(\mathbf{r})$  é usado para designar uma função (genérica) de coordenadas espaciais.

Um exemplo da força deste tipo é a força gravítica:

#### Exemplo 5. Trabalho da força gravítica.

Um corpo de massa  $m$  desloca-se no campo gravítico do ponto  $A$  com coordenadas  $(x_1, z_1)$  ao ponto  $B$  com coordenadas  $(x_2, z_2)$  ao longo de uma trajectória curvilínea  $L$  como se mostra na figura (suponhamos que o movimento ocorre no plano  $y = 0$ ).





De acordo com a definição, o trabalho é igual a  $W_{AB}(L) = \int_L \mathbf{F} d\mathbf{r}$  em que a força é a força gravítica, neste caso,  $d\mathbf{r}$  é um elemento (infinitamente pequeno) do vector de deslocamento e a integração tem de ser feita ao longo da curva  $L$ , cuja forma é indicada na figura. O vector da força gravítica só tem um componente - o componente  $z$ , i.e.  $\mathbf{F} = \{0, 0, -mg\}$ . Utilizando a definição do produto escalar e representando o vector  $d\mathbf{r}$  em componentes também  $d\mathbf{r} = \{dx, dy, dz\}$ , temos

$$\mathbf{F} d\mathbf{r} = 0 \cdot dx + 0 \cdot dy - mg \cdot dz = -mg \cdot dz$$

e, para o trabalho

$$W_{AB}(L) = \int_{z_1}^{z_2} (-mg) dz = -mg \int_{z_1}^{z_2} dz = -mg(z_2 - z_1) = -(mgz_2 - mgz_1).$$

Usando a designação introduzida no início desta secção, vê-se claramente que no caso da força gravítica a função  $u(\mathbf{r}) = -mgz$ .

Este resultado pode ser compreendido se lembrarmos o significado do integral  $\int_L \mathbf{F} d\mathbf{r}$  como soma  $\sum_i \mathbf{F}_i \cdot \Delta\mathbf{r}_i$  calculada no limite  $\Delta\mathbf{r}_i \rightarrow 0$ . Os elementos  $\Delta\mathbf{r}_i$  são pequenas deslocações, tão pequenas que a trajectória curvilínea pode ser aproximada por uma soma de traços rectilíneos ( $d\mathbf{r}$  é um elemento infinitamente pequeno; os  $\Delta\mathbf{r}_i$  são pequenas mas finitas). O produto  $\mathbf{F}_i \cdot \Delta\mathbf{r}_i$  calcula-se para cada elemento rectilíneo,  $\Delta\mathbf{r}_i$ , e os resultados somam-se. Seguindo esta lógica e representando agora o deslocamento  $\Delta\mathbf{r}_i$  como soma de um deslocamento na horizontal e de um deslocamento na vertical,  $\Delta\mathbf{r}_i = \Delta\mathbf{x}_i + \Delta\mathbf{z}_i$  (ver a figura), vê-se logo que a força efectua o trabalho só na parte do deslocamento vertical e este trabalho é positivo ( $\mathbf{F} \parallel \Delta\mathbf{z} \Rightarrow \mathbf{F} \Delta\mathbf{z} = \|\mathbf{F}\| \cdot \|\Delta\mathbf{z}\|$ ). Na parte do deslocamento horizontal, a força é perpendicular ao deslocamento e, portanto, o trabalho é igual a zero.

---

Agora, lembremo-nos que uma força (qualquer), ao efectuar trabalho sobre um corpo, faz com que a energia cinética do corpo se altere:  $W_{AB} = T_B - T_A$ . Combinando esta equação com a equação  $W_{AB} = u(\mathbf{r}_B) - u(\mathbf{r}_A)$ , temos

$$T_B - T_A = u(\mathbf{r}_B) - u(\mathbf{r}_A)$$

ou, ainda,

$$T_B - u(\mathbf{r}_B) = T_A - u(\mathbf{r}_A).$$

Estamos apenas a um passo de uma grande descoberta. Vamos substituir a função  $u(\mathbf{r})$  recentemente introduzida por uma outra função  $u(\mathbf{r}) = -U(\mathbf{r})$  ou, simplesmente,  $U$ . Temos, então

$$T_B + U_B = T_A + U_A$$

ou

$$T + U = \text{const},$$

i.e., qualquer que seja o caminho percorrido entre os pontos  $A$  e  $B$ , a soma de duas grandezas  $T + U$  conserva-se. Como  $T = \frac{mv^2}{2}$  é a energia cinética, é óbvio, por razões dimensionais, que  $U$  também tem significado de energia. E aqui está a lei de conservação da energia mecânica. Só nos resta chamar todas as grandezas envolvidas com os nomes a que nos temos habituado:

$T$  – energia cinética,

$U$  – energia potencial,

$E = T + U$  – energia mecânica.

Os sistemas em que a energia mecânica se conserva chamam-se **sistemas conservativos** e as forças, cuja acção não altera a energia mecânica do sistema, chamam-se **forças conservativas**.

É importante lembrarmos que para se chegar à conclusão sobre a conservação da energia mecânica num sistema, assumimos inicialmente que o trabalho efectuado pela força a actuar no sistema não depende da forma da trajectória mas é uma função apenas das coordenadas dos pontos inicial e final:

$$W_{AB}^{(cons.)} = -(U_B - U_A).$$

De facto, podemos usar esta propriedade para definir o que é uma força conservativa. Mas o mais importante é que isto nos dá a compreensão de o que é, realmente, a energia potencial: a energia potencial (mais precisamente, o simétrico da variação desta) é nada mais nada menos que o trabalho efectuado pela força conservativa. Daqui se seguem uma série de conclusões importantes relacionadas com o conceito da energia potencial:

1. A energia potencial é definida exclusivamente para uma força conservativa.
2. A energia potencial é definida a menos de uma constante aditiva (só sabemos definir  $\Delta U$  e não  $U_A$  e  $U_B$  em separado).

3. O conceito da energia potencial está fortemente ligado ao conceito da força. Lembrando a definição da energia potencial como o trabalho efectuado pela força conservativa  $\int_L \mathbf{F} d\mathbf{r} = -U + \text{const}$  e, também, a definição da primitiva

de uma função, podemos deduzir que  $\mathbf{F}(\mathbf{r}) = -\frac{dU(\mathbf{r})}{d\mathbf{r}}$  (ver a caixa para

significado da escrita  $\frac{dU(\mathbf{r})}{d\mathbf{r}}$ ; no caso de movimento numa só dimensão

temos  $F(x) = -\frac{dU(x)}{dx}$ ).

A importância desta última equação consiste em que ela mostra que, de ponto de vista físico, a força e a energia potencial são ambas características do mesmo fenómeno. Este

$$\mathbf{F}(\mathbf{r}) = -\frac{dU(\mathbf{r})}{d\mathbf{r}}$$

para uma força conservativa

fenómeno é a *interacção*. A interacção entre os objectos físicos (corpos, partículas, átomos etc.) é que dá origem às forças. A energia potencial, por sua vez, também só faz sentido quando houver uma interacção. Por exemplo, no caso da bem conhecida força gravítica, a *interacção gravítica* entre duas massas está na origem dessa força. Sabemos que à superfície da Terra a energia potencial gravítica é dada por  $U(z) = mgz$ , se escolhermos o eixo  $z$  a apontar para cima. A derivada da  $U(z)$  em ordem do  $z$  é igual a  $mg$ . A força é o simétrico disso,  $F(z) = -\frac{dU(z)}{dz} = -mg$ , como era de esperar. O sinal menos significa que o vector da força aponta para baixo.

Na física clássica, ambos os conceitos (tanto o da força como da energia potencial) são igualmente usadas para descrever a interacção no sistema: sabendo a



No caso de duas ou três dimensões a expressão  $\frac{dU(\vec{r})}{d\vec{r}}$  tem o significado de gradiente,  $\text{grad}(U)$ , ou  $\vec{\nabla}U$ . Repare-se, que o argumento desta função é uma grandeza escalar mas o resultado é um vector. Por definição, o gradiente é um vector com os seguintes componentes  $\text{grad}(U) = \left\{ \frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z} \right\}$ . O vector  $\text{grad}(U)$  indica o sentido em que a função  $U(\vec{r})$  cresce mais rapidamente. Para uma função unidimensional,  $U(x)$ , gradiente positivo diz nos que a função cresce no sentido  $x$  positivo enquanto  $\text{grad}(U(x)) < 0$  significa que a função  $U(x)$  diminui quando  $x$  aumenta.

força em função das coordenadas espaciais, podemos sempre determinar a função de energia potencial e vice-versa. Já na física quântica, da qual vamos falar no Capítulo 6, não é bem assim. As interações num sistema quântico caracterizam-se exclusivamente pela energia potencial e o conceito da força não tem lugar no formalismo quântico. É a função de energia potencial que entra na equação de Schrödinger na parte correspondente à interação - equação que tem um papel semelhante à 2ª lei de Newton na mecânica clássica.

---

**Exemplo 6. Trabalho da força elástica.**

Consideremos um sistema em que a única força relevante é a força elástica, por exemplo, uma mola de constante elástica  $k$ . A força elástica, como é sabido, é proporcional ao alongamento (ou à compressão) e actua no sentido contrário a este:  $F(x) = -kx$  (aqui supõe-se que  $x = 0$  corresponde à posição da extremidade livre da mola enquanto não estendida). O trabalho que a força elástica efectua quando estendemos a mola de  $x_1$  a  $x_2$  é calculado como

$$W = \int_{x_1}^{x_2} (-kx) dx = -k \int_{x_1}^{x_2} x dx = -k \left( \frac{x^2}{2} \right)_{x_1}^{x_2} = - \left( \frac{kx_2^2}{2} - \frac{kx_1^2}{2} \right).$$

Comparando este resultado com a definição para a energia potencial, i.e.

$W_{AB}^{(cons.)} = -(U_B - U_A)$ , vê-se claramente que a energia potencial relacionada com a

força elástica é igual a  $U(x) = \frac{kx^2}{2}$ .

---

### 3.3 Sistemas com forças conservativas e não conservativas.

As forças não conservativas são aquelas que não conservam a energia mecânica (um exemplo clássico deste tipo de força é a força de atrito). Num sistema físico com as forças não conservativas já não podemos utilizar a lei de conservação da energia mecânica, tão útil na resolução de muitos problemas. Mas será que existe uma maneira de incluir o efeito dessas forças na equação de conservação da energia para que seja possível continuar a usá-lo?

Vamos representar a força resultante num sistema físico como a soma de duas componentes, uma sendo a soma de todas as forças conservativas,  $\mathbf{F}^{(cons.)}$ , e a outra a soma de todas as forças não conservativas,  $\mathbf{F}^{(n.c.)}$ :

$$\mathbf{F} = \mathbf{F}^{(cons.)} + \mathbf{F}^{(n.c.)}.$$

O trabalho dessas forças todas é

$$\begin{aligned} W_{AB}(L) &= \int_L \mathbf{F} d\mathbf{r} = \int_L \left( \mathbf{F}^{(cons.)} + \mathbf{F}^{(n.c.)} \right) d\mathbf{r} = \int_L \mathbf{F}^{(cons.)} d\mathbf{r} + \int_L \mathbf{F}^{(n.c.)} d\mathbf{r} = \\ &= W_{AB}^{(cons.)} + W_{AB}^{(n.c.)}(L) \\ \Rightarrow W_{AB}(L) &= W_{AB}^{(cons.)} + W_{AB}^{(n.c.)}(L). \end{aligned}$$

O trabalho das forças conservativas, já sabemos, não depende da trajectória concreta seguida pelo corpo e é igual ao simétrico da variação da energia potencial  $W_{AB}^{(cons.)} = -(U_B - U_A)$ . O trabalho das forças não conservativas depende, em geral, da trajectória,  $L$ , e não existe uma maneira universal de o determinar a não ser através da integração directa  $\int_L \mathbf{F}^{(n.c.)} d\mathbf{r}$  tomando em conta todos os pormenores do percurso e orientação mútua dos vectores da força e de deslocamento em cada ponto deste.

Por outro lado, também sabemos que o trabalho de todas as forças, independentemente de serem conservativas ou não conservativas, resulta na variação da energia cinética:  $W_{AB}(L) = T_B - T_A$ . Substituindo  $W_{AB}(L)$  e  $W_{AB}^{(cons.)}$  pelas respectivas expressões, temos

$$T_B - T_A = -(U_B - U_A) + W_{AB}^{(n.c.)}$$

$$\text{ou } T_B + U_B = T_A + U_A + W_{AB}^{(n.c.)}$$

$$\text{e, finalmente, } E_B = E_A + W_{AB}^{(n.c.)}.$$

Esta última equação oferece-nos a saída. Sim, podemos continuar a utilizar a ideia de conservação da energia mecânica mas, no caso de um sistema com forças não conservativas, temos que incluir no balanço energético o trabalho efectuado por estas.

$$E_B = E_A + W_{AB}^{(n.c.)}$$



## Capítulo 3

# Interacções

Neste capítulo vamos falar sobre a força como manifestação de uma interacção entre os corpos. Falaremos sobre as quatro interacções fundamentais da Natureza conhecidas neste momento. Vamos apoiar-nos fortemente na relação entre a força e a energia potencial, a que chegámos no capítulo anterior, e vamos ver como se consegue descrever o comportamento de um sistema através da análise da variação espacial da energia potencial.

### 1. As forças fundamentais da Natureza

A mecânica newtoniana baseia-se fortemente no conceito de força. Basta conhecer as forças a actuar no sistema e, com a segunda lei de Newton à mão, conseguimos (melhor ou pior) descrever o movimento dos corpos e comportamento dos sistemas mecânicos. Claro, que existem, por vezes, problemas técnicos que dificultam a vida aos físicos. Não estamos a falar dessas dificuldades mas sim sobre o princípio. O conhecimento da força e da sua dependência espacio-temporal fornece toda a informação necessária para resolver um problema mecânico.

Mas de onde provêm as forças ? O que as faz aparecer ? Eis a questão.

Conhecemos uma vasta variedade de forças: força gravítica, força de atrito, força elástica, falamos da força de tensão do fio, da reacção normal da superfície sob peso de um corpo, etc. Será que todas essas forças são independentes ou algumas delas podem ser explicadas no contexto de outras, digamos, básicas ? A resposta a esta pergunta é, de facto, positiva. A maior parte das forças, das quais costumamos falar na física, é uma manifestação particular das poucas forças fundamentais da Natureza. Foram introduzidas no decurso da física, às vezes por falta da melhor explicação na altura, outras vezes por razões de conveniência. Das cinco forças mencionadas acima apenas a força gravítica é uma força fundamental; as outras quatro são, no fundo, devidas à força electromagnética (outra força fundamental). A força de atrito, por exemplo, é o resultado da interacção entre os átomos nas

superfícies dos dois corpos em contacto. Em termos práticos, no entanto, isso não é assim tão importante. Para resolver problemas da mecânica, basta-nos saber a força resultante e as suas propriedades.

No que toca à origem de uma ou de outra força fundamental, podemos avançar um pouco mais dizendo que as forças diferentes são o resultado das diferentes *interacções* entre os corpos. Neste sentido, a força é uma manifestação observável de um dado tipo de interacção.

É costume identificar-se quatro interacções fundamentais da Natureza: a interacção gravítica, a interacção electromagnética, a interacção nuclear (também chamada interacção forte) e a interacção fraca.

A *interacção gravítica* é bem conhecida: é assim que interagem dois objectos com massas não nulas. As massas entram nas equações da força e da energia potencial associada a este tipo de interacção. A força gravítica é sempre atractiva e actua a qualquer distância entre os corpos, diminuindo na proporção inversa do quadrado da distância. As forças deste tipo chamam-se forças de *longo alcance* (infinito).

Para que dois objectos interajam através da *força electromagnética*, têm que possuir carga eléctrica. Se os objectos estão em repouso, a interacção entre eles descreve-se com a lei de Coulomb. Uma vez que existem dois tipos de carga eléctrica, as cargas positivas e as cargas negativas, a força entre os corpos carregados pode ser de atracção ou de repulsão, em contraste com a força gravítica. A força colombiana é também uma força de longo alcance com o módulo a diminuir na proporção inversa do quadrado da distância entre as cargas.

A *interacção nuclear* é responsável pela estabilidade do núcleo atómico que, como se sabe consiste em dois tipos de partículas, protões e neutrões. Os protões têm carga eléctrica positiva (e por isso, interagem também pela força electromagnética) enquanto os neutrões são electricamente neutros (i.e. têm carga eléctrica igual a zero). A força nuclear, com que interagem entre si tanto os protões como os neutrões, é uma força atractiva que compensa a repulsão electrostática entre os protões e assim faz com que os núcleos atómicos fiquem estáveis. A força nuclear é uma força de *curto alcance*, actua nas distâncias da ordem de  $10^{-15}$  m (que é comparável ao tamanho de um núcleo). A essa distância, a força nuclear é muito maior que a força electrostática, mas nas distâncias superiores isso é desprezável comparativamente com a última. Para as distâncias de ordem de  $10^{-16}$  m, a atracção devida à força nuclear transforma-se em



repulsão muito forte. Isso explica por que é que os núcleos atômicos têm dimensões finitas e não colapsam.

Finalmente, a *interacção fraca* é aquela que é responsável pelo decaimento das partículas. Actua nas distâncias ainda mais curtas, da ordem de  $10^{-18}$  m. Um exemplo do processo que ocorre devido a este tipo de interacção é o decaimento de um neutrão. O neutrão livre não é estável e decai para o protão, electrão e uma partícula neutra (electricamente) chamada anti-neutrino (uma partícula de *anti-matéria*),  $n \rightarrow p + e^- + \bar{\nu}$ , com um tempo característico (período de semi-transformação) de cerca de 10 minutos.

Estes quatro tipos de forças são independentes. Não se pode descrever nenhuma delas como se fosse manifestação de uma interacção de outro tipo ou uma combinação destas. Por isso, todas elas se consideram presentemente como as forças fundamentais da Natureza. Não se pode dizer qual é a razão de existência dessas forças, mas há provas experimentais suficientes para afirmar que elas existem e são diferentes. A pergunta – *por que é que são quatro?* – também não há resposta, mas a ideia de unir todas as forças sob um mesmo telhado é intensamente explorada pelos físicos desde há muitos anos. O próprio Einstein dedicou uma grande parte da sua vida a tentar fundir a força gravítica com a electromagnética, mas sem sucesso. O progresso conseguiu-se noutro lado: na unificação das interacções electromagnética e fraca e desenvolvimento da teoria de *interacção electro-fraca*, que descreve as duas interacções com um único formalismo (prémio Nobel de 1979 para S. Glashow, A. Salam e S. Weinberg). Já a extensão desta unificação para a força nuclear é uma tarefa mais difícil, mas existe um modelo, ainda não concluído, que trata as três interacções (electromagnética, fraca e nuclear) numa base comum. Este modelo chama-se modelo de *Grande Unificação*. A força gravítica, a mais “antiga” de todas, é que continua ainda a ficar de fora ...

## 2. Campos

Uma interacção pressupõe que existem, pelo menos, dois objectos (daqui *interacção*) a exercer a influência um sobre o outro. A interacção gravítica entre o Sol e a Terra é um bom exemplo. Sabemos que o efeito do Sol sobre o movimento da Terra é dominante, mas sabemos igualmente que também existe uma interacção da Terra com

outros corpos celestes. A descrição mais completa do movimento da Terra no espaço, portanto, tem que ter em conta não só a atracção do Sol, mas também a atracção da Lua, de Vénus, de Júpiter e dos seus anéis, etc. Por outras palavras, é a força resultante de todas essas interacções que tem que entrar na equação de movimento. Nesta perspectiva, seria conveniente considerar a Terra como se esta se movesse num *campo gravítico* exterior - num campo que engloba os efeitos de todos os corpos sobre a Terra. A força exercida por este campo sobre a Terra seria, então, a soma vectorial de todas as forças gravíticas e a energia potencial seria a soma das energias potenciais associadas à atracção de cada corpo do sistema. Uma vez descrito este campo (quer dizer que se sabe como calcular a força e/ou energia potencial em qualquer ponto do espaço e em qualquer momento), podemos pôr de parte a sua origem e “esquecer” quais são realmente as fontes das forças (diz-se, também, *fontes do campo*), porque para efeitos práticos só o efeito combinado é que interessa.

Esta abordagem tem um defeito óbvio - parece que se perde o efeito que a Terra exerce sobre os outros planetas e sobre o próprio Sol (a atracção é recíproca, não esqueçamos isso). Mas este efeito pode ser desprezado em muitos casos, beneficiando deste modo da simplificação do problema.

Continuando esta lógica, podemos aplicar o conceito do campo até para o caso de apenas dois corpos. É costume falar-se do campo gravítico da Terra em que todos nós nos movemos. É habitual dizer-se que uma maçã cai sob a influência deste campo. O que se esquece, normalmente, nesta interpretação é que a maçã também atrai a Terra e que, o que nos parece apenas a queda de uma maçã, é, de facto, os dois corpos (a maçã e a Terra) a moverem-se um ao encontro do outro; isto porque a atracção é mútua e porque a força, com que a maçã atrai a Terra, é igual em módulo à força com que a Terra atrai a maçã. O efeito observável, no entanto, é muito diferente para os dois corpos por causa da diferença enorme entre as massas. O módulo da força gravítica é igual, como sabemos, a  $mg$  (onde  $m$  é a massa da maçã e  $g$  é a aceleração da gravidade) e a aceleração é igual a  $g$ . A Terra atrai a maçã com a mesma força, mas a sua aceleração, segundo Newton, é  $a_T = \frac{F}{M_T} = \frac{mg}{M_T}$  (aqui  $M_T$  é a massa da Terra).

É fácil calcular essa aceleração: por exemplo, para  $m = 0,1 \text{ kg}$ ,  $a_T \approx 1,7 \cdot 10^{-26} \cdot g$ . Isto resulta no deslocamento da Terra no sentido da maçã apenas de cerca de  $0,8 \cdot 10^{-25} \text{ m}$ ,

se o tempo de queda for 1 segundo (recorde-se,  $\Delta x = \frac{at^2}{2}$ ?). Este deslocamento é dez ordens de grandeza menor que o diâmetro de um núcleo atómico e, por isso, é perfeitamente desprezável (nem sequer tem algum sentido físico: não há nenhum objecto conhecido na Natureza com as dimensões comparáveis a este valor).

No caso do Sol e da Terra, a razão entre as massas não é tão pequena (é de ordem de  $\sim 10^6$ ), mas mesmo assim o efeito da Terra sobre o Sol é bastante pequeno comparativamente com o efeito do Sol sobre a Terra. Por isso, podemos falar da Terra a mover-se no campo gravítico do Sol.

Este modo de ver o movimento de um corpo como movimento num campo exterior conduz-nos logicamente a uma dedução estranha: o Sol cria (“emite”?) à sua volta um campo gravítico que parece de não depender da existência da Terra ou de qualquer outro objecto nas vizinhanças. Será que se tirássemos todas as estrelas e planetas do espaço, excepto o Sol, o campo gravítico ainda existiria? Se respondermos *Sim*, estaremos a contradizer à ideia de reciprocidade da *interacção* gravítica. Na verdade, pensar assim (que o campo se mantém) ou pensar o contrário, não traz nenhuma consequência práticas, ou seja, efeitos observáveis. A única maneira de verificar a existência (ou não) do campo gravítico é colocar uma massa de teste (uma sonda) algures no espaço e medir a força que esta vai (ou não vai) sofrer. Mas, assim, já temos um par de corpos...

O conceito do campo tem sido muito produtivo em física. As mesmas considerações que acabamos de aplicar à interacção gravítica, são aplicáveis a qualquer uma das outras interacções fundamentais. Assim, fala-se do campo electromagnético, campo da força nuclear, campo da força fraca. Em termos formais, apesar de ser chamado “campo da força”, o parâmetro físico que se ajusta melhor a este conceito não é a força, mas sim a energia potencial. Porém, as duas grandezas estão interligadas pela relação  $F(x) = -\frac{dU(x)}{dx}$ , que já conhecemos.

Na secção seguinte vamos ver como podemos utilizar a energia potencial e a sua dependência das coordenadas espaciais para a caracterização do movimento de um corpo. Vamos ver isso num exemplo do campo gravítico, mas lembrando que a mesma lógica também é aplicável aos campos de outra natureza.

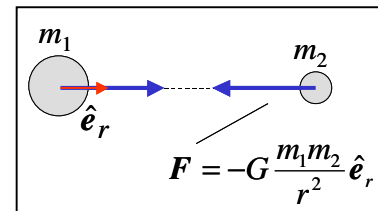
### 3. O campo gravítico e o campo electrostático

#### 3.1 O campo gravítico

A força gravítica entre dois corpos com as massas  $m_1$  e  $m_2$  é dada pela lei de gravitação universal de Newton

$$\mathbf{F} = -G \frac{m_1 m_2}{r^2} \hat{\mathbf{e}}_r,$$

em que  $G$  é a constante universal gravítica ( $G=6.67 \cdot 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2$ ),  $r$  é a distância entre os dois corpos (entre os centros de massa destes, mais precisamente) e  $\hat{\mathbf{e}}_r$  é o versor (vector unitário) do  $\mathbf{r}$ . Como a atracção é mútua, existe sempre um par de forças, uma aplicada ao corpo com massa  $m_1$  e outra aplicada ao corpo com massa  $m_2$ . Se o versor  $\hat{\mathbf{e}}_r$  na equação acima se refere ao vector  $\mathbf{r}$  com a origem no corpo 1, o  $\mathbf{F}$  é, então, a força aplicada ao corpo 2 (ver a figura).



Dada a simetria esférica neste problema (a força só depende do  $r$ ), podemos reduzi-la a uma dimensão e utilizar a equação unidimensional  $F(r) = -\frac{dU(r)}{dr}$  para a relação entre a força e a energia potencial em vez da equação geral  $\mathbf{F} = -\text{grad}(U)$ . Assim, temos  $dU = -F(r)dr$  e, explicitando a força e integrando as duas partes da equação, temos

$$\int dU = \int G \frac{m_1 m_2}{r^2} dr$$

$$U(r) = G m_1 m_2 \int \frac{dr}{r^2}$$

$$U(r) = -G \frac{m_1 m_2}{r} + \text{const}$$

A  $\text{const}$  é uma constante de integração que pode ser escolhida voluntariamente consoante as condições do problema concreto. O valor da constante não afecta o resultado prático porque em todos os problemas interessam-nos apenas duas coisas: a força, que é derivada da energia potencial (a derivada de uma constante é zero), ou a diferença  $U_1 - U_2$  entre dois pontos. É conveniente em muitos problemas considerar

que  $U \rightarrow 0$  quando  $r \rightarrow \infty$ . A constante de integração torna-se, neste caso, igual a zero.

Para usar a ideia do campo associado a um dos dois corpos, por exemplo ao corpo com a massa  $m_1$ , precisamos de um parâmetro que não dependa da massa do outro corpo. Este parâmetro chama-se *potencial*,  $\phi(r)$ , do campo gravítico e obtém-se por divisão da energia potencial  $U(r)$  por  $m_2$ :

$$\phi_1(r) = -G \frac{m_1}{r}.$$

O significado do potencial é muito simples: é a energia potencial associada a um corpo de massa de 1 kg colocado à distância  $r$  do corpo 1 (com a massa  $m_1$ ). As unidades do potencial são J/kg. Da mesma maneira, obviamente, podemos definir o potencial do campo gravítico do corpo 2:  $\phi_2(r) = -G \frac{m_2}{r}$ . O

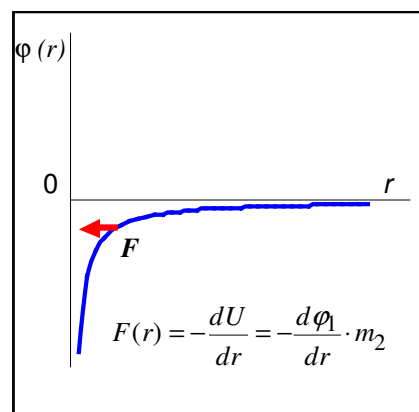


gráfico da função  $\phi(r)$  é semelhante ao da energia potencial,  $U(r)$ , uma vez que estes são directamente proporcionais:  $U(r) = \phi_1(r) \cdot m_2$  e  $U(r) = \phi_2(r) \cdot m_1$ .

A força gravítica, como sabemos, é uma força atractiva. Vamos verificar se o potencial  $\phi(r)$  que se mostra na figura resulta, realmente, numa força de atracção. A força é o simétrico da derivada da energia potencial em ordem a  $r$ . Vê-se na figura, que a derivada do  $\phi(r)$  (e da  $U(r)$ , também, obviamente) é positiva para qualquer  $r$ . A força é, então, negativa em todo o espaço ou, por outras palavras, o vector da força aponta sempre para  $r = 0$ .

Surge aqui uma pergunta: como é que as equações para a energia potencial acima concordam com a conhecida expressão  $U = mgh$ , que habitualmente usamos para calcular a energia potencial no campo gravítico à superfície da Terra? O segredo está nas palavras “à superfície da Terra”. O  $mgh$  é uma aproximação que só é válida às distâncias  $h \ll R_T$ , onde  $R_T$  é o raio da Terra ( $R_T \approx 6400$  km). Vamos demonstrar isso.

Usando a expressão geral e definindo a distância entre o centro da Terra e o objecto como  $r = R_T + h$ , temos para a energia potencial

$$U(h) = -G \frac{M_T m}{R_T + h}$$

em que  $M_T$  é a massa da Terra e  $m$  é a massa do corpo à distância  $h$  da superfície da Terra. Dividimos o numerador e o denominador por  $R_T$

$$U(h) = -G \frac{M_T m}{R_T} \cdot \frac{1}{1 + h/R_T} \approx -G \frac{M_T m}{R_T} \cdot (1 - h/R_T) = -G \frac{M_T m}{R_T} + G \frac{M_T}{R_T^2} \cdot m \cdot h$$

Usamos aqui a aproximação  $\frac{1}{1+x} \approx 1-x$  com  $x = \frac{h}{R_T} \ll 1$  (ver a caixa “Expansão

em série de Taylor”). O primeiro termo não depende do  $h$  ou, seja, é uma constante. O



A **Expansão em Série de Taylor** de uma função  $f(x)$  em torno do ponto  $x = x_0$  é dada pela fórmula

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + f''(x_0) \cdot \frac{(x - x_0)^2}{2} + \dots + f^{(n)}(x_0) \cdot \frac{(x - x_0)^n}{n!} + \dots$$

onde  $f^{(n)}(x_0)$  representa a derivada da função  $f(x)$  de ordem  $n$ ,  $f^{(n)}(x_0) = \frac{d^n f(x)}{dx^n}$ ,

calculada no ponto  $x = x_0$  (primeiro, calcula-se a derivada em função do  $x$  e depois o valor desta para  $x = x_0$ ; o resultado será, então, um número). A expansão substitui a função  $f(x)$  pela série infinita de polinómios. Esta pode ser uma boa solução quando se procura a descrição aproximada da função nas vizinhanças de um ponto. Dois ou três primeiros termos são suficientes em muitas situações.

Em física, o caso de  $x_0 = 0$  tem um interesse especial: para um  $x$  muito pequeno ( $x \ll 1$ ) podemos escrever

$$f(x) \approx f(0) + f'(0) \cdot x$$

(em alguns casos pode ser necessário adicionar  $+ f''(0) \cdot \frac{x^2}{2}$ ).

Exemplos:

$\sin x \approx x$	$e^{-x} \approx 1 - x$	$(1 \pm x)^n \approx 1 \pm nx$
$\cos x \approx 1 - \frac{x^2}{2}$	$\ln(1+x) \approx x$	$\sqrt{1+x} \approx 1 + \frac{1}{2}x, (n = \frac{1}{2})$
$\tan x \approx x$	$\frac{1}{1+x} \approx 1 - x$	$\frac{1}{(1 \pm x)^m} \approx 1 \mp mx$
$e^x \approx 1 + x$	$\frac{1}{1-x} \approx 1 + x$	$\frac{1}{\sqrt{1+x}} \approx 1 - \frac{1}{2}x, (m = \frac{1}{2} \text{ ou } n = -\frac{1}{2})$

segundo termo é proporcional a  $h$  e é esse que nos interessa. Designamos  $G \frac{M_T}{R_T^2} = g$  e temos, então,  $U(h) \approx \text{const} + mgh$ . Uma vez que, de ponto de vista prático, estamos interessados apenas na diferença entre as energias potenciais de um ponto para outro e não no valor da energia potencial em si, podemos omitir a constante e chegamos à expressão  $U(x) = mgh$ . A omissão da constante é equivalente à suposição que a energia potencial na superfície da Terra ( $h = 0$ ) é igual a zero.

Podemos tirar duas conclusões importantes deste cálculo:

- 1)  $U(x) = mgh$  é uma aproximação válida só quando  $h \ll R_T$  (i.e. à superfície da Terra);
- 2) A aceleração da gravidade  $g$ , a que habitualmente atribuímos um valor de  $9.81 \text{ m/s}^2$  não é, de facto, uma constante universal. A aceleração depende da massa do planeta e do seu raio:  $g = G \frac{M_T}{R_T^2}$ . Por exemplo, a aceleração da gravidade é diferente no equador e nos pólos porque a Terra não é uma esfera ideal – é um elipsóide ligeiramente achatado nos pólos. Além disso, torna-se claro como determinar a aceleração da gravidade à superfície de outro planeta: basta substituir a massa e o raio da Terra pela massa e pelo raio desse planeta.

### 3.2 Comparação do campo gravítico com o campo electrostático.

#### Potencial atractivo e potencial repulsivo.

A lei de Coulomb para a força electrostática é muito semelhante à lei de gravitação de Newton. Ambas as forças são inversamente proporcionais ao quadrado da distância; as cargas entram na equação de Coulomb do mesmo modo que as massas na lei de Newton. Há, no entanto, uma diferença substancial que resulta que alguns fenómenos naturalmente presentes no campo electrostático são impossíveis no campo gravítico.

Todas as observações existentes indicam que a interacção gravítica resulta sempre em atracção entre os corpos. Pelo contrário, a interacção electrostática pode resultar tanto em atracção como em repulsão entre as cargas, dependendo dos sinais dessas. Duas cargas com os sinais opostos atraem-se, mas, se forem ambas positivas

ou ambas negativas, repelem-se. A ausência da repulsão gravítica na Natureza obriga-nos a supor que a massa (“carga gravítica”) só pode ser positiva (ou só negativa, o que não afecta as equações, mas ... será que ficaríamos agradados se nos pedissem para pagarmos “menos 5 kg de batatas”?). A reflexão deste facto nas equações é um sinal “-“ na lei de Newton e um sinal “+” na lei de Coulomb.

Repare-se na forma das curvas da energia potencial  $U(r)$  (ou do potencial  $\phi(r)$ ) para a atracção e para a repulsão (ver a tabela que compara o campo gravítico com o campo electrostático). O potencial atractivo tem a derivada positiva (mas não é por ser negativo que é atractivo); o potencial repulsivo tem a derivada negativa.

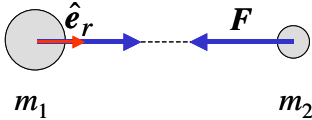
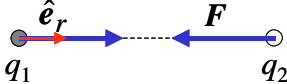
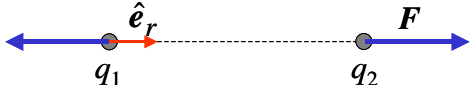
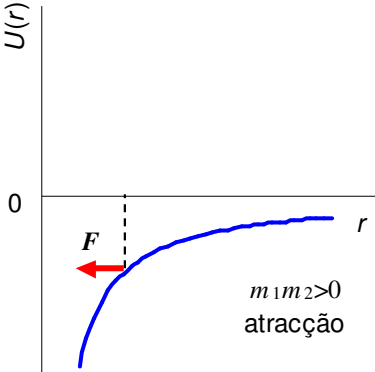
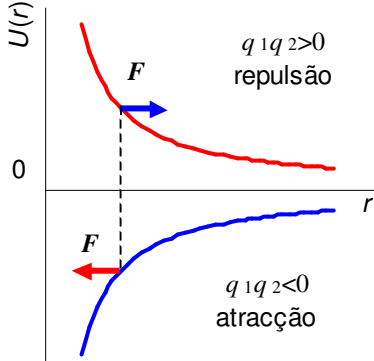
### **Forças gravíticas e forças electrostáticas no Universo**

A interacção electromagnética é muito mais forte do que a interacção gravítica. Comparando, por exemplo, a força de atracção gravítica com a força de repulsão electrostática entre dois electrões, chegamos a um valor inimaginável:

$$\frac{F_g}{F_e} = \frac{G}{k} \cdot \frac{m_e^2}{e^2} \sim 10^{-43} !$$

Pergunta-se, então, porque é que ao falarmos sobre o movimento dos corpos celestes só tomamos em conta a força gravítica ? A resposta está no facto de que existem duas variedades de cargas eléctricas e só uma de massa. A grande escala, as forças de atracção e de repulsão eléctrica entre as cargas anulam umas às outras, mas as forças gravíticas produzem um efeito cumulativo.



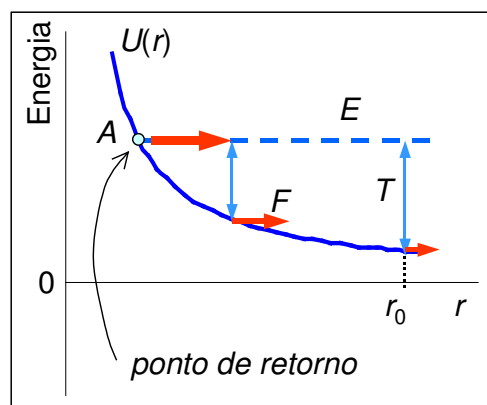
Campo gravítico	Campo electrostático
Força	
$\mathbf{F} = -G \frac{m_1 m_2}{r^2} \hat{\mathbf{e}}_r$ $(G = 6.67 \cdot 10^{-11} \text{ N} \cdot \text{m}^2 / \text{kg}^2)$  <p><math>m_1 &gt; 0, m_2 &gt; 0</math></p>	$\mathbf{F} = k \frac{q_1 q_2}{r^2} \hat{\mathbf{e}}_r$ $(k = \frac{1}{4\pi\epsilon_0} = 9 \cdot 10^{-11} \text{ N} \cdot \text{m}^2 / \text{C}^2)$  <p><math>q_1 \cdot q_2 &lt; 0</math></p>  <p><math>q_1 \cdot q_2 &gt; 0</math></p>
Energia potencial	
$U(r) = -G \frac{m_1 m_2}{r}$  <p><math>m_1 m_2 &gt; 0</math> atracção</p>	$U(r) = k \frac{q_1 q_2}{r}$  <p><math>q_1 q_2 &gt; 0</math> repulsão</p> <p><math>q_1 q_2 &lt; 0</math> atracção</p>
Potencial do campo	
$\varphi(r) = -G \frac{m}{r}$ <p>(potencial do campo gravítico criado pela massa <math>m</math>)</p>	$\varphi(r) = k \frac{q}{r}$ <p>(potencial do campo eléctrico criado pela carga <math>q</math>)</p>

## 4. Movimento num potencial

Vamos mostrar que o conhecimento da energia potencial de um corpo num campo, qualquer que seja a natureza deste, em função das coordenadas espaciais fornece a informação completa para a descrição do movimento neste campo.

### 4.1 Potencial repulsivo.

A figura ao lado mostra o gráfico da energia potencial de uma força repulsiva em função da coordenada,  $U(r)$ . O  $r$  aqui corresponde à distância ao centro da força – por exemplo, uma carga positiva colocada em  $r=0$  e uma outra, também positiva, a uma distância  $r$  variável. Este potencial é repulsivo porque em qualquer parte do



espaço o vector da força aponta para fora:  $F(r) = -\frac{dU(r)}{dr} > 0$ .

Repare-se, que o eixo vertical neste gráfico está em unidades de energia e, portanto, podemos representar no mesmo gráfico não só a energia potencial, mas também a energia mecânica,  $E$ , e a energia cinética,  $T$ . Suponhamos que o sistema é conservativo. Então,  $E = \text{const}$ , como se mostra na linha horizontal tracejada. A energia cinética,  $T(r) = E - U(r)$ , está marcada no gráfico com as setas verticais azuis.

Imaginemos que um corpo se encontra inicialmente a uma distância  $r_0$  do centro da força e tem a energia mecânica  $E$  e a energia cinética  $T \neq 0$ . Suponhamos que o corpo se move inicialmente no sentido  $r$  negativo. O que vai acontecer ao corpo à medida que se aproxima ao  $r=0$ ? Verifica-se que a energia cinética vai diminuindo gradualmente até chegar ao ponto A, em que  $T=0$  e  $E=U$ . No contexto de dinâmica, podemos dizer que a força, que actua no sentido contrário ao vector velocidade é que desacelera e faz parar o corpo.

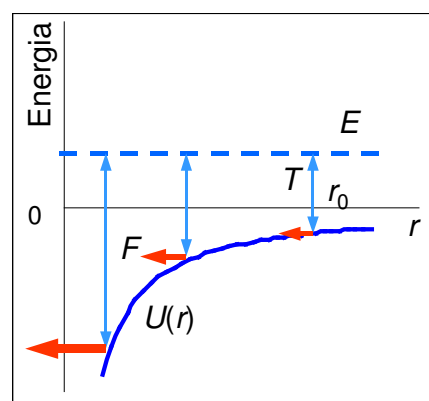
O que vai suceder quando o corpo parar no ponto A? A força continua a actuar no sentido  $r$  positivo resultando em aceleração no mesmo sentido. Isto significa que a paragem no ponto A é só por um instante e que no momento seguinte o corpo ganha a

velocidade no sentido positivo acelerando à medida que se afasta (a energia cinética aumenta com a distância). Por isso, o ponto *A* chama-se *ponto de retorno*. Depois de inverter o sentido do movimento, o corpo afasta-se para o infinito.

A coordenada do ponto de retorno depende, obviamente, da energia inicial do corpo: quanto maior for a energia no ponto de partida, mais próximo ao centro se consegue chegar. Esta é a explicação da utilidade dos aceleradores de partículas e da razão dos físicos necessitarem de energias cada vez mais elevadas para testarem a Natureza: para chegarem com as sondas cada mais perto do objecto de investigação.

## 4.2 Potencial atractivo.

Num potencial atractivo, que se mostra na figura ao lado, o comportamento do corpo é diferente. Uma partícula, situada inicialmente no ponto  $r_0$  e com o vector velocidade a apontar no sentido  $r$  negativo, vai continuar a mover-se no mesmo sentido, acelerando com uma aceleração crescente (em módulo) uma vez que o módulo da força aumenta à medida em que o corpo se aproxima do centro da força.



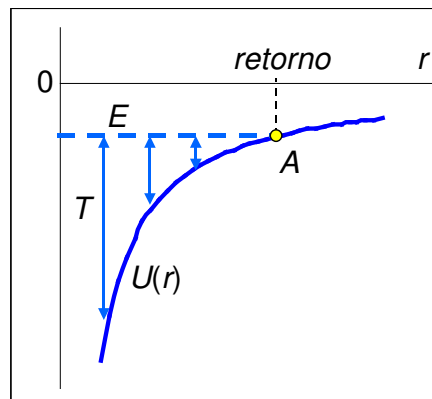
A energia cinética (setas verticais azuis na figura)

aumenta. Se  $U(r) \propto -\frac{1}{r}$  (lê-se proporcional), como é o caso da interacção gravítica e

da atracção electrostática entre duas cargas com sinais diferentes, a partícula cai, no final, para o centro da força,  $r = 0$ . Na realidade, pode existir um limite (distância mínima), tal como o raio da Terra no caso da queda de uma maçã, por exemplo. Aliás, a curva da energia potencial na figura corresponde só às distâncias  $r > R_T$ . A superfície da Terra funciona como uma barreira do potencial repulsivo impedindo a penetração do corpo ao cair para dentro do planeta. A natureza deste potencial é, no entanto, diferente - são as ligações entre os átomos constituintes da crosta da Terra que não deixam chegar ao centro do planeta e não as forças gravíticas.

Se no ponto inicial o corpo se movia no sentido  $r$  positivo, a força actuava no sentido contrário e a energia cinética diminuía na medida em que o corpo se afasta. Com a energia  $E > 0$ , como na figura acima, não existe nenhum limite para o movimento e o corpo segue para  $r = \infty$ .

A situação já é diferente se a energia mecânica é negativa, como se mostra na figura seguinte. Este facto, que a energia mecânica pode ser negativa, não deve surpreender-nos: como  $E = T + U$  e a energia potencial pode ser negativa, é perfeitamente possível que  $E$  seja negativa (a energia cinética,  $T = \frac{mv^2}{2}$ , é que



não pode ser negativa). Vê-se que a partícula, neste caso, não pode afastar-se para além do ponto A em que a energia cinética se torna igual a zero. A partícula fica confinada no espaço  $r < r_A$ . Assim, uma pedra atirada para cima volta a cair para o chão. A altura máxima, a que a pedra sobe, depende da velocidade inicial. Para uma velocidade inicial maior,  $E$  é maior e o ponto de retorno situa-se a um  $r$  também maior. Os níveis de energia correspondem aos estados ligados dos quais falaremos na secção seguinte. Para isso, temos de considerar um potencial mais realista.

Se a velocidade inicial for suficiente para assegurar  $E > 0$ , o corpo escapa ao campo gravítico da Terra e afasta-se para  $r = \infty$ .

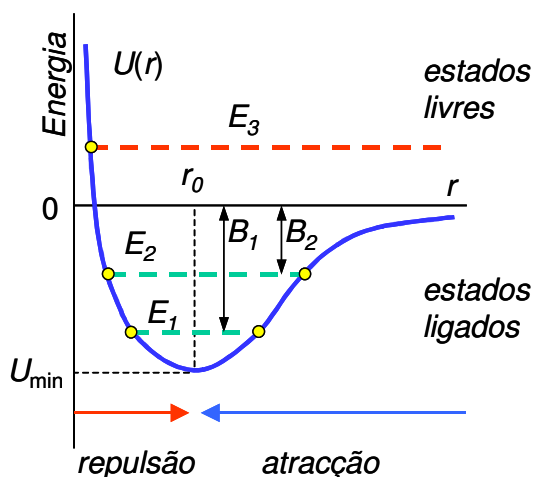
#### 4.3 Potencial realista – uma molécula diatómica.

Em sistemas reais, os potenciais puramente atractivos ou puramente repulsivos não se encontram. Por exemplo, a lei de Coulomb considera as cargas como sendo pontuais. Para os objectos extensos, a lei de Coulomb pode ser aplicada directamente apenas às distâncias muito superiores das dimensões dos corpos. Já no espaço mais próximo temos de dividir os objectos em partes pequenas (muito mais pequenas relativamente à distância), para poder aplicar a lei de Coulomb, e somar (integrar) as forças resultantes das interacções de cada elemento de um corpo com cada elemento de outro corpo. O mesmo é verdade para a lei universal de gravitação de Newton.

Para sermos mais concretos, consideremos uma molécula diatómica,  $H_2$ , por hipótese. Cada um dos dois átomos de hidrogénio é um sistema composto por duas partículas com carga eléctrica oposta: um protão, com carga positiva no centro, e um electrão, com carga negativa, a circular em volta do protão. Na molécula  $H_2$ , a distância entre os átomos (entre os protões, mais precisamente) é comparável às

dimensões de cada um, de modo que, apesar de os átomos serem sistemas electricamente neutros, não se pode ignorar o facto de estes serem feitos de partículas carregadas.

O cálculo da energia potencial deste sistema,  $U(r)$ , onde  $r$  é a distância entre os protões, não é simples. O resultado está representado qualitativamente na figura. A interacção entre os átomos tem carácter atractivo a grandes distâncias, mas às distâncias curtas aparecem forças repulsivas, o que se entende: se não existisse a parte repulsiva, o sistema sofria um colapso. A combinação das curvas potenciais de dois tipos resulta na formação de um vale (diz-se *poço potencial*).



Como de costume, é conveniente escolher a função da energia potencial de tal maneira que  $U(r) \rightarrow 0$  quando  $r \rightarrow \infty$  (note-se que a energia potencial está definida a menos de uma constante aditiva). Para a energia total mecânica,  $E$ , podemos distinguir duas regiões de valores possíveis nos quais o comportamento do sistema é muito diferente:  $E > 0$  e  $U_{\min} < E < 0$  ( $E < U_{\min}$  implicasse  $T < 0$  o que é impossível). No caso de  $E < 0$ , que corresponde aos níveis de energia  $E_1$  e  $E_2$  na figura, a possível distância entre os átomos está limitada por dois pontos de retorno em torno de  $r_0$  (como já mencionámos, a situação  $E < U$ , que se verifica para além desses pontos, não corresponde a nenhuma realidade física - na física clássica, pelo menos). Os dois átomos nos níveis  $E_1$  e  $E_2$  não podem afastar-se muito, nem aproximar-se muito, o que quer dizer que se forma um sistema diatómico estável – uma molécula. Aos estados de energia do sistema com essa propriedade chamam-se *estados ligados*, ao contrário dos *estados livres* (nível  $E_3$  na figura). Com  $E = E_3$  existe só um ponto de retorno; os átomos podem ser separados por qualquer distância excepto nas distâncias muito pequenas, i.e. ficam livres e não podem formar uma molécula.

De acordo com a física clássica, o sistema ligado pode existir em qualquer estado da energia  $U_{\min} < E < 0$  (não é bem assim na física quântica: esta prevê que os níveis de energia permitidos sejam discretos, mas a condição  $U_{\min} < E < 0$  mantém-se). Ao estado ligado com a energia  $E$  mais baixa possível no sistema chama-se *estado*

*fundamental* (na física clássica este coincide com  $E = U_{\min}$  mas num sistema quântico a energia do estado fundamental é  $> U_{\min}$ ). Os estados ligados com a energia superior à energia do estado fundamental chamam-se *estados excitados*.

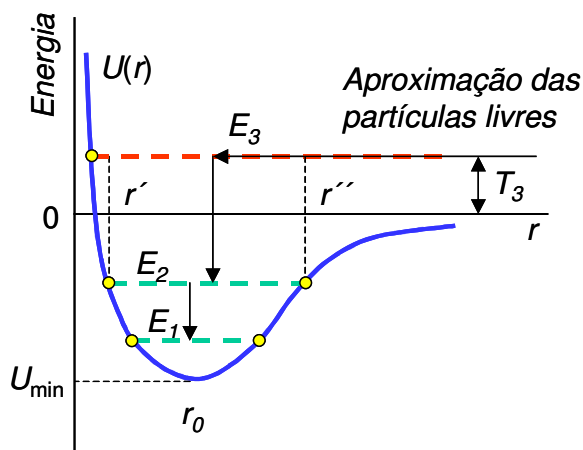
Uma molécula isolada permanecerá no estado fundamental durante um tempo infinito (sistema conservativo), mas se absorver alguma energia do exterior, por exemplo em forma da luz ou calor, pode aumentar a sua energia e subir para um nível excitado. Se a energia absorvida for suficientemente grande, o sistema pode saltar para um estado livre (por exemplo, de  $E_1$  para  $E_3$  na figura) em que já não existe limite superior para a distância entre os átomos e, assim, a molécula desintegra-se: temos um fenómeno de *dissociação*. A energia mínima necessária para transferir o sistema de um estado ligado para o estado livre chama-se *energia de ligação* deste estado ligado. A energia de ligação para os níveis de energia  $E_1$  para  $E_2$  é designada na figura por  $B_1$  e  $B_2$  – do Inglês *binding energy*. Verifica-se facilmente que  $B_i = -E_i$ .

Os estados excitados “vivem” um tempo finito determinado por um conjunto de factores que não podemos discutir aqui. O princípio geral da física diz que sempre que existe para um sistema físico a possibilidade de reduzir a sua energia, o sistema vai fazê-lo transferindo o excesso de energia para o exterior. No caso de uma molécula excitada, esta transferência ocorre por emissão da radiação electromagnética (luz visível, ultravioleta ou infravermelho) ou por colisão com outra partícula (outra molécula, por exemplo). A molécula neste processo não pode ser considerada um sistema conservativo, como é óbvio. A conservação da energia tem que contar com essa parcela da energia emitida para o espaço ou transferida para outro corpo.

Para um sistema mecânico, o equivalente deste processo de desexcitação seria, por exemplo, a perda da energia cinética devido ao atrito.

Para dissociar uma molécula, como já verificámos, é preciso fornecer energia à molécula do exterior. O processo inverso – a formação de uma molécula a partir de dois átomos livres – tem de ser acompanhado por libertação de energia. A figura a seguir ilustra como ocorre este processo.

Imaginemos dois átomos muito afastados a aproximarem-se um do outro com a energia cinética  $T_3$ . Como a energia potencial às distâncias grandes  $U \approx 0$ , a energia mecânica  $E_3 \approx T_3$ . Ao aproximar-se de uma distância comparável com o  $r$  que corresponde ao vale da energia potencial, torna-se possível



a transição para um nível de estado ligado. Suponhamos que a transição ocorre para o nível  $E_2$ . Neste nível, a distância entre os átomos permitida é entre  $r'$  e  $r''$  (ver a figura). Se no intervalo de tempo durante o qual os dois átomos a colidir ficam separados por uma distância  $r' < r < r''$  (é muito curto este tempo, de ordem de  $10^{-12}$  s) ocorre uma transferência da energia  $E_3 - E_2$  para o exterior, forma-se uma molécula. Essa energia pode ser transferida para o exterior do sistema, que é neste caso um sistema de dois átomos, pela forma da radiação electromagnética ou por colisão com uma terceira partícula. Se a molécula se formou no estado excitado, como neste exemplo, vai seguir uma transição para o estado fundamental, também com a libertação de energia.

### Pilhas de hidrogénio

Nos processos químicos reais, a situação complica-se pelo facto de os componentes necessários para a fusão raramente se encontrarem na natureza prontos para serem usados. Mais frequentemente, primeiro precisamos de quebrar algumas ligações, i.e. gastar energia.

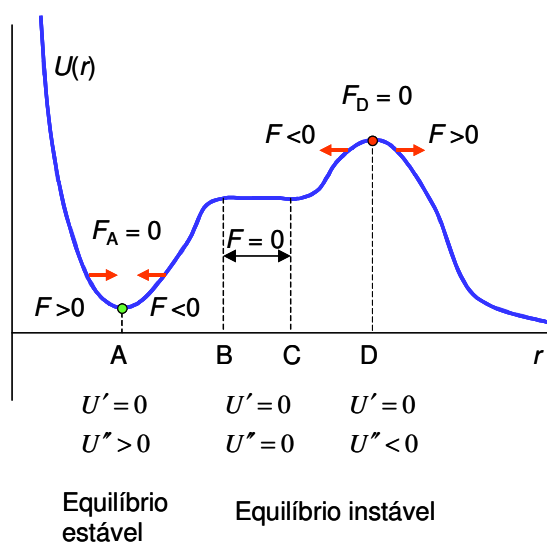
O hidrogénio pode vir a tornar-se uma fonte de energia importante num futuro muito próximo. A combustão do hidrogénio é, simplesmente, uma reacção de oxidação  $2H_2 + O_2 \rightarrow 2H_2O$ . A formação de moléculas de água ao juntar dois átomos de hidrogénio e um átomo de oxigénio é acompanhada pela libertação de energia de um modo idêntico, do ponto de vista físico, ao processo de formação da molécula  $H_2$ . Mas os átomos de hidrogénio e de oxigénio não existem na natureza isolados e para os obter têm que ser quebradas as ligações das moléculas  $H_2$  e  $O_2$ . O balanço energético global deste processo será, então,  $Q = 2 \cdot |B_{H_2O}| - 2 \cdot |B_{H_2}| - |B_{O_2}|$ , em que  $B$  é a energia de ligação da respectiva molécula (utilizamos a energia de ligação em módulo porque é negativa – ver primeira figura nesta secção). A reacção será exotérmica (com libertação de energia) se  $Q > 0$  (e neste caso é) ou endotérmica, se for  $Q < 0$ .

Este exemplo simples mostra como se ganha energia num processo químico, mas não só. Esta é a base física do qualquer processo de fusão incluindo a fusão nuclear.

## 4.4 Equilíbrio.

Mais um aspecto relacionado com o perfil da curva de energia potencial que precisamos de discutir antes de avançarmos para o movimento oscilatório: é a questão da existência de pontos ou regiões de equilíbrio. O significado da *posição de equilíbrio* é o seguinte. Se um corpo for colocado nesta posição com a velocidade inicial igual a zero, permanecerá nesta posição durante um tempo infinito (pressupõe-se, obviamente, que o sistema é isolado ou, seja, que não está sujeito a nenhuma força exterior). Para satisfazer esta condição não é suficiente que a velocidade inicial seja nula. É necessário que a aceleração também seja igual a zero. Caso contrário, mesmo se em  $t = 0$   $v = 0$ , já no instante seguinte o corpo vai ganhar a velocidade (recorde-se,  $v = v_0 + at$  ?). A condição  $a = 0$ , de acordo com a 2ª lei de Newton, significa que a força, exercida pelo campo sobre o corpo, no ponto de equilíbrio também tem que ser igual a zero.

A figura representa várias situações em que existe uma posição (ou uma região) de equilíbrio. Sabendo que  $F(r) = -\frac{dU(r)}{dr}$ , é fácil de identificar essas posições no gráfico de  $U(r)$ : são aquelas em que a derivada da energia potencial em ordem do  $r$  é igual a zero,  $U' = 0$ . Para a curva da figura esta condição está satisfeita no ponto A, na região entre os pontos B e C, e no ponto D.



O que vai suceder se desviarmos, um pouco, o corpo da posição de equilíbrio? Se a posição inicial era no ponto A, a força que aparecerá ao se desviar o corpo da posição de equilíbrio vai actuar no sentido do ponto A, tentando, assim, restaurar o equilíbrio (força restauradora). No ponto D isto não acontece: a força vai acelerar o corpo no sentido de o afastar do equilíbrio. Por isso, o equilíbrio no ponto A é

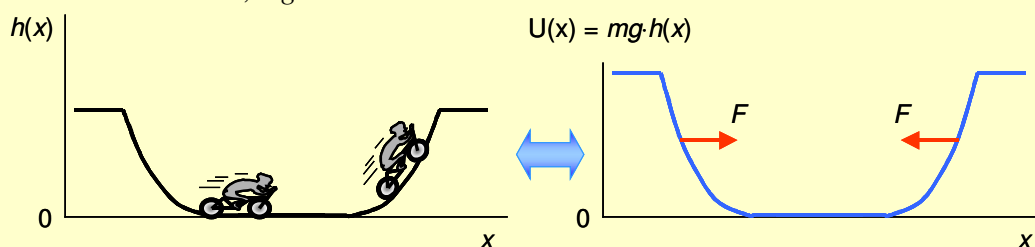


*equilíbrio estável* enquanto no ponto D é *equilíbrio instável*. No contexto matemático, o equilíbrio estável tem lugar no ponto em que a função  $U(r)$  tem um mínimo, ou seja, em que  $U' = 0$  e  $U'' > 0$  (a primeira condição é equivalente à  $F = 0$ ). No equilíbrio instável  $U' = 0$  e  $U'' < 0$ . Existe mais uma possibilidade:  $U' = 0$  e  $U'' = 0$ . Esta situação realiza-se na região BC. Se colocarmos um corpo, digamos no meio dessa região, e o desviarmos ligeiramente para a esquerda ou para a direita, não aparecerá nenhuma força, nem restauradora, nem destruidora. Mas, uma vez deslocado o corpo nunca mais voltará à posição inicial a não ser através de uma intervenção do exterior (no ponto A é a força do campo que restitui a posição inicial; na região BC a força é nula em toda a região excepto os pontos extremos). Por isso, o equilíbrio na região BC tem que ser considerado instável.

O movimento em torno de um ponto de equilíbrio estável tem um interesse especial. A força de restauração, que aparece sempre que o desviamos desta posição, resulta na tendência de voltar ao equilíbrio tornando o movimento a ser de carácter oscilatório. O movimento oscilatório é o assunto central do próximo capítulo.

### As curvas potenciais e $mgh$

A interpretação das curvas potenciais e do comportamento de um corpo no campo descrito por estas pode tornar-se mais claro se usarmos a analogia com o campo gravítico à superfície da Terra. A energia potencial no campo gravítico à superfície da Terra é directamente proporcional à altura acima da superfície ou de outro ponto de referência:  $U(h) = mgh$ . Imaginemos uma bicicleta numa rampa, cujo perfil pode ser descrito por uma função  $h(x)$ . Então, a energia potencial da bicicleta na rampa é uma função do  $x$ :  $U(x) = mg \cdot h(x)$ . Desta relação é óbvio que a forma dos gráficos  $h(x)$  e  $U(x)$  será idêntica, estando a diferença apenas no factor de escala constante,  $mg$ .



Deste modo, para compreender o movimento dos corpos no campo de uma força, qualquer que seja a sua natureza, podemos interpretar o gráfico  $U(x)$  como sendo o perfil de uma rampa com uma bicicleta em cima, com a vantagem de se saber muito bem como se move uma bicicleta numa rampa. O movimento dos corpos no campo real será semelhante !

## Capítulo 4

# Oscilações

Oscilação é um processo muito comum na Natureza. Oscilam as cordas de guitarra quando um músico toca, oscila o pêndulo de um relógio, oscila o carro quando apanha um buraco. Oscila também o *clock* de computador, determinando a frequência com que o processador executa as operações. O relógio atômico, o mais preciso que a tecnologia moderna pode oferecer, é baseado nas oscilações das camadas electrónicas do átomo de césio. As oscilações dão origem às ondas de todo o tipo e também estão envolvidas no processo de recepção do sinal que as ondas transportam.

Estes fenómenos físicos, embora de natureza diferente, têm muito em comum e seguem as leis semelhantes. Isto permite-nos estudar as características gerais das oscilações nos exemplos simples. Este é o caminho que vamos seguir neste capítulo. Vamos estudar um sistema mecânico simples, composto por uma massa e uma mola, começando por considerar um sistema sem atrito e introduzindo, passo a passo, os efeitos que aproximam um sistema idealizado a um sistema real. E não esqueçamos que o nosso objectivo é mais ambicioso do que estudar apenas este sistema concreto: é, sim, obter uma visão generalizada dos processos oscilatórios qualquer que seja a natureza destes.

## 1. Oscilações harmónicas simples

Consideremos, para começar, um sistema simples composto por um bloco de massa  $m$  que desliza sem atrito numa superfície horizontal e que está ligado a uma parede através de uma mola com a constante elástica  $k$  (ver figura). Quando a mola não está distendida nem comprimida, tem um comprimento  $l_0$  e o bloco fica em posição  $x_0$  sem se mover. Vamos agora pegar no bloco com a mão e desloca-lo para a posição  $x = x_0 + \Delta x$  de modo que a mola fica distendida com um comprimento  $l_0 + \Delta l$  (obviamente,  $\Delta l = \Delta x$ ). Quem fizer esta experiência verifica que é preciso fazer um

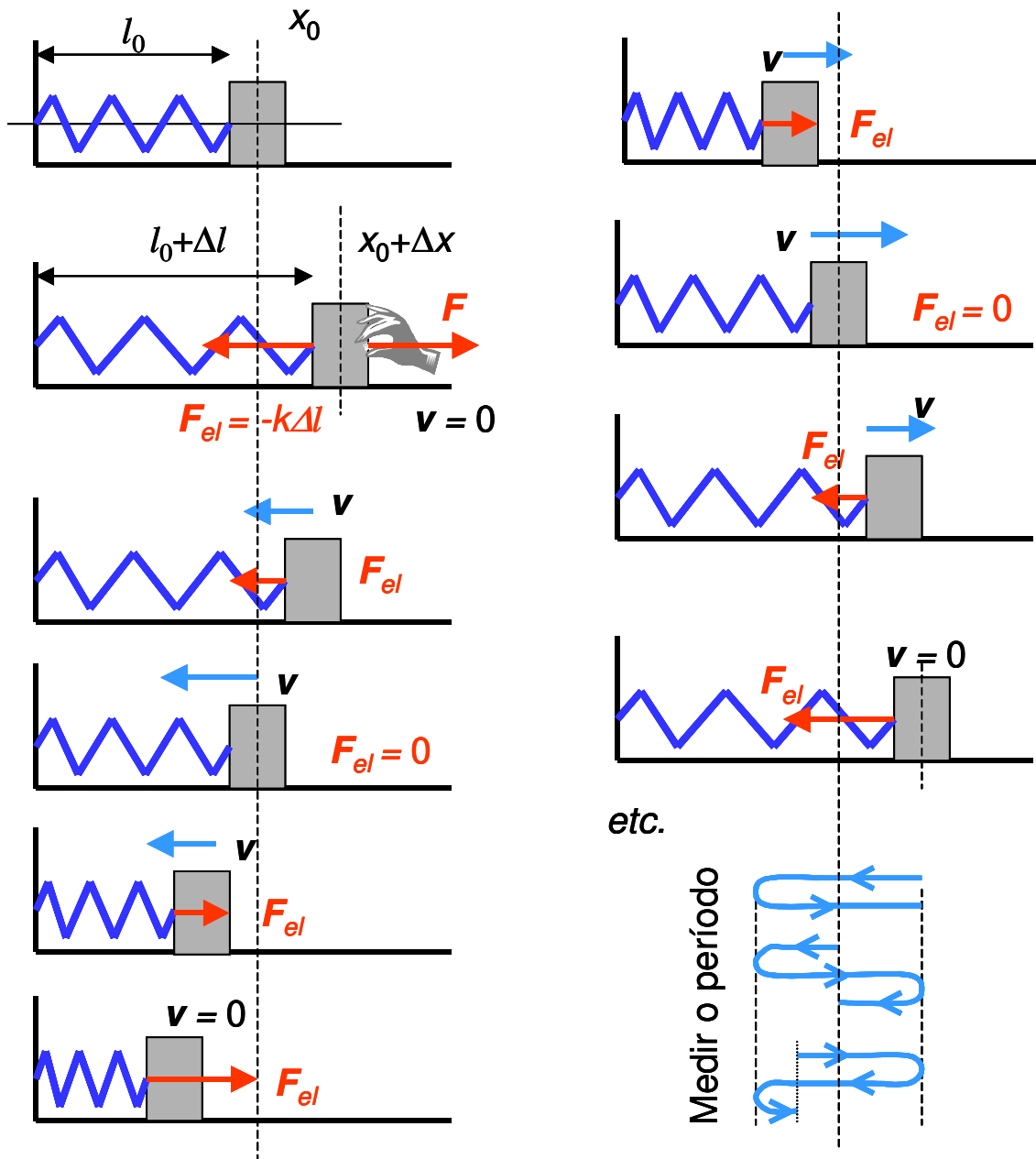
esforço para distender a mola porque esta tem tendência para voltar ao comprimento natural. Verifica-se, também, que este esforço é tanto maior quanto mais distendermos a mola. Em termos físicos, este facto experimental descreve-se com uma força, criada pela mola distendida e chamada *força elástica*, que actua no sentido contrário ao vector  $\Delta l$  e que é proporcional ao valor absoluto do desvio do comprimento natural  $l_0$

$$F_{el} = -k\Delta l$$

(quando esticamos a mola, a força elástica tenta comprimi-la, mas se a comprimirmos, a força elástica vai tentar expandi-la). Esta lei é também conhecida como a *lei de elasticidade de Hooke* pelo nome de um físico inglês Robert Hooke que a descobriu em 1660.

Quando largamos o bloco, este entra em movimento no sentido ao  $x_0$ , ganhando velocidade (ver figura). O valor absoluto da força elástica diminui à medida que o comprimento da mola se aproxima do seu comprimento natural,  $l_0$ . Ao passar o ponto  $x_0$ , o sentido da força elástica inverte-se. Agora, a força elástica em vez de acelerar o bloco, como no início, contraria o movimento fazendo com que o bloco desacelere e pare num certo ponto à esquerda do  $x_0$ . Mas a força elástica continua a actuar obrigando o bloco a voltar para trás, etc. O movimento terá um carácter repetitivo (periódico) e na ausência do atrito, como no nosso exemplo, continua durante tempo infinito. Este tipo de movimento chama-se movimento *oscilatório* e o ponto  $x_0$ , em que a força elástica é igual a zero, chama-se *ponto de equilíbrio*.

Podemos já introduzir os parâmetros que caracterizam o movimento oscilatório. O desvio máximo do bloco da posição de equilíbrio – é igual tanto à direita como à esquerda – chama-se *amplitude* de oscilação,  $A$ . O intervalo de tempo mínimo, em que o sistema volta ao estado inicial, chama-se *período*,  $T$ . É indiferente em que instante de tempo começa a contar o período, mas é importante que o sistema dê uma volta completa, i.e. regresse à mesma posição e se mova no mesmo sentido. Por exemplo, se começarmos a contar o tempo no instante em que largamos o bloco, o período será igual ao tempo que este demora a chegar à esquerda e voltar à posição extrema do lado direito. Podemos também começar a contagem quando o bloco passa à posição de equilíbrio movendo-se para a esquerda, mas neste caso o período vai incluir as passagens pelo ponto extremo do lado esquerdo, pela posição de equilíbrio a mover-se à direita, pelo ponto extremo à direita e, finalmente, na chegada à posição de equilíbrio outra vez.



O inverso do período chama-se *frequência*,  $f = \frac{1}{T}$ . A frequência é igual ao número de oscilações (voltas completas) que o sistema efectua por 1 segundo. A unidade de frequência é um Hertz sendo  $1 \text{ Hz} = 1 \text{ s}^{-1}$  (Heinrich Hertz é um físico alemão do século XIX que descobriu as ondas electromagnéticas).

O nosso objectivo é a descrição matemática do movimento oscilatório. Começamos por aplicar a segunda lei de Newton ao nosso sistema.

$$ma = -kx$$

onde do lado direito da equação já colocámos a força elástica (o peso do corpo e a força da reacção normal do plano horizontal anulam-se uma à outra e, por isso, não

estão incluídas na equação). O  $x$  aqui é a coordenada do bloco – escolhemos a origem do  $x$  na posição de equilíbrio para simplificar o cálculo (i.e.,  $x_0 = 0$ ). Lembrando que

$$a = \frac{d^2 x}{dt^2}, \text{ temos}$$

$$m \frac{d^2 x}{dt^2} = -kx.$$

Fazendo umas transformações simples e designando a segunda derivada do  $x$  em ordem ao tempo por  $\ddot{x}$  (é uma convenção habitual), chegamos a

$$\ddot{x} = -\frac{k}{m}x.$$

Convém designar a razão  $k/m$  por um parâmetro  $\omega_0^2 = \frac{k}{m}$ , cujo significado já vamos ver a seguir. É importante notar para efeitos de resolução da equação diferencial que como  $m$  e  $k$ , tanto  $\omega_0^2$  são constantes e não dependem do tempo.

Assim,

$$\ddot{x} = -\omega_0^2 x,$$

- eis a equação diferencial de segunda ordem cuja solução,  $x(t)$ , descreve a trajectória do corpo. Só nos falta resolvê-la... Na matemática existem métodos rigorosos para encontrar a solução (ou as soluções), mas vamos simplesmente tentar “adivinhar” a função cuja segunda derivada é igual ao simétrico da própria função:  $f(t)'' = -f(t)$  ignorando, para já, a constante  $\omega_0^2$ . Podemos invocar, também, algumas considerações físicas acerca das propriedades da função que estamos a procurar. Sabemos, por exemplo, que esta função deve ser periódica (porque tem que descrever um movimento repetitivo) e que o  $x$  oscila, em função do tempo, entre duas posições extremas passando pelo 0 no meio. Vendo de perto, parece que estamos a descrever um seno ou co-seno... Realmente, as funções  $\sin(t)$  e  $\cos(t)$  têm essas propriedades e ainda  $\sin(t)'' = -\sin(t)$  e  $\cos(t)'' = -\cos(t)$ . Falta ainda incluir na solução a constante  $\omega_0^2$ . É fácil verificar que  $\sin(\omega_0 t)'' = -\omega_0^2 \sin(\omega_0 t)$  e  $\cos(\omega_0 t)'' = -\omega_0^2 \cos(\omega_0 t)$ .

Verificamos, então, que ambas as funções  $x(t) = \sin(\omega_0 t)$  e  $x(t) = \cos(\omega_0 t)$  são soluções da equação de movimento. Qual aquela que preferimos? De facto, sabe-se da matemática (e é fácil de demonstrar), que se existirem várias soluções de uma

equação diferencial linear (i.e. que não contém produtos das derivadas), uma combinação linear dessas soluções também satisfaz a equação. Há ainda mais um pormenor que deve ser tomado em conta: a equação de movimento contém a segunda derivada do  $x$  em ordem ao tempo,  $\ddot{x}$ . Qualquer que seja o método que escolhemos para resolver a equação, o caminho da  $\ddot{x}(t)$  para o  $x(t)$  tem que passar inevitavelmente pela dupla integração. E, como sabemos, cada operação de integração vai fazer aparecer uma constante indefinida. Deste modo, o resultado final vai conter duas constantes.

Tendo em conta estas considerações, escrevemos a solução geral da seguinte forma

$$x(t) = a \cdot \sin(\omega_0 t) + b \cdot \cos(\omega_0 t)$$

onde  $a$  e  $b$  são constantes de integração. Deixa-se ao aluno a tarefa de verificar que a função acima é solução da equação diferencial  $\ddot{x} = -\omega_0^2 x$  independentemente dos valores das constantes  $a$  e  $b$  (para isso, substituem-se  $x$  e  $\ddot{x}$  na equação pela função  $x(t) = a \cdot \sin(\omega_0 t) + b \cdot \cos(\omega_0 t)$  e a sua segunda derivada, respectivamente, e verifica-se que a equação se transforma numa igualdade).

Repare que os valores das constantes  $a$  e  $b$  não estão definidos. A equação não nos diz nada acerca dos valores concretos. Ainda mais – quaisquer valores de  $a$  e  $b$  satisfazem a equação! Começámos por escrever a equação do movimento para um processo físico bem concreto, em que  $x(t)$  tem um valor exacto em qualquer instante de tempo, mas chegamos a uma função geral que abrange todos os movimentos que um sistema composto por um corpo de massa  $m$  e uma mola de constante elástica  $k$  pode ter! Agora, temos que encontrar uma maneira de determinar os valores das constantes para que a função  $x(t)$  descreva o nosso exemplo. A informação sobre esses valores não está na equação diferencial, mas sim nas condições iniciais, tais como, por exemplo, a posição inicial e a velocidade inicial, ou posição e aceleração iniciais ou em qualquer outro momento, ou ainda as posições em  $t_1$  e em  $t_2$ , etc. Como temos duas constantes a definir, precisamos de mais duas equações para além da equação diferencial.

Antes de avançarmos, vamos modificar a solução em forma  $x(t) = a \cdot \sin(\omega_0 t) + b \cdot \cos(\omega_0 t)$  para uma forma mais cómoda de analisar. Para isso,

em vez de um par de constantes independentes  $a$  e  $b$  introduzimos um outro par de outras constantes independentes  $A$  e  $\delta$  ligadas às  $a$  e  $b$  através das equações

$$\begin{cases} a = -A \cdot \sin \delta \\ b = A \cdot \cos \delta \end{cases}$$

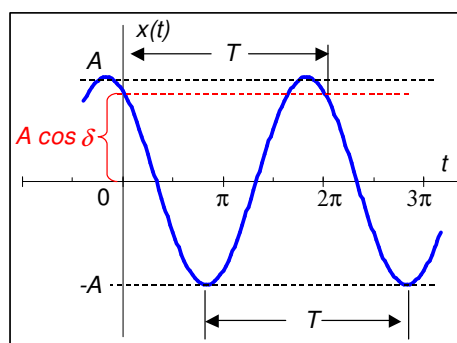
Deste modo, temos para a função  $x(t) = A \cos \delta \cdot \cos(\omega_0 t) - A \sin \delta \cdot \sin(\omega_0 t)$  ou

$$x(t) = A \cdot \cos(\omega_0 t + \delta)$$

(utilizamos aqui a fórmula de trigonometria  $\cos(\alpha + \beta) = \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta$ ). Esta é a forma mais habitual da função que descreve uma *oscilação harmónica simples*. O gráfico desta função está representado na figura ao lado. Daqui torna-se claro o significado dos parâmetros:  $A$  corresponde ao desvio máximo da posição de equilíbrio e chama-se *amplitude* de oscilação;  $\delta$  determina o valor da função em  $t = 0$ :  $x(0) = A \cdot \cos \delta$  e chama-se *fase inicial*. Em muitas situações pode-se escolher a origem do eixo do tempo de tal maneira que a fase inicial seja igual a zero. Neste caso,  $x(0) = A$ .

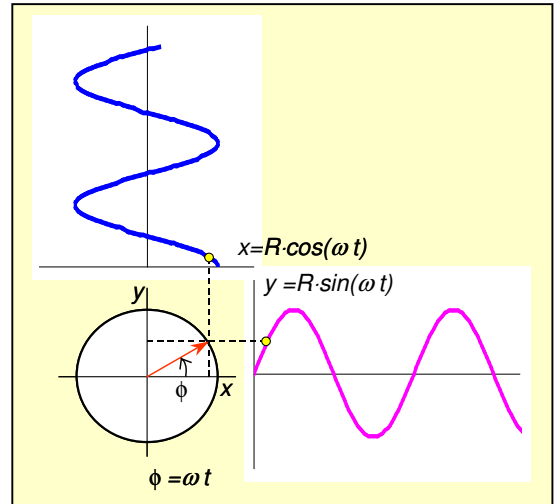
Sabendo o significado físico das constantes  $A$  e  $\delta$ , podemos perceber porque é que a equação diferencial não fornece informação sobre estas – é porque a amplitude  $A$  é determinada pela acção da nossa mão ao estender a mola e, portanto, somos nós, e não a segunda lei de Newton, que definimos a amplitude das oscilações. O valor da fase inicial  $\delta$  é apenas uma questão de escolha da origem do eixo do tempo.

Com a constante  $\omega_0$  é diferente. Esta é definida pela constante elástica da mola e pela massa do corpo e, portanto, não depende da nossa vontade, como é o caso da  $A$  e do  $\delta$ . É fácil verificar que  $\omega_0$  caracteriza a taxa (frequência) com que o sistema oscilatório volta ao estado inicial (i.e. a mesma posição e com o movimento no mesmo sentido): chama-se *frequência própria* do sistema. Muitas vezes dá-se a este termo um nome mais extenso – *frequência própria circular* – para a distinguir da frequência  $f$  introduzida acima (diz-se também *frequência linear*). O intervalo de tempo mínimo, necessário para que o sistema volte ao estado inicial, chama-se *período* de oscilação,  $T$ . Como o período do co-seno é igual a  $2\pi$ , podemos escrever a



### Frequência linear e frequência angular

Das duas frequências, a frequência linear tem um significado físico mais transparente – é o número de oscilações (voltas completas) que o sistema faz por segundo. A frequência angular é  $2\pi$  vezes isso. A razão pela qual surge o factor  $2\pi$  é muito simples: porque este é o período das funções trigonométricas seno e co-seno. Então, porque é que é *angular*? Isto tem a ver com o facto de o movimento oscilatório poder ser visto como projecção do movimento circular no plano  $xy$  para um dos eixos (ver a figura). O termo *angular* vem desta representação, o que faz todo o sentido para um movimento circular, mas introduz alguma confusão no caso do movimento oscilatório dos sistemas lineares, tais como uma massa e uma mola, por exemplo.



$$\omega = 2\pi f$$

$$f = \frac{1}{T}$$

seguinte relação  $\omega_0 T = 2\pi$ . Lembrando que  $T = \frac{1}{f}$  chegamos à relação entre as duas frequências  $f_0$  e  $\omega_0$

$$\omega_0 = 2\pi f_0$$

(o índice “0” sublinha que a frequência que está em causa é a frequência própria definida como  $\omega_0^2 = \frac{k}{m}$ ; como veremos mais tarde, há situações em que o sistema oscila com uma frequência diferente da frequência própria).

Como já verificámos no capítulo sobre cinemática, o conhecimento do  $x(t)$  permite-nos calcular outros parâmetros cinemáticos, nomeadamente a velocidade e a aceleração

$$x(t) = A \cdot \cos(\omega_0 t + \delta)$$

$$v(t) = \dot{x}(t) = -A\omega_0 \sin(\omega_0 t + \delta)$$

$$a(t) = \dot{v}(t) = -A\omega_0^2 \cos(\omega_0 t + \delta)$$

Daqui, podemos calcular a energia cinética associada ao sistema

$$T = \frac{mv^2}{2} = \frac{mA^2\omega_0^2}{2} \sin^2(\omega_0 t + \delta).$$

A partir desta equação verifica-se que a energia cinética varia em função do tempo entre 0 e o valor máximo de  $T_{\max} = \frac{mA^2\omega_0^2}{2}$ . Para calcular a energia potencial temos



que recordar a relação entre a força e a energia potencial associada a esta força,

$F(x) = -\frac{dU(x)}{dx}$ . Como a força neste sistema é a força elástica,  $F(x) = -kx$ , obtemos

a energia potencial com o seguinte cálculo

$$U(x) = -\int F(x)dx = k \int x dx = \frac{kx^2}{2}$$

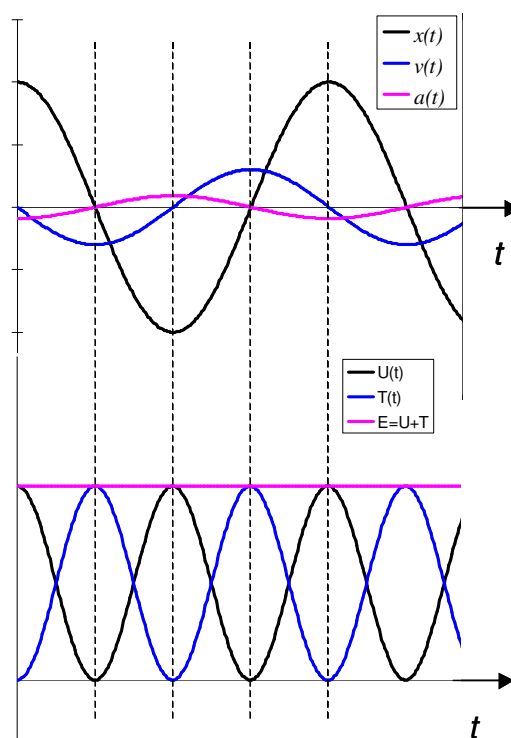
em que escolhemos a constante de integração  $C = 0$  por razões de conveniência. Ao substituir  $x$  por  $x(t)$  chegamos à energia potencial em função do tempo

$$U(t) = \frac{kA^2}{2} \cos^2(\omega_0 t + \delta).$$

Somando  $T(t) + U(t)$  e lembrando que  $\omega_0^2 = \frac{k}{m}$  (é assim que a definimos) temos a energia mecânica

$$\begin{aligned} E = T + U &= \frac{mA^2\omega_0^2}{2} \sin^2(\omega_0 t + \delta) + \frac{kA^2}{2} \cos^2(\omega_0 t + \delta) = \\ &= \frac{kA^2}{2} \sin^2(\omega_0 t + \delta) + \frac{kA^2}{2} \cos^2(\omega_0 t + \delta) = \\ &= \frac{kA^2}{2} \end{aligned}$$

i.e., a energia mecânica do sistema não depende do tempo, apesar de a energia cinética e a energia potencial variar, cada uma, em função do tempo. Por outras palavras, o sistema oscilatório composto por uma mola e uma massa é um sistema conservativo (convém lembrar aqui que a nossa suposição inicial era que não existe atrito no sistema). Como a força elástica é a única força que entrou na equação de movimento por nós composto,  $ma = -kx$ , podemos concluir que a força elástica é uma força conservativa. A variação do  $x(t)$ ,  $v(t)$  e  $a(t)$ , tanto como  $T(t)$  e  $U(t)$ , com o tempo mostra-se na figura ao lado.



## 2. Oscilações num sistema arbitrário

Vamos mostrar nesta secção que as oscilações harmónicas simples podem surgir em muitas situações, com forças de natureza muito diferente, e não apenas num sistema composto por uma mola e uma massa, como acabámos de ver na secção anterior. Para isso, consideremos um sistema conservativo abstracto, cuja energia potencial se descreve com uma função  $U(x)$ . Suponhamos também que esta função tem um mínimo em  $x = x_0$ , i.e. uma posição de equilíbrio estável. Como acabámos de verificar no capítulo anterior, o movimento em torno de um mínimo de energia potencial tem carácter repetitivo, periódico, devido à força restauradora que surge assim que o sistema é desviado do  $x = x_0$  e que actua no sentido de o devolver para a posição de equilíbrio. Este tipo de movimento chama-se, em geral, *movimento oscilatório*.

Vamos supor que os desvios do sistema da posição de equilíbrio em  $x = x_0$  são suficientemente pequenos para podermos utilizar a expansão da função  $U(x)$  em série de Taylor em torno do ponto de equilíbrio

$$U(x) = U(x_0) + U'(x_0) \cdot (x - x_0) + U''(x_0) \cdot \frac{(x - x_0)^2}{2} + \dots$$

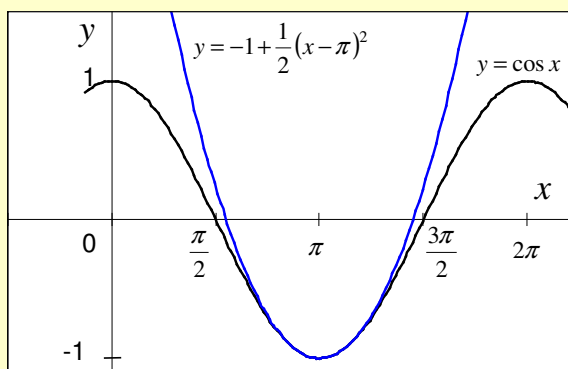
(recorde-se que a escrita  $U'(x_0)$  significa a primeira derivada da função  $U(x)$  calculada no ponto  $x = x_0$ , i.e.  $U'(x_0)$  é um número;  $U''(x_0)$  é a segunda derivada calculada em  $x = x_0$  etc.). O primeiro termo desta equação é uma constante e não tem interesse para a definição da energia potencial - podemos até ignorá-lo. O segundo termo é nulo porque a função  $U(x)$  tem um mínimo em  $x = x_0$  (essa era a nossa suposição inicial) e, portanto,  $U'(x_0) = 0$ . Só nos resta o terceiro termo, acerca do qual podemos dizer que este é positivo uma vez que no ponto em que uma função tem um mínimo, a segunda derivada é positiva. Podemos designar então,  $U''(x_0) = k > 0$ . Vamos também supor, para facilitar as contas, que  $x_0 = 0$ . É fácil constatar que não se perde nada ao fazer esta suposição, é apenas uma questão de escolha de um referencial adequado. Deste modo, a equação para a função  $U(x)$  torna-se

$$U(x) \approx \frac{1}{2} k x^2 + \text{const}.$$



### Um co-seno ou uma parábola ?

Um co-seno em pontos extremos aproxima-se muito bem de uma parábola. Se fizermos a expansão do co-seno em série de Taylor nas vizinhanças do ponto  $x = x_0$ , obtemos  $\cos x \approx \cos(x_0) - \sin(x_0) \cdot (x - x_0) - \cos(x_0) \cdot \frac{(x - x_0)^2}{2}$ . A figura mostra um exemplo para  $x_0 = \pi$ . Neste ponto,  $\cos \pi = -1$  e  $\cos \pi = 0$  de modo que  $\cos x \approx -1 + \frac{(x - \pi)^2}{2}$ .



Isto significa que, qualquer que seja a forma concreta da função correspondente à energia potencial  $U(x)$ , é sempre possível aproximá-la de uma parábola nas proximidades do ponto de equilíbrio estável. A forma concreta da função vai influenciar apenas os limites para o desvio do ponto de equilíbrio (amplitude de oscilações) dentro dos quais a aproximação pode ser considerada aceitável. Mas  $U(x) = \frac{1}{2}kx^2$  é a energia potencial de um sistema com força elástica  $F(x) = -kx$ , como vimos na secção anterior quando falámos de um sistema composto por uma mola (um elemento elástico) e uma massa (um elemento de inércia) ! Já sabemos que um sistema destes representa um oscilador harmónico simples cujo movimento se descreve com uma função harmónica (seno, coseno ou  $e^{\pm i\omega t}$ ).

Podemos concluir daqui que qualquer sistema que tenha uma posição de equilíbrio estável está sujeita às oscilações e estas oscilações têm carácter de oscilações harmónicas simples, desde que a amplitude seja pequena (i.e., a expansão em série de Taylor seja uma boa aproximação).

### 3. Oscilações amortecidas

As oscilações harmónicas simples são um caso idealizado. Um sistema pode oscilar com a mesma amplitude durante um tempo infinito apenas na ausência das forças não conservativas, tais como atrito e resistência do ar ou outros efeitos que levam à não conservação da energia mecânica como, por exemplo, o aquecimento dos componentes do sistema devido ao movimento. E, realmente, a única força que tomamos em conta quando escrevemos a equação do movimento  $m \frac{d^2 x}{dt^2} = -kx$  é a força elástica, que é uma força conservativa. Na realidade, como sabemos, existe sempre dissipação de energia que se traduz na diminuição da amplitude das oscilações (a nossa experiência do dia a dia permite-nos tirar esta conclusão mesmo sem efectuarmos cálculos). Como podemos tomar este efeito em conta?

A segunda lei de Newton permite fazê-lo facilmente. Basta acrescentar à força elástica mais uma força – a força de atrito

$$m \frac{d^2 x}{dt^2} = -kx - \alpha \frac{dx}{dt}.$$

A força de atrito é assumida aqui sendo proporcional à velocidade,  $F_{at} = -\alpha v$ , o que é verdade em muitas situações (nas velocidades grandes, no entanto, podem aparecer termos proporcionais às potências mais elevadas da velocidade). O sinal menos indica que a força de atrito actua no sentido oposto ao vector velocidade. O coeficiente de proporcionalidade  $\alpha$  é o coeficiente de atrito.

Dividindo ambas as partes da equação por  $m$  e introduzindo novas constantes  $\omega_0^2 = \frac{k}{m}$  (frequência própria, como antes) e  $\lambda = \frac{\alpha}{2m}$  (*coeficiente de amortecimento* – é um parâmetro que caracteriza a taxa com que as oscilações se apagam, como vamos ver a seguir), chegamos à equação

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = 0$$

que contém derivadas da segunda e da primeira ordem da função  $x(t)$  (utilizamos aqui as notificações habituais  $\dot{x} = \frac{dx}{dt}$  e  $\ddot{x} = \frac{d^2 x}{dt^2}$ ). Que função  $x(t)$  seria a solução para esta equação diferencial? Algumas propriedades desta função, podemos prever desde já. Em primeiro lugar, a função tem que ser periódica com o período igual ou próximo a

$T = \frac{2\pi}{\omega_0}$ , como no caso sem atrito. Além disso, sabemos que a amplitude tem que

diminuir com o tempo, com uma taxa relacionada, de alguma maneira, com o coeficiente de amortecimento  $\lambda$  (a presença do termo  $2\lambda\dot{x}$  na equação de movimento é a diferença principal comparativamente ao caso anterior). À semelhança da oscilação harmónica simples adivinhamos que a solução conterá, provavelmente, seno ou co-seno, ou um exponencial com expoente complexo que é, no fundo, uma combinação das funções trigonométricas. Tendo em conta estes argumentos vamos procurar uma solução da forma  $x(t) = ae^{i\omega t}$  com dois parâmetros,  $a$  e  $\omega$ , que ainda não sabemos bem o que são, mas já podemos dizer que  $a$  se mede com as mesmas unidades que o  $x$  e, portanto, é uma medida de deslocamento; o  $\omega$  tem unidades  $s^{-1}$  e deve estar ligado à frequência das oscilações. A função exponencial foi utilizada em vez do seno ou co-seno apenas porque é mais fácil trabalhar com ela do que com as funções trigonométricas.

As derivadas do  $x$  em ordem ao tempo são  $\dot{x} = i\omega ae^{i\omega t}$  e  $\ddot{x} = i^2\omega^2 ae^{i\omega t} = -\omega^2 ae^{i\omega t}$ , e ao substituir  $x$ ,  $\dot{x}$  e  $\ddot{x}$  na equação diferencial pelas respectivas funções temos

$$-\omega^2 ae^{i\omega t} + 2\lambda i\omega ae^{i\omega t} + \omega_0^2 ae^{i\omega t} = 0$$

que depois de cortar  $ae^{i\omega t}$  se transforma numa equação quadrática

$$-\omega^2 + 2\lambda i\omega + \omega_0^2 = 0$$

cujas duas soluções em ordem de  $\omega$  (o resto dos parâmetros conhecemos) são

$$\omega_{1,2} = i\lambda \pm \sqrt{\omega_0^2 - \lambda^2}.$$

Isto significa que existem duas funções, que satisfazem a equação diferencial,  $a_1 e^{i\omega_1 t}$  e  $a_2 e^{i\omega_2 t}$ . Repare que as constantes  $a_1$  e  $a_2$  não são necessariamente iguais. A equação diferencial não nos fornece informação nenhuma acerca destas constantes (o factor  $ae^{i\omega t}$  corta-se nas transformações acima independentemente do valor do  $a$ ). Mas isto já não nos surpreende: a solução de uma equação diferencial de segunda ordem deve trazer duas constantes de integração. Assim, a solução geral obtém-se como uma combinação linear dessas duas soluções

$$x(t) = a_1 e^{i\omega_1 t} + a_2 e^{i\omega_2 t}$$

ou, mais completo,

$$x(t) = a_1 e^{i(\lambda + \sqrt{\omega_0^2 - \lambda^2})t} + a_2 e^{i(\lambda - \sqrt{\omega_0^2 - \lambda^2})t}$$

ou, ainda,

$$x(t) = a_1 e^{(-\lambda + i\sqrt{\omega_0^2 - \lambda^2})t} + a_2 e^{(-\lambda - i\sqrt{\omega_0^2 - \lambda^2})t}$$

e, finalmente,

$$x(t) = e^{-\lambda t} \left[ a_1 e^{i\sqrt{\omega_0^2 - \lambda^2} t} + a_2 e^{-i\sqrt{\omega_0^2 - \lambda^2} t} \right].$$

Nesta expressão, a soma entre parênteses parece descrever a parte oscilatória (recorde-se,  $\frac{e^{ix} + e^{-ix}}{2} = \cos(x)$ ?). Isto é verdade se o termo  $\sqrt{\omega_0^2 - \lambda^2}$  for real, o que pode não se verificar na realidade. Há portanto, duas situações distintas a considerar.

*Situação 1:  $\lambda < \omega_0$  (amortecimento pequeno).*

Neste caso, a expressão sob raiz quadrado  $\omega_0^2 - \lambda^2$  é positiva e, portanto, podemos substituir a soma de dois exponenciais na expressão para  $x(t)$  por um co-seno. Para isso, substituímos um par de constantes independentes  $a_1$  e  $a_2$  por um outro par  $A$  e  $\delta$  (à semelhança com o que fizemos anteriormente):  $a_1 = \frac{1}{2}A_0 e^{i\delta}$  e

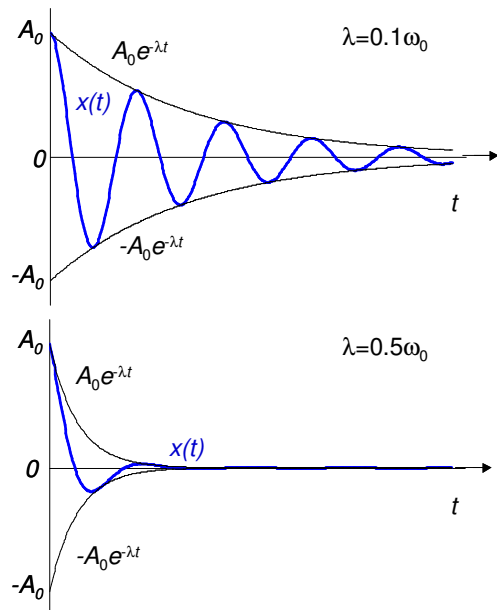
$a_2 = \frac{1}{2}A_0 e^{-i\delta}$ . Assim, temos

$$x(t) = A_0 e^{-\lambda t} \frac{1}{2} \left[ e^{i(\sqrt{\omega_0^2 - \lambda^2} t + \delta)} + e^{-i(\sqrt{\omega_0^2 - \lambda^2} t + \delta)} \right]$$

e

$$x(t) = A_0 e^{-\lambda t} \cos(\sqrt{\omega_0^2 - \lambda^2} t + \delta).$$

Esta é a função que descreve as oscilações amortecidas. Comparando com a respectiva função para as oscilações harmónicas simples, reparamos que a frequência angular já não é igual à frequência própria,  $\omega_0$ , mas é inferior a esta  $\omega = \sqrt{\omega_0^2 - \lambda^2}$ . A diferença é tanto maior



quanto maior for o coeficiente de amortecimento  $\lambda$ . Outro efeito do atrito, aliás já previsto por nós, é que a amplitude de oscilação agora é uma função do tempo,  $A(t) = A_0 e^{-\lambda t}$ . O coeficiente de amortecimento  $\lambda$  determina a rapidez com que diminui a amplitude. No limite do amortecimento desprezável, i.e. quando  $\lambda = 0$ , a função que descreve as oscilações amortecidas transforma-se numa equação de oscilações harmónicas simples com uma frequência igual à frequência própria do sistema  $\omega = \omega_0$  e a amplitude independente do tempo:  $x(t) = A_0 \cos(\omega_0 t + \delta)$ , como era de esperar. Os gráficos da função  $x(t)$  para dois valores do coeficiente de amortecimento estão representadas ao lado.

*Situação 1:  $\lambda > \omega_0$  (amortecimento significativo).*

Se for assim, a expressão sob raiz quadrado  $\omega_0^2 - \lambda^2$  é negativa e, consequentemente, a frequência  $\omega = \sqrt{\omega_0^2 - \lambda^2}$  tem um valor imaginário. Qual é a interpretação física deste facto? Vamos introduzir um parâmetro real  $\gamma = \sqrt{\lambda^2 - \omega_0^2}$ . Então,  $\omega = i\gamma$  e para  $x(t)$  temos

$$\begin{aligned} x(t) &= e^{-\lambda t} \{ a_1 e^{i(i\gamma)t} + a_2 e^{-i(i\gamma)t} \} \\ x(t) &= e^{-\lambda t} \{ a_1 e^{-\gamma t} + a_2 e^{\gamma t} \} \\ x(t) &= a_1 e^{-(\lambda+\gamma)t} + a_2 e^{-(\lambda-\gamma)t} \end{aligned}$$

Como  $\lambda$  é sempre maior que  $\gamma$  (o  $\gamma$  foi definido assim), ambos os termos descrevem um processo de decréscimo exponencial e não existe nenhuma componente oscilatória na equação. A equação descreve, neste caso, a situação em que o amortecimento é tão forte que o sistema está impedido de oscilar. Se o desviarmos da posição de equilíbrio e o largarmos, este voltará lentamente à posição inicial mas nunca a ultrapassa, como acontecia no caso do amortecimento fraco (ver a figura).

## 4. Oscilações forçadas. Ressonância.

Existem situações em que as oscilações surgem como resultado da acção de uma força periódica  $F(t)$  exterior ao sistema. O movimento de um baloiço empurrado por uma pessoa ou a oscilação de um prédio sob efeito de ondas sísmicas são dois exemplos. As oscilações deste tipo chamam-se *oscilações forçadas*. O período da variação da força exterior, à partida, não está correlacionada com o período com que o

sistema oscila sozinho – o período próprio do sistema  $T_0 = \frac{2\pi}{\omega_0}$  (lembramos que para

uma massa  $m$  ligada a uma mola de constante elástica  $k$  a frequência própria é  $\omega_0 = \sqrt{k/m}$ ). Como vai oscilar o sistema sob o efeito da força  $F(t)$  ?

Começamos, como sempre, por escrever a segunda lei de Newton. A força resultante é, neste caso, o somatório de três forças – da força elástica  $F_{el} = -kx$ , da força de atrito  $F_{at} = -\alpha v$  e da força exterior  $F(t)$ :

$$m\ddot{x} = -kx - \alpha\dot{x} + F(t)$$

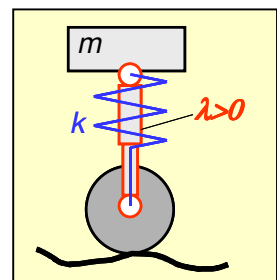
$$m\ddot{x} + \alpha\dot{x} + kx = F(t)$$

$$\ddot{x} + \frac{\alpha}{m}\dot{x} + \frac{k}{m}x = \frac{F(t)}{m}$$

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = \frac{F(t)}{m}$$

### Molas e amortecedores.

Um bom exemplo de um sistema oscilatório com amortecimento é a suspensão das rodas de um carro. As molas da suspensão, com o carro em cima, formam um sistema oscilatório com a frequência própria  $\omega_0 = \sqrt{\frac{k}{m}}$  em que  $m$  é a massa do carro e  $k$  é a constante elástica das molas. Quando andamos de carro, não queremos, obviamente, que este oscile ao apanhar cada buraco na estrada. Por isso, o sistema de suspensão tem os amortecedores – os elementos com  $\lambda > 0$  que reduzem as oscilações.





em que usamos as designações habituais  $\lambda = \frac{\alpha}{2m}$  (coeficiente de amortecimento) e

$\omega_0 = \sqrt{\frac{k}{m}}$ . Para que a resolução desta equação seja possível, temos que especificar a forma funcional da força exterior. Como sabemos da matemática, qualquer função periódica pode ser decomposta numa série dos *senos* e *co-senos*. Então, podemos substituir a força exterior por uma força harmónica  $F(t) = F_0 \cos(\omega_f t)$ , sem correr o risco de perder a generalidade na nossa caracterização do processo (repare-se que a frequência da força exterior  $\omega_f$  não é necessariamente igual a  $\omega_0$ ):

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = \frac{F_0}{m} \cos(\omega_f t)$$

A solução desta equação diferencial de segunda ordem não uniforme (porque tem uma função do lado direito  $\neq 0$ ) não é trivial. A oscilação forçada real inclui duas fases – a fase transitória, em que se iniciam as oscilações, e a fase estacionária, em que as oscilações se estabilizam e a amplitude se torna constante no tempo. Vamos concentrar o nosso interesse na fase estacionária e, em vez de mergulharmos na resolução da equação, apresentamos já a solução e dedicamos a nossa atenção à análise do resultado.

A solução estacionária da equação das oscilações forçadas com uma força  $F(t) = F_0 \cos(\omega_f t)$  é

$$x(t) = A(\omega_f) \cdot \cos(\omega_f t + \delta(\omega_f))$$

em que

$$A(\omega_f) = \frac{F_0/m}{\sqrt{(\omega_0^2 - \omega_f^2)^2 + 4\lambda^2 \omega_f^2}}$$

é a amplitude das oscilações forçadas e

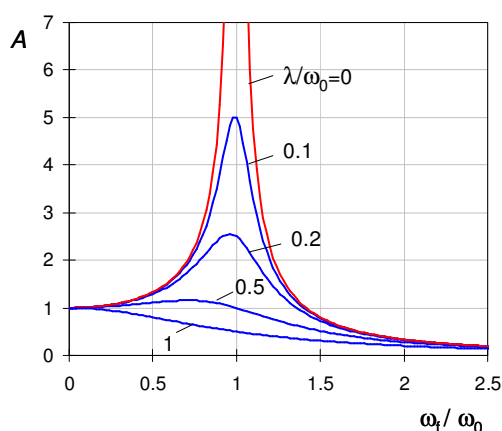
$$\delta(\omega_f) = \arctan\left(\frac{2\lambda\omega_f}{\omega_f^2 - \omega_0^2}\right)$$

é a fase, ambas dependentes da frequência da força exterior.

O facto de a amplitude das oscilações forçadas depender da frequência da força  $\omega_f$ , e não apenas da amplitude desta  $F_0$ , é notável. Ao analisarmos a expressão para  $A(\omega_f)$ , notamos que esta função tem um máximo em  $\omega_f = \sqrt{\omega_0^2 - 2\lambda^2}$  (ver a caixa)

$$A_{\max} = \frac{F_0/m}{2\lambda\sqrt{\omega_0^2 - \lambda^2}}.$$

O gráfico da amplitude em função da frequência da força exterior para vários valores de  $\lambda$  é representado ao lado. A amplitude máxima é tanto maior quanto menor for o amortecimento e tende para o infinito com  $\lambda \rightarrow 0$ . Este fenómeno de aumento significativo da amplitude de oscilações, quando a frequência da força exterior se aproxima da frequência própria do sistema, é designado por *ressonância*.



O efeito de ressonância tem um papel muito importante em todas as áreas da ciência e tecnologia e até na nossa vida. Todos os sistemas mecânicos e eléctricos, atómicos e nucleares são sistemas oscilatórios, que têm uma ou várias frequências

#### O máximo da função $A(\omega_f)$ .

Sabe-se que no ponto em que uma função tem um máximo ou um mínimo a derivada da função é igual a 0. Então, o que temos de fazer para encontrar o máximo ou o mínimo da  $A(\omega_f)$  é calcular o valor de  $\omega_f$  em que  $\frac{dA(\omega_f)}{d\omega_f} = 0$ :

$$\frac{dA(\omega_f)}{d\omega_f} = \frac{F_0}{m} \cdot \frac{1}{2} \cdot \frac{2(\omega_0^2 - \omega_f^2) \cdot (-2\omega_f) + 8\lambda^2 \omega_f}{((\omega_0^2 - \omega_f^2)^2 + 4\lambda^2 \omega_f^2)^{3/2}} = 0$$

Para que esta expressão seja igual a 0, o numerador tem que ser igual a 0 (o denominador nunca é =0):

$$\begin{aligned} 2(\omega_0^2 - \omega_f^2) \cdot (-2\omega_f) + 8\lambda^2 \omega_f &= 0 \\ (-(\omega_0^2 - \omega_f^2) + 2\lambda^2) \cdot \omega_f &= 0 \end{aligned}$$

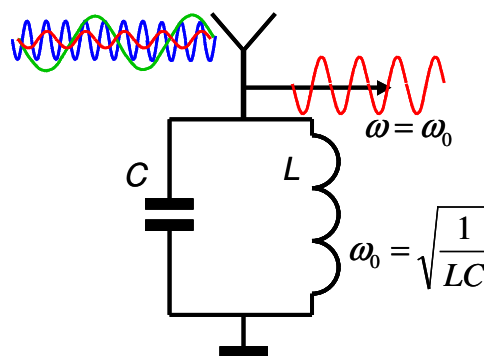
Esta equação tem duas soluções: a solução  $\omega_f = 0$  não nos interessa porque corresponde ao caso estático; a outra dá  $\omega_0^2 - \omega_f^2 = 2\lambda^2 \Rightarrow \omega_f^2 = \omega_0^2 - 2\lambda^2$ .

Para provar que esta frequência corresponde a um máximo e não a um mínimo da amplitude, em princípio, temos que calcular a segunda derivada da  $A(\omega_f)$  e mostrar que ela é negativa neste ponto. Neste caso concreto não há necessidade de o fazer: vê-se logo que  $A(\omega_f)$  aumenta quando  $\omega_f$  se aproxima a  $\omega_0$ , seja de um lado ou do outro (por causa do termo  $(\omega_0^2 - \omega_f^2)^2$ ).

próprias, e, quando estão sujeitos a uma excitação do exterior, podem entrar em ressonância. O efeito da ressonância pode ser tanto benéfico como altamente destrutivo. Eis apenas alguns exemplos.

### Ressonância ao serviço do homem.

**Rádio.** Há milhares de estações de rádio à nossa volta a emitir ondas electromagnéticas para espaço, mas com um bom receptor de ondas conseguimos sintonizar a que mais gostamos. Como? Como conseguimos distinguir sinais das emisoras diferentes que induzem na antena do nosso rádio (apenas um pedaço de um fio!) correntes de praticamente



igual amplitude? As ondas são emitidas com frequências diferentes (por exemplo, a estação “90FM” emite na frequência 90 MHz). A antena está ligada a um circuito eléctrico composto por uma bobina e um condensador - circuito  $LC$  ( $L$  é a indutância da bobina e  $C$  a capacitância do condensador) que é um sistema oscilatório com frequência própria,  $\omega_0 = \sqrt{\frac{1}{LC}}$ . O que oscila neste caso é a corrente eléctrica. Todas

as ondas são igualmente recebidas na antena e transmitidas ao circuito  $LC$  induzindo neste as oscilações. Se a frequência dos sinais vindos do exterior for longe da frequência  $\omega_0$ , estes terão pouco efeito sobre o oscilador - a corrente no circuito continua a oscilar, mas com uma amplitude insignificante. Ao mesmo tempo, os sinais cuja frequência coincide com a frequência própria do sistema, resultam em oscilações com uma amplitude muito maior. Portanto, o circuito  $LC$  permite seleccionar a frequência desejada num mar de ondas à nossa volta. Basta ajustar  $\omega_0$ , variando a indutância ou a capacitância dos elementos do circuito.

No circuito  $LC$ , a indutância da bobine  $L$  representa o elemento de inércia e o inverso da capacitância  $1/C$  é equivalente à “elasticidade” – dois componentes necessários para que um sistema seja capaz de efectuar as oscilações (recorde-se que, para uma mola  $\omega_0 = \sqrt{\frac{k}{m}}$ ). O amortecimento, que num sistema real existe sempre, é

devido à resistência eléctrica dos elementos do circuito  $R$  (dos fios da bobina e das ligações). Quanto menor for  $R$ , maior a ressonância, i.e. mais alto e mais agudo o pico no gráfico da amplitude das oscilações em função da frequência da força exterior (neste caso da onda electromagnética incidente na antena) e, conseqüentemente, melhor a selectividade do aparelho.

*Micro-ondas.* Nunca se perguntou a si próprio porque é que as micro-ondas aquecem a água num copo de vidro mas não aquecem o copo? O segredo está, outra vez, na ressonância. A frequência das micro-ondas nos fornos de cozinha está escolhida para ser próxima à frequência de rotação das moléculas de água no meio líquido (os fornos de cozinha usam a frequência de 2,45 GHz;  $\lambda \approx 12,5$  cm). Por isso, as substâncias que têm água na sua composição absorvem as micro-ondas com uma eficiência muito superior às que não têm.

#### *Os perigos da ressonância.*

As obras de construção civil devem contar com as forças exteriores periódicas, na origem das quais podem estar tanto processos naturais (terramotos, vento) como a actividade humana (assim como um grupo de soldados a marchar numa ponte ou as vibrações devido ao trânsito das máquinas pesadas, tais como um comboio, por exemplo). Um exemplo de um desastre muito conhecido é a queda da ponte suspensa sobre o estreito de Tacoma (*Tacoma Narrows*) nos Estados Unidos em 1940. Com um vento relativamente forte, cerca de 60 km/h, o tabuleiro da ponte começou a oscilar e acabou por cair. O vento era forte mas constante e, no entanto, foi este vento que deu origem às oscilações. Inicialmente, o vento causou as turbulências do ar em volta do tabuleiro da ponte. As turbulências do ar resultaram em pequenos desvios do tabuleiro da posição de equilíbrio que por sua vez afectaram as turbulências e, de certo modo as oscilações. Criou-se um *feedback* positivo no sistema, a estrutura já não conseguia dissipar a energia fornecida do exterior e a ponte ruiu.

Este desastre é o resultado de um conjunto de processos complexos em que não é fácil isolar o efeito da ressonância na sua forma pura. No entanto, a lição aprendida pelos engenheiros foi que é extremamente importante saber calcular as frequências próprias dos diferentes elementos do sistema e projectar a obra de tal maneira que estas estejam longe das frequências de eventuais forças exteriores, qualquer que seja a sua origem. Os dispositivos de amortecimento são importantes, também – estes

permitem absorver a energia das oscilações e são especialmente importantes quando a frequência da força é próxima da frequência da ressonância (recorde-se que quanto maior for o coeficiente de amortecimento, menor a altura do pico de ressonância no gráfico de amplitude em função da frequência da força exterior).

Um efeito de natureza semelhante é a vibração das asas dos aviões ou de alguns dos seus componentes (dos *flaps*, por exemplo) cuja origem está também nas turbulências do ar ao contornar o aparelho em vôo. Este efeito é bem conhecido pelos construtores de aeronaves desde há muito tempo (tendo já causado algumas quedas) mas ainda não está completamente compreendido. O nome genérico que se dá a este fenómeno é *flutter*, que significa vibrações, pulsações, movimento irregular, agitação, etc.

## Capítulo 5

# Ondas

Neste capítulo vamos discutir as propriedades gerais das ondas. As ondas são um fenómeno físico sem o qual a vida do homem moderno (e a vida em geral) seria impensáveis. De facto, desde sempre que encaramos as ondas todos os dias: uma pedra que cai na água calma de um lago provoca ondas na superfície da água; a luz do Sol, sem a qual não existiria vida na Terra, é uma onda electromagnética; a onda sonora é um importante meio de comunicação entre os seres vivos. Dois dos cinco sentidos do homem são baseados na detecção das ondas.

A descoberta das ondas electromagnéticas por Heinrich Rudolf Hertz em 1888 e o desenvolvimento que se seguiu revolucionaram a vida do homem. O rádio, a televisão, telemóveis, o forno de microondas, GPS, radares etc., etc. Têm todos por base a utilização das ondas de natureza electromagnética.

Além destes exemplos clássicos dos fenómenos ondulatórios existem outros, menos óbvios mas talvez mais importantes. No início do século XX percebeu-se que a Natureza em geral é dualística: qualquer objecto da Natureza não é nem um corpo nem uma onda, mas sim os dois ao mesmo tempo. Este é o princípio de dualidade onda-partícula da matéria em geral. É verdade, que ao se descrever o movimento da Terra à volta do Sol ou de uma bola num campo de ténis, não passa pela cabeça de ninguém pensar nestes como sendo ondas e não corpos. Mas isto é apenas porque as propriedades corpusculares da matéria dominam a este nível macroscópico e mesmo no mundo microscópico a representação de um electrão como apenas uma partícula por vezes pode ser suficiente. No entanto, trata-se sempre de uma aproximação que pode ser boa ou menos boa, dependendo das condições do problema e dos objectivos práticos. Porém, existem exemplos de situações contrários. A experiência com uma fenda dupla, que dá o mesmo resultado tanto com as ondas da luz (vamos discuti-la no fim deste capítulo) como com os electrões (no capítulo a seguir), é um excelente exemplo.

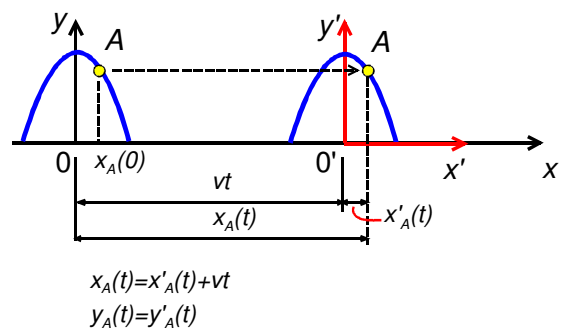
Finalmente, um conceito básico da mecânica quântica é uma função chamada função de onda por ter carácter oscilatório e por se propagar no espaço como uma onda. O significado físico desta função de onda é muito diferente da das ondas clássicas – esta

caracteriza a distribuição no espaço da probabilidade de detectar um objecto numa dada região do espaço. Seja como for, esta função segue as mesmas regras gerais que se aplicam também às outras ondas.

## 1. Propriedades gerais das ondas.

Pode-se definir uma onda como uma oscilação que se propaga no espaço. As ondas do mar, as ondas sonoras e as ondas electromagnéticas são os exemplos mais comuns deste fenómeno. Um processo ondulatório, por definição, desenvolve-se no espaço e no tempo e pode ser descrito com uma função de dois argumentos,  $\Phi(\mathbf{r}, t)$ . Em princípio, todos os processos físicos têm lugar no espaço e no tempo tal, por exemplo, como as oscilações de um corpo pendurado numa mola, cujo movimento se descreve com a posição vertical em função do tempo  $z(t)$ . Mas em contraste com a oscilação de uma mola, que está confinada no espaço dentro de certos limites, uma onda estende-se por todo o espaço até ser absorvida por obstáculos ou pelo meio em que se propaga.

Quando se fala de uma onda, normalmente referimo-nos a um processo periódico, quer no espaço, quer no tempo. No entanto, pode ser considerada também uma onda constituída apenas por um impulso que em muitos aspectos é semelhante a uma onda periódica.



Consideremos, por exemplo, um impulso a propagar-se ao longo de uma corda gerado por abanar a corda com a mão uma única vez. Para tornar a discussão mais concreta, descrevamos a forma do impulso com uma parábola invertida como se mostra na figura. Suponhamos que no instante  $t = 0$  o impulso está em posição  $x = 0$  e a forma deste se descreve com a função  $y(x, 0) = a^2 - x^2$  em que  $a$  é uma constante que determina largura e altura do impulso. Se a forma do impulso não varia na medida em que este evolui ao longo da corda, podemos utilizar a mesma função para o descrever em qualquer instante de tempo. Para isso, basta substituir  $x$  por  $x'$ , onde  $x'$  é a coordenada do pico do impulso num novo referencial que se desloca com o impulso com a mesma velocidade e, em que, por consequência, o impulso fica imóvel, i.e.  $y(x', t) = a^2 - x'^2$

(note-se que  $y$  não depende do  $t$  neste referencial). Este novo referencial fica a uma distância igual a  $vt$  do antigo em qualquer instante de tempo. A relação entre o  $x$  e o  $x'$ , pode ser verificada comparando as coordenadas de um ponto arbitrário do impulso (ponto  $A$  na figura) num referencial e no outro. Daí vem  $x = x' + vt \Rightarrow x' = x - vt$  e ao substituir  $x'$  em  $y(x', t) = a^2 - x'^2$  temos  $y(x, t) = a^2 - (x - vt)^2$ .

Repare, que o termo  $vt$  que aparece na função não tem nada a ver com a forma concreta da função escolhida para o exemplo, mas é apenas o resultado da transformação de coordenadas do referencial em que o impulso está imóvel no referencial em que ele se move com a velocidade  $v$ . Isto leva-nos a uma conclusão importante: qualquer que seja a função, que descreve uma onda, a coordenada espacial e o tempo entram no argumento da função em combinação  $x - vt$ . Se o impulso se desloca para a esquerda,  $v$  substitui-se por  $-v$  e, então,  $x - vt$  muda para  $x + vt$ .

Agora voltemos às ondas periódicas. Matematicamente, a periodicidade no espaço com um período espacial, digamos  $\lambda$ , e no tempo com período  $T$  significa que a função de onda satisfaz as seguintes condições

$$\Phi(x + \lambda, t) = \Phi(x, t) \text{ e}$$

$$\Phi(x, t + T) = \Phi(x, t) .$$

Já conhecemos o parâmetro  $T$  do capítulo anterior sobre as oscilações – é o *período*. O  $\lambda$ , que caracteriza a periodicidade no espaço, é designado por *comprimento de onda*.

Para uma função de onda  
 $\Phi(x, t) = \Phi(x \pm vt)$

Ao combinar estas duas equações numa e utilizando o facto de  $\Phi(x, t) = \Phi(x - vt)$  (suponhamos por uns instantes que a onda se move para a direita) temos

$$\Phi(x + \lambda, t + T) = \Phi(x, t)$$

$$\Phi[(x + \lambda) - v(t + T)] = \Phi(x - vt)$$

$$\Phi[(x - vt) + (\lambda - vT)] = \Phi(x - vt) .$$

Esta igualdade tem que ser válida para qualquer instante de tempo e para qualquer  $x$ . Logo,  $\lambda - vT = 0$  e

$$\lambda = vT .$$

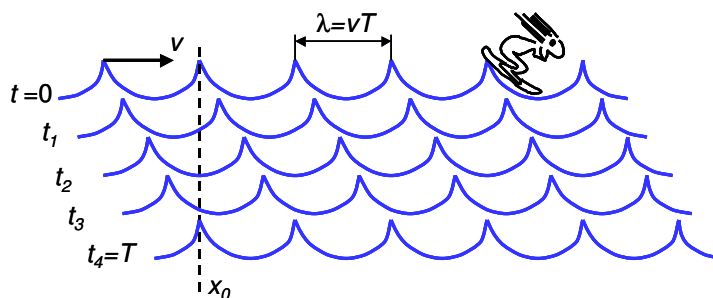


Lembrando agora que o período é igual ao inverso da frequência  $T = \frac{1}{f}$  chegamos a uma relação fundamental entre o comprimento de onda, a frequência da onda e a velocidade com que esta se propaga no espaço

$$v = \lambda f .$$

Relação entre a velocidade de onda, o comprimento de onda e a frequência desta

$$v = \lambda f$$



Uma questão pertinente é saber qual o significado físico da função  $\Phi(x,t)$ ? Isto depende da natureza da onda em causa. Para ondas do mar é a altura da onda,  $\Phi(x,t) = z(x,t)$ , para ondas sonoras pode ser ou o deslocamento das moléculas do meio da posição média ou a variação local da pressão,  $\Delta p(x,t)$ , se o som se propaga num gás. No caso das ondas electromagnéticas é o vector de intensidade do campo eléctrico ou do campo magnético,  $\mathbf{E}(x,t)$  e  $\mathbf{B}(x,t)$ .

Voltaremos ainda a falar sobre estes exemplos concretos das ondas, mas por agora vamos abordar outra questão importante. As ondas, qualquer que seja a natureza destas, estão claramente ligadas a um movimento: propagam-se no espaço e têm uma velocidade associada à propagação. O que é que se move, o que é que se propaga, exactamente? Estamos falar da velocidade do quê? Quando se fala de um movimento, surge logo a imagem de um corpo, de uma massa, a deslocar-se de um ponto do espaço para outro. Mas as ondas não transportam a massa: uma gaivota a flutuar nas ondas do mar não está a ser levada pelas ondas – a gaivota permanece no mesmo sítio, move-se apenas na vertical. Um rádio a tocar música não faz com que as ondas sonoras emitidas por este transportem o ar de um sítio para o outro. Não são as mesmas moléculas que estão em contacto com a membrana do altifalante, a bater nos ouvidos dos ouvintes, por vezes à distância de muitos metros. As ondas transportam a energia, isso sim. Caso contrário, não se poderia explicar como é que uma onda sonora faz vibrar a membrana do nosso ouvido e porque é que a luz do Sol aquece. Além disso, as ondas podem transportar informação. Este será o assunto que vamos considerar na secção 4 deste capítulo.

Acerca da velocidade da onda podemos dizer, para já, que esta é a velocidade com que se move uma crista da onda, um vale ou um outro ponto qualquer na onda. Mas esta não é a velocidade com que se desloca a água – esta oscila só na vertical. Na secção seguinte vamos dar uma definição mais formal da velocidade da onda.

## 2. Ondas harmónicas

### 2.1 Função da onda harmónica

Muitos processos ondulatórios podem ser descritos com funções periódicas das quais as funções harmónicas, *seno* e *co-seno*, são mais simples. Mesmo se o perfil de uma onda real (na superfície do mar, por exemplo) não coincide exactamente com a forma destas funções, nós sabemos que qualquer função periódica pode ser representada como uma série de Fourier – uma soma dos *senos* e *co-senos*. E, portanto, faz todo o sentido estudar, antes de mais, as propriedades fundamentais das ondas harmónicas.

Como apurámos na secção anterior, as variáveis do espaço e do tempo  $x$  e  $t$  entram na função da onda numa combinação  $x - vt$  ou  $x + vt$ , em que  $v$  é a velocidade de propagação da onda e os sinais  $-$  ou  $+$  correspondem às ondas a propagar-se ao longo do eixo  $x$  quer no sentido positivo (sinal  $-$ ), quer no sentido negativo (sinal  $+$ ). As ondas harmónicas não são excepção desta regra. Podemos escrever, então, a função correspondente a uma onda harmónica como

$$\Phi(x, t) = A \cdot \sin(k(x - vt))$$

em que  $A$  caracteriza a amplitude de variação da grandeza física  $\Phi$  e, portanto,  $A$  mede-se nas mesmas unidades que  $\Phi$ . Ainda não sabemos o que é o parâmetro  $k$ , mas precisamos dele para que a expressão sob o *seno* seja adimensional. Para isto,  $k$  tem que ter unidades  $m^{-1}$ . Para apurar o significado físico deste parâmetro, lembremo-nos que a função da onda é periódica no espaço e que a medida da periodicidade no espaço é o comprimento de onda  $\lambda$ :

$$\Phi(x + \lambda, t) = \Phi(x, t).$$

No caso da onda harmónica isto implica que

$$\sin(k(x + \lambda) - kvt) = \sin(kx - kvt)$$

$$k(x + \lambda) - kvt = kx - kvt + 2\pi n$$

$$k\lambda = 2\pi n$$

onde  $n$  é um número inteiro. Quando falamos da periodicidade (seja no tempo ou no espaço) interessa-nos o intervalo mínimo em que a função se repete. Portanto, temos para  $k$

$$k = \frac{2\pi}{\lambda}.$$

O parâmetro  $k$  assim definido chama-se *número de onda*.

Para a periodicidade no tempo podemos escrever uma equação semelhante

$$\Phi(x, t + T) = \Phi(x, t)$$

$$\sin(kx - kv(t + T)) = \sin(kx - kv t)$$

$$kx - kv(t + T) = kx - kv t + 2\pi$$

$$kvT = 2\pi$$

$$kv = \omega$$

Na última passagem aproveitámos a relação entre o período e a frequência  $T = \frac{2\pi}{\omega}$ , já apresentada no capítulo anterior sobre as oscilações. Daqui segue uma equação importante para a velocidade de propagação de uma onda

$$v = \frac{\omega}{k}$$

ou, ainda,

$$v = f\lambda,$$

o que já era de esperar (substituímos aqui  $\omega$  e  $k$  por  $\omega = 2\pi f$  e  $k = \frac{2\pi}{\lambda}$ , respectivamente).

Utilizando as relações entre  $\omega$ ,  $k$  e  $\lambda$  e ainda  $f = 1/T$ , podemos escrever a função de onda sob várias formas

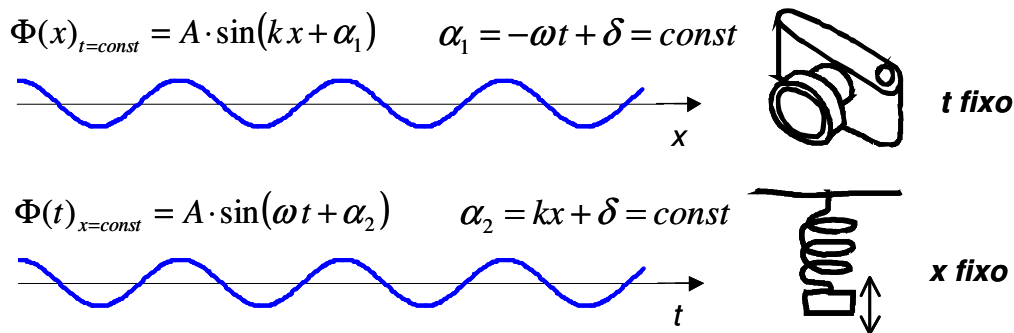
$$\Phi(x, t) = A \cdot \sin(k(x - vt) + \delta)$$

$$\Phi(x, t) = A \cdot \sin(kx - \omega t + \delta)$$

$$\Phi(x, t) = A \cdot \sin\left[2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right) + \delta\right],$$

sendo a segunda mais comum e compacta. A primeira forma tem a vantagem de explicitar a velocidade da onda como coeficiente do tempo (desde que a combinação  $x \pm vt$  esteja posta em evidência). Na terceira forma, os parâmetros que caracterizam a periodicidade no espaço e no tempo estão explicitados. Nestas equações, introduzimos também a fase inicial  $\delta$  que depende da escolha dos referenciais do tempo e no espaço (em  $t = 0$  e  $x = 0$  temos  $\Phi(x, t) = A \cdot \sin \delta$ ).

A apresentação gráfica de uma função de duas variáveis é difícil. É mais fácil representar a onda num gráfico unidimensional fixando um dos argumentos, i.e. para  $t$  fixo, ou para  $x$  fixo (ver as figuras). O primeiro gráfico é como se fosse uma fotografia instantânea da onda (uma fotografia das onda do mar é um bom exemplo). No segundo caso, estamos a observar a variação temporal da onda num dado ponto do espaço (o movimento de uma bóia no mar, por exemplo). Para uma onda harmónica este movimento é nada mais nada menos que as oscilações harmónicas.



Para as ondas harmónicas, podemos especificar melhor o que se entende sobre a velocidade de uma onda. Já mencionámos que as ondas não transportam a massa. O movimento de translação, associado aos fenómenos ondulatórios, é o movimento do estado da matéria: de uma crista da onda, por exemplo. No caso de uma onda harmónica,  $\Phi(x, t) = A \cdot \sin(kx - \omega t + \delta)$ , a posição de uma crista caracteriza-se por  $kx_{crista} - \omega t + \delta = \frac{\pi}{2}$  em que o seno é igual a 1. Resolvendo esta equação em ordem do

$\omega = 2\pi f$ $k = \frac{2\pi}{\lambda}$ $v = f\lambda$ $v = \frac{\omega}{k}$ $vT = \lambda$	<p><b>Várias formas da função da onda harmónica</b></p> $\Phi(x, t) = A \cdot \sin(kx - \omega t + \delta)$ $\Phi(x, t) = A \cdot \sin\left[2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right) + \delta\right]$ $\Phi(x, t) = A \cdot \sin(k(x - vt) + \delta)$
---	---

$x_{crista}$ , temos  $kx_{crista}(t) = \frac{\pi}{2} + \omega t - \delta$ , ou  $x_{crista}(t) = x_0 + \frac{\omega}{k}t$ , em que  $x_0 = \frac{\pi/2 - \delta}{k}$  é uma constante (é a posição da crista da onda em  $t = 0$ ). Isto significa que a crista se desloca a uma velocidade igual a  $\frac{\omega}{k}$ . E é precisamente isso que nós chamamos velocidade da onda. Voltando à função da onda  $\Phi(x, t) = A \cdot \sin(kx - \omega t + \delta)$ , reparamos que em todos estes pontos  $x_{crista}(t)$  a fase da onda, i.e. o argumento do *seno*, é constante. Por outras palavras, podemos dizer que a velocidade  $v = \frac{\omega}{k}$ , de que estamos a falar, corresponde à velocidade com que se propaga no espaço uma dada fase da onda ( $\pi/2$ , neste exemplo). Daqui vem o termo *velocidade de fase*.

Para as ondas não harmónicas, também podemos aplicar o termo de velocidade de fase, embora a expressão  $v = \frac{\omega}{k}$  já não seja válida porque a frequência angular, definida como  $\omega = 2\pi f = \frac{2\pi}{T}$ , só se aplica às ondas harmónicas (consequência do período das funções harmónicas *seno* e *co-seno* ser  $2\pi$ ; é por isso que temos que utilizar  $\sin\left(2\pi \frac{t}{T}\right) = \sin(\omega t)$  e não  $\sin\left(\frac{t}{T}\right)$  para descrever uma onda com o período igual a  $T$ ). No entanto, a velocidade de fase expressa através da frequência  $f = \frac{1}{T}$  como  $v = \frac{2\pi f}{k} = f\lambda$  pode ser utilizada para qualquer onda periódica.

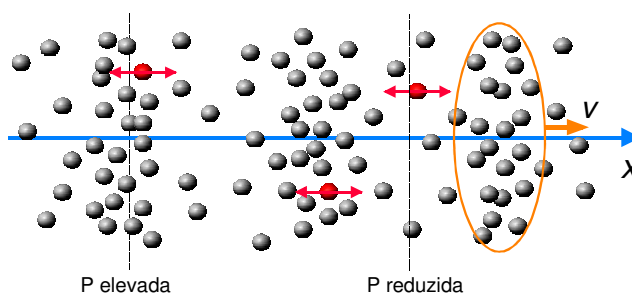
## **2.2 Exemplos das ondas harmónicas**

O significado físico da função  $\Phi(x, t)$ , como já foi dito, depende da natureza da onda em causa. Se fixarmos uma extremidade de uma corda numa parede e abanarmos a outra extremidade podemos observar uma onda unidimensional a evoluir ao longo da corda. A função da onda, neste caso,  $\Phi(x, t) = z(x, t)$ , onde  $z(x, t)$  corresponde à coordenada dos pontos da corda ao longo do eixo  $z$  perpendicular à corda (estendida ao longo do  $x$ ) e orientada de tal maneira que as oscilações ocorrem no plano  $xOz$ .

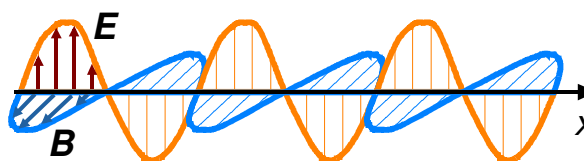
Para as ondas do mar a função da onda descreve também a altura da onda, embora o processo seja mais complexo porque ocorre em duas dimensões. A função da onda depende de duas coordenadas  $(x, y)$  no plano da superfície da água  $\Phi(\mathbf{r}, t) = z(x, y, t)$ .

Se existir uma simetria cilíndrica, como no caso da onda excitada por uma pedra, as coordenadas  $(x, y)$  podem ser substituídas por uma única coordenada radial  $r$ .

Outro exemplo de ondas que enfrentamos todos os dias é o das ondas sonoras. As ondas sonoras são a propagação de uma perturbação local que resulta na compressão ou descompressão do meio. A alteração local da pressão do ar afecta as zonas subjacentes, estas afectarão as zonas vizinhas, etc., e assim a onda sonora propaga-se. A membrana de um altifalante a vibrar obriga as moléculas do ar em contacto com ela a acompanhar o seu movimento, quer empurrando-os para frente, quer puxando-os para trás e criando, assim, ora compressão, ora descompressão local do ar. Deste modo, as ondas sonoras podem ser descritas, quer em termos da variação da pressão do ar relativamente à pressão de equilíbrio, quer em termos de deslocamento das moléculas do ar relativamente à posição média. No primeiro caso,  $\Phi(x,t) = \Delta p(x,t)$  e no segundo -  $\Phi(x,t) = \xi(x,t)$ . É importante notar que as moléculas do ar oscilam ao longo do eixo  $x$  de modo que  $x$  em  $\xi(x,t)$  é a posição média de uma dada molécula e  $\xi$  é a coordenada instantânea desta molécula medida em relação à sua posição média.



No caso das ondas electromagnéticas, a função da onda descreve a variação espaço-temporal da intensidade do campo eléctrico



$E(x,t)$  ou do campo magnético  $B(x,t)$ . Os vectores do campo eléctrico e do campo magnético de uma onda electromagnética são perpendiculares à direcção de propagação da onda e, também, perpendiculares entre si como se mostra na figura.

Em geral, essas ondas não são necessariamente ondas harmónicas mas, como já foi dito, em todo o caso podem ser representadas como uma soma (finita ou infinita) das ondas harmónicas com as frequências diferentes. Um som puro (emitido por um diapasão, por exemplo) é uma onda harmónica porque tem uma só frequência (por vezes chama-se uma onda harmónica simples). Uma nota musical emitida por um instrumento ou cantada é composta por um conjunto de ondas harmónicas com frequências

diferentes. O espectro de frequências é um espectro discreto composto por uma frequência fundamental e frequências mais altas, múltiplas da frequência fundamental - as harmónicas (ver secção 4.2). É por isso que estas ondas por vezes se chamam ondas harmónicas compostas. Quando falamos emitimos som com um espectro de frequências ainda mais rico, praticamente contínuo.

## **2.3 Polarização**

Há uma diferença importante entre, por exemplo, as ondas numa corda e as ondas sonoras. Numa corda, cada ponto oscila numa direcção perpendicular à direcção em que a onda se propaga. Estes dois eixos formam um plano que pode ser orientado de muitas formas. Se a corda estiver na horizontal, as oscilações podem ocorrer na vertical ou na horizontal, ou num plano inclinado, mas sempre perpendicularmente à corda. As ondas deste tipo chamam-se *ondas transversais*. Se o plano em que ocorrem as oscilações se mantém constante, diz-se, neste caso, que a onda está *polarizada* neste plano. As ondas electromagnéticas também são ondas transversais porque os vectores do campo eléctrico e do campo magnético são perpendiculares ao vector da velocidade.

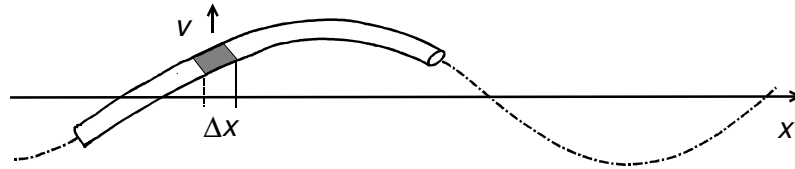
Em contraste com estas ondas, a onda sonora é uma onda *longitudinal* porque as moléculas do meio oscilam na mesma direcção em que a onda se propaga. O plano de polarização não está definido para as ondas longitudinais e, portanto, o conceito de polarização não se aplica.

Uma onda do mar é um caso mais complexo. A onda é aparentemente transversal mas o único plano de polarização possível é o plano vertical. Isto é uma consequência do facto de o nosso espaço ter três dimensões. A onda ocupa o plano  $(x, y)$  e a única dimensão “disponível” para uma onda transversal é a dimensão ao longo do eixo  $z$ . De facto, uma análise mais detalhada mostra que uma onda na superfície da água é uma combinação de uma onda transversal com uma onda longitudinal; daqui resulta que as moléculas da água perto da superfície se movem ao longo de uma trajectória elíptica. É assim que se move, por exemplo, um pedaço de algas suspenso na água. Por isso se dá por vezes às ondas superficiais o nome de ondas elípticas.

## **2.4 Energia da onda**

Consideremos uma onda harmónica numa corda vibrante (ver a figura). Como já sabemos, um elemento da corda  $\Delta x$  efectua uma oscilação harmónica simples: para  $x$

fixo  $y(x,t) = y(t) = A \cdot \sin(\omega t + \alpha)$ , em que o termo  $kx$  está incluído na constante  $\alpha$ . Se a massa por unidade de comprimento da corda for  $\mu$ , a energia associada ao movimento de um elemento  $\Delta x$  (com a massa igual a  $\mu \Delta x$ ) pode ser calculada através da energia



cinética máxima (ver capítulo sobre oscilações):  $E = \frac{1}{2} \mu \Delta x v_{\max}^2$ . Repare, o  $v$  aqui não

é a velocidade da onda mas sim a velocidade com que se move o elemento da corda na

direcção perpendicular à corda. Como  $v = \frac{dy}{dt} = A \omega \cdot \cos(\omega t + \alpha)$ , temos para a

velocidade máxima  $v_{\max} = A \omega$ . Então,  $E = \frac{1}{2} \mu \Delta x (A \omega)^2$  e para a energia por unidade

de comprimento  $\varepsilon = \frac{E}{\Delta x}$  temos

$$\varepsilon = \frac{1}{2} \mu A^2 \omega^2.$$

Para outros tipos de ondas, é óbvio que vão aparecer na equação outros parâmetros específicos para um dado tipo, mas as características gerais, tais como a amplitude e a frequência, mantêm-se. Uma conclusão importante, então, é que a energia transportada por uma onda é proporcional ao quadrado da amplitude e ao quadrado da frequência

$$\varepsilon \propto A^2 \omega^2.$$

A energia da onda

$$\varepsilon \propto A^2 \omega^2$$

Isto é válido para qualquer onda clássica.

### 3. A equação da onda

Nesta secção vamos derivar a equação geral da onda, utilizando um exemplo concreto de uma onda a propagar-se ao longo de uma corda esticada, tal como uma corda de um instrumento musical. O caminho é semelhante ao que seguimos para compor a equação do movimento oscilatório: identificar as forças que afectam o movimento de um elemento da corda e aplicar a 2ª lei de Newton ao sistema.

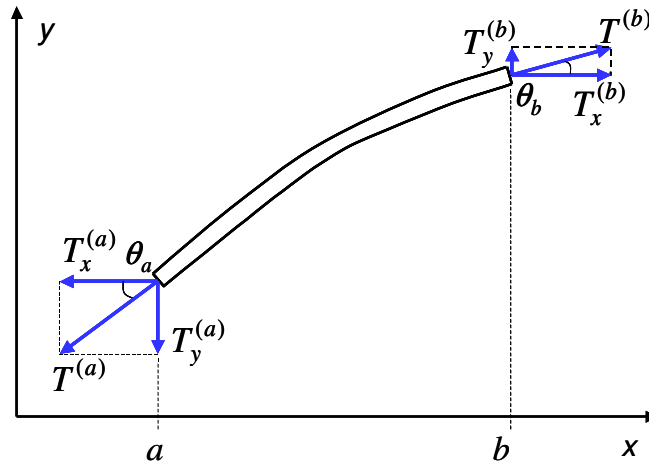


Vamos centrar o nosso interesse na fase da vibração estabilizada deixando de fora a fase inicial, em que as oscilações se iniciam por uma excitação exterior. Vamos supor também que no sistema não existe atrito ou outro mecanismo de dissipação da energia das vibrações. E, finalmente, supomos que a corda é infinita ignorando assim os efeitos dos bordos.

O desenho ao lado representa um pequeno fragmento da corda esticada ao longo do eixo  $x$ , que vibra no plano  $xy$ . A única força relevante neste sistema, a que faz com que ao picar a corda num ponto a excitação se propague ao longo dela, é a força da tensão da corda. Esta força é responsável pela integridade da corda e pela distribuição do excesso local da tensão às partes adjacentes da corda, que ocorre com uma velocidade finita – a velocidade de propagação da onda.

É importante desde já marcar claramente o nosso objectivo do ponto de vista matemático. O nosso objectivo formal é compor a equação para a função da onda transversal (a corda vibra ao longo dos  $yy$ ) que se propaga na corda ao longo do eixo  $x$ , i.e. a equação para uma função  $y(x,t)$ .

Representamos a força de tensão, que actua ao longo da corda, como soma de dois vectores  $\mathbf{T}^{(a)} = \mathbf{T}_x^{(a)} + \mathbf{T}_y^{(a)}$ , na extremidade (a), e  $\mathbf{T}^{(b)} = \mathbf{T}_x^{(b)} + \mathbf{T}_y^{(b)}$ , na extremidade (b), tal como se mostra na figura. Como não existe movimento ao longo do



eixo  $x$ , podemos concluir que  $\mathbf{T}_x^{(a)} = -\mathbf{T}_x^{(b)}$  (caso contrário existiria a aceleração ao longo do eixo  $x$  e, como consequência, movimento nesta direcção). Designaremos as normas das ambas as forças por  $\|\mathbf{T}_x^{(a)}\| = \|\mathbf{T}_x^{(b)}\| = T_x$ . Sabemos que os componentes  $y$  das forças não são iguais nem simétricos. O componente da força resultante ao longo do eixo  $Oy$  é  $F_y = T_y^{(a)} - T_y^{(b)}$ . Os componentes  $T_y^{(a)}$  e  $T_y^{(b)}$  podem ser expressos através da  $F_x$  e dos ângulos  $\theta_A$  e  $\theta_B$ :

$$\begin{cases} T_y^{(a)} = T_x \cdot \tan \theta_a \\ T_y^{(b)} = T_x \cdot \tan \theta_b \end{cases}$$

e para  $F_y$  vem, então,

$$F_y = T_x \cdot \tan \theta_a - T_x \cdot \tan \theta_b.$$

Lembre-mo-nos que o nosso objectivo é chegar a uma equação para  $y(x,t)$  e, portanto, temos que exprimir todas as grandezas, que entram na equação, em termos da função  $y$  e de variáveis  $x$  e  $t$ . Tendo em vista este objectivo, substituímos os tangentes pelas derivadas

$$\tan \theta_A = \left. \frac{\partial y}{\partial x} \right|_a \quad \text{e} \quad \tan \theta_B = \left. \frac{\partial y}{\partial x} \right|_b$$

(utilizamos aqui as derivadas parciais porque  $y$  é uma função de duas variáveis,  $x$  e  $t$ ).

Assim,

$$F_y = T_x \left( \left. \frac{\partial y}{\partial x} \right|_a - \left. \frac{\partial y}{\partial x} \right|_b \right).$$

A componente da força  $F_y$  produz um movimento acelerado (cada ponto da corda oscila ao longo do eixo  $y$ ) e, de acordo com a segunda lei de Newton,  $F_y = ma_y$ . Massa  $m$  aqui é a massa daquele pequeno elemento da corda entre os pontos  $x = a$  e  $x = b$ . A equação final não deve depender do comprimento do elemento da corda que estamos a considerar  $\Delta x = b - a$ . Mas a massa não pode desaparecer de todo porque esta caracteriza a inércia do sistema – um elemento essencial para que haja oscilações. O mais provável nestas circunstâncias é que vai aparecer na equação a massa por unidade de comprimento da corda  $\mu$ . Então, escrevendo  $m = \mu \Delta x$  e lembrando que a aceleração

é a segunda derivada da coordenada  $a_y = \frac{\partial^2 y}{\partial t^2}$  temos

$$\mu \Delta x \cdot \frac{\partial^2 y}{\partial t^2} = T_x \left( \left. \frac{\partial y}{\partial x} \right|_a - \left. \frac{\partial y}{\partial x} \right|_b \right)$$

ou ainda

$$\mu \cdot \frac{\partial^2 y}{\partial t^2} = T_x \frac{\left. \frac{\partial y}{\partial x} \right|_a - \left. \frac{\partial y}{\partial x} \right|_b}{\Delta x}.$$

No limite  $\Delta x \rightarrow 0$  teremos do lado direito da equação a segunda derivada do  $y$  em ordem a  $x$ :

$$\lim_{\Delta x \rightarrow 0} \frac{\left. \frac{\partial y}{\partial x} \right|_a - \left. \frac{\partial y}{\partial x} \right|_b}{\Delta x} = \frac{\partial^2 y}{\partial x^2}.$$

Vem finalmente

$$\frac{\partial^2 y}{\partial t^2} = v^2 \frac{\partial^2 y}{\partial x^2},$$

onde  $v = \sqrt{\frac{T}{\mu}}$  é a velocidade de onda. Utilizamos aqui o facto de que para amplitudes

pequenas  $T_x \approx T$  (porque  $\tan \theta \rightarrow 0$  quando  $\theta \rightarrow 0$ ): a corda de uma guitarra com comprimento de cerca de 70 cm oscila com amplitudes não superiores a  $\approx 1$  mm. Porque

é que concluímos que  $\sqrt{\frac{T}{\mu}}$  é a velocidade da onda? Por duas razões. Em primeiro lugar

porque tem unidades m/s ( $\sqrt{\frac{N}{kg/m}} = \sqrt{\frac{kg \cdot m}{s^2} \cdot \frac{m}{kg}} = \sqrt{\frac{m^2}{s^2}} = \frac{m}{s}$ ) e em segundo lugar,

verifica-se que a função da onda harmónica só satisfaz a equação se o coeficiente em

frente da  $\frac{\partial^2 y}{\partial x^2}$  for igual a  $\frac{\omega}{k}$ , que é a velocidade da onda harmónica (ver caixa).

Para generalizar a equação obtida falta substituir a função da onda, que neste exemplo é o desvio da posição de equilíbrio em direcção  $y$ , por uma função de onda geral  $\Phi(x, t)$ .

A função da onda harmónica não é a única solução da equação da onda. De facto, pode-se demonstrar que qualquer função de dois argumentos  $x$  e  $t$  satisfaz a equação da onda desde que estes se combinem numa expressão  $x \pm vt$ . Outra observação importante é que a velocidade, evidenciada na equação da onda, depende da natureza

física das ondas ( $\sqrt{\frac{T}{\mu}}$  surgiu só porque consideramos um exemplo concreto de onda

numa corda). Por exemplo, para as ondas electromagnéticas em vácuo o que tem que aparecer na equação é a velocidade da luz  $c^2$ .

**A equação geral da onda:**

$$\frac{\partial^2 \Phi}{\partial t^2} = v^2 \frac{\partial^2 \Phi}{\partial x^2}$$

**Uma onda harmónica satisfaz a equação da onda.**

A função que descreve uma onda harmónica é  $\Phi(x, t) = A \sin(kx - \omega t)$

$$\frac{\partial \Phi}{\partial t} = -\omega A \cos(kx - \omega t)$$

$$\frac{\partial \Phi}{\partial x} = k A \cos(kx - \omega t)$$

$$\frac{\partial^2 \Phi}{\partial t^2} = -\omega^2 A \sin(kx - \omega t)$$

$$\frac{\partial^2 \Phi}{\partial x^2} = -k^2 A \sin(kx - \omega t)$$

Substituímos as segundas derivadas na equação da onda e temos

$$-\omega^2 A \sin(kx - \omega t) = v^2 \cdot (-k^2 A \sin(kx - \omega t)) \Rightarrow \omega^2 = v^2 k^2 \Rightarrow v = \frac{\omega}{k},$$

que é uma tautologia (igualdade) para uma onda harmónica. Isto significa que uma onda harmónica é a solução da equação da onda (uma das soluções, de facto).

**Uma função  $\Phi(x, t) = \Phi(x \pm vt)$  satisfaz a equação da onda.**

Para provar isto, temos que nos lembrar da regra de derivação de uma função composta:  $\frac{\partial}{\partial x} \Phi[u(x, t)] = \frac{d\Phi}{du} \cdot \frac{\partial u}{\partial x}$ . No nosso caso,  $u = x \pm vt$ . Vamos escolher

apenas o sinal menos para facilitar. Então,  $\frac{\partial u}{\partial x} = 1$  e  $\frac{\partial u}{\partial t} = -v$ .

$$\frac{\partial \Phi}{\partial t} = \frac{d\Phi}{du} \cdot \frac{\partial u}{\partial t} = -v \frac{d\Phi}{du}$$

$$\frac{\partial^2 \Phi}{\partial t^2} = \frac{d}{du} \left( -v \frac{d\Phi}{du} \right) \cdot \frac{\partial u}{\partial t} = -v \frac{d^2 \Phi}{du^2} \cdot (-v) = v^2 \frac{d^2 \Phi}{du^2}$$

$$\frac{\partial \Phi}{\partial x} = \frac{d\Phi}{du} \cdot \frac{\partial u}{\partial x} = \frac{d\Phi}{du}$$

$$\frac{\partial^2 \Phi}{\partial x^2} = \frac{d}{du} \left( \frac{d\Phi}{du} \right) \cdot \frac{\partial u}{\partial x} = \frac{d^2 \Phi}{du^2} \cdot 1 = \frac{d^2 \Phi}{du^2}.$$

Substituímos as segundas derivadas na equação da onda e temos

$$\left( v^2 \frac{d^2 \Phi}{du^2} \right) = v^2 \left( \frac{d^2 \Phi}{du^2} \right), \text{ que é uma tautologia (igualdade) para qualquer que seja}$$

a função  $\Phi(x, t) = \Phi(x \pm vt)$ .

## 4. Sobreposição das ondas harmónicas.

### 4.1 Interferência construtiva e destrutiva:

*Sobreposição de duas ondas planas com fases diferentes que se propagam no mesmo sentido.*

Imaginemos duas ondas harmónicas com amplitudes e frequências iguais a propagarem-se no mesmo sentido com a mesma velocidade, mas que têm as fases diferentes

$$\Phi_1(x, t) = A \cdot \sin(kx - \omega t + \delta_1)$$

$$\Phi_2(x, t) = A \cdot \sin(kx - \omega t + \delta_2)$$

(o facto de as frequências e as velocidades serem iguais implica que os comprimentos de onda também são iguais porque  $v = f\lambda$ ). Qual será o resultado da sobreposição destas ondas? Formalmente temos

$$\Phi_s(x, t) = \Phi_1(x, t) + \Phi_2(x, t)$$

$$\Phi_s(x, t) = A \cdot \sin(kx - \omega t + \delta_1) + A \cdot \sin(kx - \omega t + \delta_2)$$


$$\Phi_s(x, t) = 2A \cdot \sin\left(kx - \omega t + \frac{\delta_1 + \delta_2}{2}\right) \cos\left(\frac{\delta_1 - \delta_2}{2}\right).$$

Reescrevemos esta última fórmula da seguinte maneira

$$\Phi_s(x, t) = A_s \cdot \sin\left(kx - \omega t + \frac{\delta_1 + \delta_2}{2}\right)$$

$$\text{onde } A_s = 2A \cdot \cos\left(\frac{\delta_1 - \delta_2}{2}\right).$$

Concluimos, então, que o resultado de sobreposição dessas duas ondas harmónicas também é uma onda harmónica, cuja frequência  $\omega$  e velocidade  $v = \frac{\omega}{k}$  (e, então, o comprimento da onda  $\lambda = \frac{2\pi}{k}$ , também) são iguais às das ondas originais, mas a amplitude e a fase são diferentes. O facto mais notável é que a amplitude da onda

  $\sin a + \sin b = 2 \cdot \sin\left(\frac{a+b}{2}\right) \cos\left(\frac{a-b}{2}\right)$

resultante depende da diferença de fases das ondas a sobrepor-se e, em função desta, pode variar entre 0 e o dobro das amplitudes de cada uma das ondas originais. Esta propriedade é específica para o mundo das ondas e não tem analogia no mundo das partículas. Ao adicionar uma bola de bilhar a uma outra bola, teríamos sempre duas bolas, mas nunca 1,5 ou mesmo nenhuma! Parece que as ondas, ao contrário das bolas, interferem uma com a outra e até podem aniquilar-se. Daqui surge o termo *interferência*. Nos casos extremos temos

para  $\delta_1 - \delta_2 = 2\pi n \Rightarrow A_s = 2A$  - diz-se que existe *interferência construtiva*, e

para  $\delta_1 - \delta_2 = \pi + 2\pi n \Rightarrow A_s = 0$  - temos *interferência destrutiva*.

Na prática, raramente encontramos ondas planas. Mais vezes observamos a interferência das ondas emitidas por fontes pontuais (ou os que podem ser consideradas como tais), i.e. ondas esféricas ou circulares. Nestes casos, o padrão de interferências é mais complexo. Se a sobreposição das ondas planas resulta em anulação (ou ampliação) das ondas em todo o espaço, para ondas esféricas isto acontece só em certas regiões do espaço. No entanto, a razão física para isto é a mesma: se as duas ondas chegam a um dado ponto com as fases iguais (i.e.  $(kx - \omega t + \delta)_1 = (kx - \omega t + \delta)_2 + 2\pi n$ ), há interferência construtiva neste ponto; se as fases são opostos (i.e.  $(kx - \omega t + \delta)_1 = (kx - \omega t + \delta)_2 + \pi + 2\pi n$ ), as ondas anulam-se. Vamos discutir essas situações na secção 7 deste capítulo.

## **4.2 Onda estacionária:**

*Sobreposição de duas ondas com a mesma frequência que se propagam em sentidos opostos.*

Um caso de grande importância é a sobreposição de duas ondas com as mesmas características, mas que se propagam em sentidos opostos. O sentido de propagação da onda reflecte-se no sinal com que se combinam  $kx$  e  $\omega t$  no argumento do *seno*:  $kx - \omega t$  descreve uma onda a propagar-se no sentido  $x$  positivo enquanto  $kx + \omega t$  corresponde a uma onda que se propaga no sentido  $x$  negativo (compare com  $x - vt$  e  $x + vt$ , onde  $v = \frac{\omega}{k}$ ). Portanto, podemos escrever as funções dessas duas ondas como

$$\Phi_1(x, t) = A \cdot \sin(kx - \omega t + \delta_1)$$

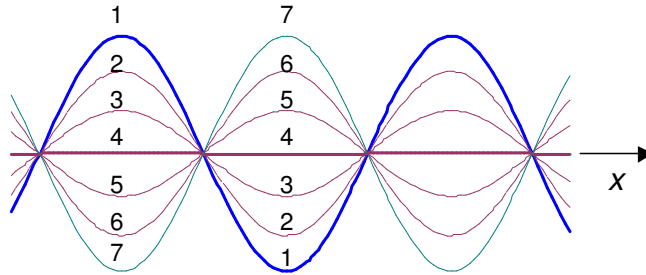
$$\Phi_2(x, t) = A \cdot \sin(kx + \omega t + \delta_2).$$

Somando  $\Phi_s(x, t) = \Phi_1(x, t) + \Phi_2(x, t)$ , temos

$$\Phi_s(x, t) = 2A \cdot \sin\left(kx + \frac{\delta_1 + \delta_2}{2}\right) \cos\left(\omega t - \frac{\delta_1 - \delta_2}{2}\right).$$

Ao olhar bem para esta função reparamos com espanto, que uma das propriedades inerentes da onda, que destacamos sempre – a combinação das coordenadas espaciais e do tempo numa expressão  $x \pm vt$  – já não está explicitamente presente em  $\Phi_s(x, t)$ . As variáveis  $x$  e  $t$  separaram-se e isto resulta num efeito novo: em certos pontos do eixo  $x$ , tais que  $kx + \frac{\delta_1 + \delta_2}{2} = \pi n$  ( $n$  é um número inteiro), a função da onda  $\Phi_s(x, t)$  é igual a zero em qualquer instante de tempo.

A existência dos pontos  $x$  em que a função da onda é sempre igual a zero é um efeito mais notório da sobreposição das ondas a viajar em sentidos opostos. Estes pontos chamam-se *nodos*. O que é que acontece nos pontos intermédios em



que  $kx + \frac{\delta_1 + \delta_2}{2} = \frac{\pi}{2} + \pi n$ ? Nesses pontos a função da onda torna-se igual a

$\Phi_s(t) = \pm 2A \cdot \cos\left(\omega t - \frac{\delta_1 - \delta_2}{2}\right)$  ou, seja, varia no tempo entre  $-2A$  e  $+2A$ . A figura

seguinte mostra o perfil da onda  $\Phi_s(x, t)$  em vários instantes de tempo cuja sequência está indicada junto às curvas.

A onda, que a figura mostra, chama-se onda *estacionária*. Em contraste com uma onda *progressiva*, tal como uma onda harmónica simples, por exemplo, que se propaga no espaço, a onda estacionária não se propaga – os nodos estão sempre no mesmo sítio e os ventres, ora aparecem, ora desaparecem, mas não saem do espaço entre os nodos. Repare, que para uma onda harmónica progressiva esta situação é impossível – mesmo que num dado instante de tempo, num dado ponto do espaço a função da onda fosse

igual a zero, já no instante seguinte não seria (porque  $A \cdot \sin(kx - \omega t + \delta) = 0$  quando  $kx - \omega t + \delta = 2\pi n$  ou, seja,  $kx = 2\pi n + \omega t - \delta$ ; isto significa que  $x$ , em que a função da onda é igual a zero, varia com o tempo).

Pode-se sugerir, então, que se for assim, a velocidade da onda estacionária deve ser igual a zero. Note-se que a velocidade de uma onda é a velocidade com que se propaga a fase da onda – uma crista, por exemplo, ou um vale, ou um ponto em que a função da onda é nula etc. – i.e. velocidade de fase (veja-se a primeira secção deste capítulo). Com base neste entendimento temos que concluir que a velocidade da onda estacionária é, realmente, nula. O mesmo também se verifica formalmente. Como os argumentos  $x$  e  $t$  de qualquer função de onda se combinam em  $x \pm vt$  e na função da onda estacionária o  $x$  e o  $t$  se encontram separados, a única maneira de satisfazer a primeira condição é assumir que  $v = 0$  (assim, o  $x$  já pode aparecer sozinho no argumento da função da onda).

Vamos explorar um pouco mais as ondas estacionárias. Em que situações é que estas podem surgir? A situação mais comum é quando uma onda progressiva se reflecte de um obstáculo e se sobrepõe a si própria. Por exemplo, uma onda sonora ao reflectir-se numa parede vai viajar em sentido contrário e sobrepor-se à onda que continua a ser emitida. Como ambas as ondas são essencialmente a mesma onda e diferem apenas no sentido de propagação, temos as condições para formação de uma onda estacionária. Outro exemplo é uma corda de um instrumento musical fixa em duas extremidades. “Fixa em duas extremidades” significa que a corda não pode vibrar nos pontos de fixação nem para além deles, i.e. a onda, gerada na corda de uma ou outra maneira, estará contida entre estes dois pontos. Por exemplo, quando se pica a corda de uma guitarra, a excitação propaga-se ao longo da corda em ambos os sentidos até encontrar um ponto de fixação onde se reflecte. As ondas reflectidas em cada extremidade da corda viajam ao encontro de uma outra e sobrepõem-se formando uma onda estacionária.

Consideremos uma corda de comprimento  $L$  com uma extremidade em  $x = 0$  e a outra em  $x = L$ . A corda não vibra nos pontos de fixação, o que significa que a onda estacionária que se forma na corda deve ter nodos nestes pontos, i.e.

$$kx + \frac{\delta_1 + \delta_2}{2} = \pi n \quad \text{em } x = 0 \text{ e em } x = L \quad (n \text{ é um número inteiro}).$$



Em  $x = 0$  temos  $\frac{\delta_1 + \delta_2}{2} = \pi n$  e para  $x = L$  temos  $kL + \frac{\delta_1 + \delta_2}{2} = \pi n'$  (em princípio,  $n$  e  $n'$  são números inteiros diferentes). Ao combinarmos as duas equações obtemos  $kL + \pi n = \pi n' \Rightarrow kL = \pi(n' - n) \Rightarrow kL = \pi m$  onde  $m$  também é um número inteiro.

Ao substituírmos o número de onda  $k$  pela sua definição  $k = \frac{2\pi}{\lambda}$  chegamos, finalmente, à seguinte equação

$$\lambda_m = \frac{2L}{m} \text{ com } m = 1, 2, 3, \dots$$

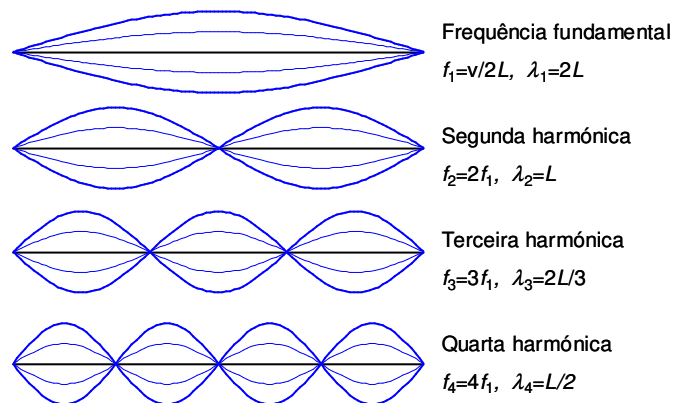
Esta equação mostra que as ondas estacionárias numa corda só podem existir com certos comprimentos de onda  $\lambda_1 = 2L$ ,  $\lambda_2 = L$ ,  $\lambda_3 = \frac{2}{3}L$  etc. (ver a figura). Em termos da frequência este facto exprime-se como

$$f_m = \frac{v}{2L} m$$

(utilizamos aqui  $f = \frac{v}{\lambda}$ ):  $f_1 = \frac{v}{2L}$ ,  $f_2 = \frac{v}{2L} 2 = 2f_1$ ,  $f_3 = \frac{v}{2L} 3 = 3f_1$  etc. Isto quer dizer que *a)* existe numa frequência mínima com que uma corda pode vibrar,  $f_1$  (esta frequência chama-se *frequência fundamental*) e *b)* as outras frequências permitidas são os múltiplos da frequência fundamental e chamam-se *harmónicas* da respectiva ordem

$$f_m = m f_1 \quad (m = 2, 3, 4, \dots).$$

Este é o resultado da aplicação das condições de fronteira à função da onda estacionária, e é de facto a base física de funcionamento dos instrumentos musicais não só de cordas mas também de percussão, os sopros e todos os outros instrumentos acústicos.



### **Lá de um violino e de uma flauta: iguais mas diferentes**

As notas idênticas de todos os instrumentos musicais afinam-se de forma a terem a mesma frequência fundamental. Esta é a razão física da consonância (aquilo que soa agradavelmente ao ouvido) dos Lá-s dos instrumentos diferentes. No entanto, um ouvido experiente é capaz de distinguir os diferentes instrumentos mesmo se estes tocarem a mesma nota. Onde está, então, a diferença? A diferença está no facto de ao produzir-se um certo som, além da frequência fundamental excitam-se também as harmónicas de ordem superior. A contribuição relativa de diferentes harmónicas para o som emitido pelo instrumento – o *espectro de harmónicas* – varia de um instrumento musical para outro. Esta é a explicação física para o facto de diferentes instrumentos terem *timbres* diferentes.

Outra diferença substancial entre o som produzido por diferentes instrumentos musicais está na fase, em que a onda sonora está a ser formada depois de uma excitação inicial – o ataque. Nesta fase, além da frequência fundamental e das harmónicas, também estão presentes frequências altas que não são múltiplos da frequência fundamental. Isto resulta no som inconfundível de um instrumento.

(<http://pt.wikipedia.org/w/index.php?title=Timbre&oldid=1014587>)

## **4.3 Batimentos:**

*Sobreposição de duas ondas planas que se propagam no mesmo sentido com frequências ligeiramente diferentes.*

Outro caso que tem interesse considerar é quando se sobrepõem duas ondas, que se propagam no mesmo sentido, com mesma velocidade e amplitudes iguais, mas que têm frequências um pouco diferentes. Alguns instrumentos musicais têm várias cordas para cada nota, que, em princípio, devem ser afinadas para a mesma frequência (por exemplo, uma guitarra com cordas duplas ou um piano, que tem duas ou três cordas para cada nota). No entanto, quando as cordas correspondentes à mesma nota estão ligeiramente desafinadas, produz-se um efeito sonoro específico, chamado efeito de *batimentos*, em que a amplitude da onda resultante não é constante, mas oscila com uma frequência baixa. Os batimentos permitem a um ouvido experiente detectar pequenas diferenças entre as frequências. Vamos ver como se descreve formalmente este efeito.

O essencial aqui é que as duas ondas têm frequências quase iguais. Podemos representar, então, as funções dessas ondas  $\Phi_1(x,t) = A \cdot \sin[k_1 x - \omega_1 t + \delta_1]$  e  $\Phi_2(x,t) = A \cdot \sin[k_2 x - \omega_2 t + \delta_2]$  na seguinte forma

$$\Phi_1(x,t) = A \cdot \sin[(k + \Delta k)x - (\omega + \Delta\omega)t + \delta_1]$$

$$\Phi_2(x,t) = A \cdot \sin[(k - \Delta k)x - (\omega - \Delta\omega)t + \delta_2]$$

Aqui  $\omega$  corresponde à frequência média  $\omega = \frac{\omega_1 + \omega_2}{2}$  e  $\Delta\omega = \frac{\omega_1 - \omega_2}{2}$ . Ao supor que as frequências são “quase iguais” constatamos que  $\Delta\omega \ll \omega$  e que  $\omega_1 \approx \omega_2 \approx \omega$ . O mesmo é válido para o número de onda  $k$ . Repare, uma vez que as velocidades das ondas são iguais (no caso de um instrumento musical é a velocidade do som no ar), para  $\omega_1 = \omega + \Delta\omega$  tem-se necessariamente  $k_1 = k + \Delta k$  (e não  $k - \Delta k$ ) porque a frequência e o número de onda estão ligadas à velocidade através da equação  $v = \frac{\omega}{k}$ , i.e. ao maior  $\omega$  corresponde maior  $k$ .

Somando e utilizando a fórmula para a soma de dois *senos* temos

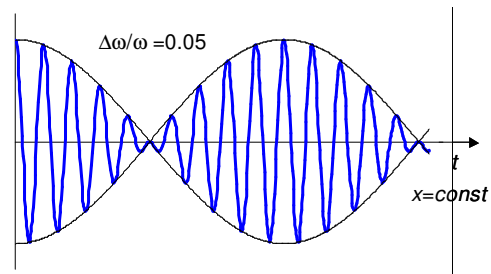
$$\Phi_s(x, t) = 2A \cdot \sin\left[kx - \omega t + \frac{\delta_1 + \delta_2}{2}\right] \cos\left[\Delta kx - \Delta\omega t + \frac{\delta_1 - \delta_2}{2}\right]$$

que podemos também escrever da seguinte maneira:

$$\Phi_s(x, t) = A_s(t) \cdot \sin\left[kx - \omega t + \frac{\delta_1 + \delta_2}{2}\right] \text{ em que}$$

$$A_s(t) = 2A \cdot \cos\left[\Delta kx - \Delta\omega t + \frac{\delta_1 - \delta_2}{2}\right]$$

O resultado da sobreposição de duas ondas com frequências ligeiramente diferentes é uma onda com frequência igual à média dessas frequências e com amplitude que varia no tempo, com uma frequência muito mais baixa do que as frequências de cada uma das ondas  $\Delta\omega \ll \omega$  - os batimentos. A figura ao lado mostra um exemplo de batimentos.



Olhamos mais uma vez para a função  $\Phi_s(x, t)$  e perguntamos: qual é que é a velocidade desta onda? Ao localizar a combinação  $kx - \omega t = k\left(x - \frac{\omega}{k}t\right)$  e comparando-o com  $x - vt$ , que qualquer função da onda deve ter, facilmente identificamos  $v = \frac{\omega}{k}$ , o que não constitui surpresa nenhuma – é a velocidade de fase, como já sabíamos. Mas a verdadeira surpresa está no facto de o  $x$  e o  $t$  se encontrarem, também, no argumento do *co-seno* em combinação  $\Delta kx - \Delta\omega t = \Delta k\left(x - \frac{\Delta\omega}{\Delta k}t\right)$  e, seguindo a mesma lógica, podemos identificar uma outra velocidade nesta expressão,

$u = \frac{\Delta\omega}{\Delta k}$  (designamos esta velocidade com um símbolo diferente para a distinguir da velocidade de fase  $v$ ). Parece que existem duas velocidades associadas a uma onda destas...

Formalmente, a interpretação deste facto é simples: a velocidade  $v = \frac{\omega}{k}$  está relacionada com a onda harmónica de frequência  $\omega$  (linha azul da figura) enquanto  $u = \frac{\Delta\omega}{\Delta k}$ , pelos vistos, corresponde à velocidade da onda de frequência baixa  $\Delta\omega$ , ou, seja, daquela onda que modula a amplitude da onda de frequência base  $\omega$  (linhas pretas da figura). A velocidade  $u$  chama-se *velocidade de grupo* porque caracteriza a velocidade com que se propaga um agrupamento (um pacote) de ondas como aqueles que se vêem na figura.

Para descobrir a importância física da velocidade de grupo, temos que responder, primeiro, a duas questões fundamentais: *a)* o que é uma onda real e em que é que difere de uma onda harmónica; e *b)* como é que uma onda transmite a informação. Vamos voltar a estas questões um pouco mais adiante, mas antes disso verificaremos se existe alguma relação entre as duas velocidades. Para isso, substituímos a razão  $\frac{\Delta\omega}{\Delta k}$  pela derivada  $\frac{d\omega}{dk}$ . Podemos fazer isto porque desde início assumimos que  $\Delta\omega$  e  $\Delta k$  são grandezas pequenas.

Nas transformações que se seguem, as grandezas  $\omega$ ,  $k$  e  $v$  consideram-se todas variáveis (é uma suposição mais geral do que assumir algumas delas constantes; no fim do cálculo veremos se valeu a pena fazê-lo). Lembremo-nos ainda da definição da derivada como limite da razão das variações pequenas (veja-se a caixa na secção sobre trabalho e energia). Isto permite-nos considerar a derivada  $\frac{d\omega}{dk}$  como sendo a razão de  $d\omega$  e  $dk$ .

Como o nosso objectivo é encontrar a relação entre  $u$  e  $v$ , vamos tentar explicitar  $v$  em  $u = \frac{d\omega}{dk}$ :

$$\omega = vk \quad \Rightarrow \quad u = \frac{d\omega}{dk} = \frac{d(vk)}{dk} = v + k \frac{dv}{dk}.$$

Multiplicando o último termo por  $1 = \frac{d\lambda}{d\lambda}$  temos

$$u = v + k \cdot \frac{dv}{dk} \cdot \frac{d\lambda}{d\lambda} = v + k \cdot \frac{dv}{d\lambda} \cdot \frac{d\lambda}{dk}.$$

Como  $\lambda = \frac{2\pi}{k}$ ,  $\frac{d\lambda}{dk} = -\frac{2\pi}{k^2}$  e, então

$$u = v + k \cdot \frac{dv}{d\lambda} \cdot \left( -\frac{2\pi}{k^2} \right) = v - \frac{2\pi}{k} \cdot \frac{dv}{d\lambda}$$


E, finalmente, para a velocidade de grupo vem

$$u = v - \lambda \frac{dv}{d\lambda}.$$

Daqui se vê que quando a velocidade da onda num meio não varia com o comprimento de onda, a velocidade de grupo é igual à velocidade de fase:  $u = v$ . Nos meios nos quais a velocidade de onda depende do comprimento de onda (ou da frequência – pode-se exprimir a mesma relação em termos de frequência), a velocidade de grupo é diferente da velocidade de fase. Os meios em que  $v = v(\lambda)$  chamam-se *meios dispersivos*. Um exemplo bem conhecido é a dispersão da luz branca num prisma. Este efeito é devido precisamente ao facto da velocidade de propagação da luz no vidro ser diferente para diferentes comprimentos de onda.

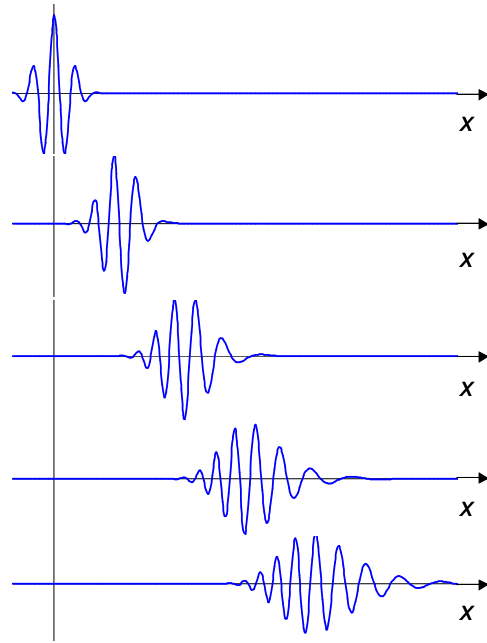
Voltaremos à questão do significado físico da velocidade de grupo. Em primeiro lugar, notamos que duas velocidades de onda apareceram somente porque somamos duas ondas com frequências diferentes. Cada uma dessas duas ondas é uma onda monocromática que, em geral, se propaga no meio com a sua própria velocidade. (Monocromática significa que a frequência  $\omega$  é definida com exactidão.) A sobreposição destas, no entanto, já não é uma onda monocromática como é, aliás, qualquer onda real. As ondas reais não são monocromáticas por duas razões. Primeiro porque na prática não existem emissores que emitam numa frequência exacta. A emissão ocorre sempre numa banda de frequências, que pode ser muito estreita, mas sempre finita (mesmo o laser emite a luz com comprimentos de onda dentro de um intervalo  $(\lambda - \Delta\lambda, \lambda + \Delta\lambda)$ ). A segunda razão é que uma onda real, ao contrário de uma onda harmónica, é finita no tempo – tem sempre início e fim. Estas partes da onda real não podem ser descritas com um simples *seno* ou *co-seno* com uma só frequência. Lembremo-nos, por exemplo, da fase inicial do som emitido por um instrumento musical (o ataque) que é sempre mais rico em frequências do que o som contínuo.

Concluimos, então, que uma onda real não é uma onda harmónica. E ainda bem! Porque uma onda harmónica é pouco útil do ponto de vista de transmissão de informação. Um som contínuo, de frequência e amplitude invariáveis, que continua desde sempre e nunca acaba, não transmite informação nenhuma. Mas se variarmos pelo menos um dos parâmetros da onda, seja a frequência ou amplitude, ou ainda interrompermos a onda por uns instantes e começarmos de novo, isso, sim, permite-nos transmitir a informação. No último caso, a informação pode ser transmitida, por exemplo, em forma de código de Morse em que a cada letra do alfabeto e a cada número está atribuída uma sequência de sinais de duração diferente: sinais curtos e sinais longos. Assim, três sinais curtos seguidos por três sinais longos e mais outros três impulsos curtos,  $\bullet\bullet\bullet---\bullet\bullet\bullet$ , codificam o sinal S.O.S. As ondas electromagnéticas modulados em amplitude ou em frequência permitem transmitir sons com sinais de rádio (costuma-se designar esses dois modos de transmissão por AM e FM, respectivamente, do Inglês *amplitude modulation* e *frequency modulation*).

Em comunicações digitais, a informação transmite-se em forma de sequências de uns e zeros, em que 1 é representado como presença do sinal e 0 como ausência. Um código de 8 bits 10110010 representa-se, então, com a seguinte sequência de sinais rectangulares . Um impulso rectangular é composto por muitas ondas harmónicas e contém uma larga gama de frequências, desde  $f = 1/2T$  ( $T$  é a duração de um bit) e até as frequências muito elevadas, até  $f = \infty$ , em princípio, para um impulso ideal com tempo de subida infinitamente pequeno (para perceber por que precisamos das frequências altas basta compararmos a parte da frente de um impulso rectangular com a subida de um *seno*). Na caixa em baixo mostra-se como se representa um impulso rectangular por uma série de funções harmónicas. A representação exacta requer um número infinito de ondas. Com algumas dezenas consegue-se uma aproximação razoável.

A velocidade de propagação das ondas com frequências (comprimentos de onda) diferentes num meio é, em geral, diferente. Isto leva a que umas partes do impulso se propagam mais rápido do que outras e, como resultado, a forma do sinal varia ao longo da distância. O exemplo ao lado mostra a propagação no espaço de um impulso composto por 100 ondas harmónicas de frequências diferentes, num meio dispersivo em que a velocidade de fase varia com o número de onda  $k$  como  $v(k) = \frac{1}{\sqrt{1 + \alpha k}}$ , i.e. as

ondas com frequências maiores são mais lentas ( $\alpha > 0$  é uma constante). Vê-se na figura que à medida que o impulso se propaga alarga-se e fica distorcido. As ondas com frequências baixas ultrapassam as que têm uma frequência alta (compare-se a parte da frente e a de trás do impulso no fim de curso). A velocidade com que se propaga o impulso como um todo – a velocidade do grupo de ondas – é, obviamente, diferente da velocidade de fase de cada uma das ondas harmónicas que o constituem. E como a unidade de informação, um bit, é o impulso todo e não cada onda harmónica em separado (essas



ondas não transmitem a informação por serem ondas monocromáticas – não esqueçamos!), é a velocidade do impulso que caracteriza a rapidez com que a informação é transmitida. Eis uma conclusão importante: a informação transmite-se por ondas com velocidade igual à velocidade de grupo.

Outra observação relevante é que quando transmitimos a informação por ondas temos que ter em conta a degradação da forma do impulso à medida que este se propaga no espaço. Ao emitir dois impulsos (dois bits de informação) separados por um intervalo de tempo demasiado curto arriscamo-nos a não poder distingui-los a partir de certa distância por causa do alargamento dos impulsos. O efeito do alargamento é tanto maior quanto maior for a gama das frequências das ondas harmónicas que constituem o sinal. Este efeito, em conjunto com a absorção de ondas no meio, limita o número de bits que podem ser transmitidos por segundo. (O coeficiente de absorção das ondas num meio também é uma função da frequência. As frequências altas, normalmente, são mais absorvidas do que as frequências baixas. Isto também faz com que a forma do impulso varie ao longo da distância.)

### Um impulso quadrado como soma de ondas harmónicas.

De acordo com a teoria de Fourier, uma função periódica  $f(x)$  com período igual a  $L$  pode ser representada como uma soma infinita (ou finita em alguns casos) de funções harmónicas, i.e.

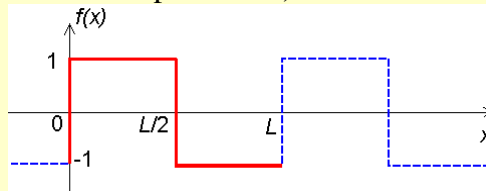
$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(2\pi n \frac{x}{L}\right) + \sum_{n=1}^{\infty} b_n \sin\left(2\pi n \frac{x}{L}\right)$$

em que os coeficientes  $a_n$  e  $b_n$  podem ser determinados a partir das equações

$$a_n = \frac{2}{L} \int_0^L f(x) \cos\left(2\pi n \frac{x}{L}\right) dx$$

$$b_n = \frac{2}{L} \int_0^L f(x) \sin\left(2\pi n \frac{x}{L}\right) dx$$

(aqui utilizamos  $x$  como uma variável genérica que pode significar uma coordenada espacial ou tempo, ou ainda outro parâmetro).



Um impulso quadrado em  $x \in [0, L]$  pode ser descrito com uma função

$$f(x) = 2 \left[ H\left(\frac{2x}{L}\right) - H\left(\frac{2x}{L} - 1\right) \right] - 1,$$

em que  $H(x)$  é a função de Heaviside definida como

$$H(x) = \begin{cases} 0, & \text{para } x < 0 \\ 1/2, & \text{para } x = 0 \\ 1, & \text{para } x > 0 \end{cases}$$

Como  $f(x) = f(L - x)$ , a função  $f(x)$  é ímpar e, então,  $a_0 = a_n = 0$ . Os coeficientes  $b_n$  calculados com o integral referido em cima são

$$b_n = \frac{4}{\pi n} \times \begin{cases} 0, & \text{para } n \text{ par} \\ 1, & \text{para } n \text{ ímpar} \end{cases}$$

Deste modo,

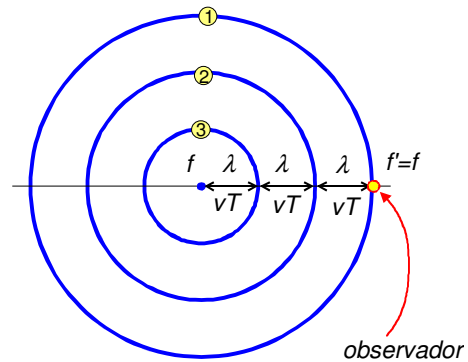
$$f(x) = \sum_{n=1,3,5,\dots}^{\infty} \frac{4}{\pi n} \sin\left(2\pi n \frac{x}{L}\right)$$

(verifique com Excel !).

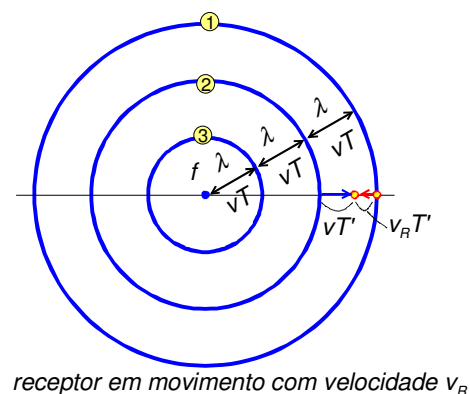


## 5. Efeito de Doppler

Imaginemos uma fonte a emitir ondas com uma dada frequência  $f$ . Qual será a frequência das ondas  $f'$  detectadas por um receptor distante? À partida, não se vê porque é que esta seria diferente da frequência emitida pelo emissor... No caso estático, quando o emissor e o receptor estão em repouso, as duas frequências são realmente iguais. É fácil verificar isto na figura ao lado que representa a posição das frentes da onda num dado instante de tempo. Seria mais fácil de analisar a situação se a onda em causa fosse uma sequência de *beep*'s curtos emitidos com um período  $T$ . Neste caso, as linhas circulares correspondiam ao lugar geométrico dos pontos nos quais se ouve um *beep* num dado instante. Os números ao lado das linhas indicam a ordem com que os *beep*'s foram emitidos. O espaçamento entre os círculos é, em termos temporais, igual a  $T$  e, em termos de espaço, é igual ao produto do  $T$  vezes a velocidade de propagação das ondas:  $vT$ , que é igual ao comprimento de onda  $\lambda$ , como sabemos. Para um observador imóvel, os *beep*'s chegarão com intervalos de tempo iguais a  $T$  entre eles e, então, a frequência medida por este observador será igual a  $1/T$ , i.e. será igual à frequência dos sinais emitidos pelo emissor.



Já não é assim se o observador estiver em movimento. A figura seguinte demonstra a situação em que o receptor se move ao encontro das ondas. O ponto de encontro com o *beep* seguinte será algures à esquerda da posição inicial do observador. A distância que a onda tem que percorrer até chegar ao receptor agora é menor e, portanto, o intervalo de tempo  $T'$  entre os *beep*'s que o observador



vai medir será também menor:  $T' < T$ . Vê-se na figura que a distância  $\lambda = vT$  entre as frentes da onda agora é percorrida parcialmente pela onda e, parcialmente, pelo receptor de modo que  $T$  e  $T'$  se relacionam através da seguinte equação  $vT = vT' + v_R T'$ , em

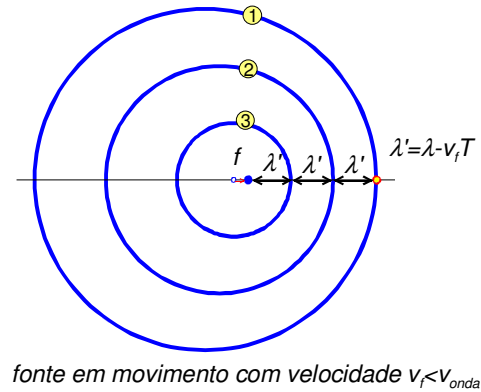
que  $v$  é a velocidade da onda e  $v_R$  é a velocidade do receptor. Daqui resulta

$T' = T \frac{v}{v + v_R}$  e para as frequências  $\frac{1}{f'} = \frac{1}{f} \cdot \frac{v}{v + v_R}$ . Finalmente,

$$f' = f \cdot \left( 1 + \frac{v_R}{v} \right).$$

A frequência detectada é, realmente, diferente da frequência emitida pela fonte e é tanto maior quanto maior a velocidade do receptor. Esta expressão foi obtida considerando o receptor a mover-se no sentido da fonte. Para descrever a situação em que o receptor se afasta do emissor das ondas, basta inverter o sinal da velocidade  $v_R$  na equação:  $f' = f \cdot \left( 1 - \frac{v_R}{v} \right)$ . Neste caso, a frequência detectada será inferior à emitida e se o observador se mover com velocidade igual à velocidade da onda,  $f'$  torna-se igual a 0 o que faz sentido. Para  $v_R > v$ , a equação não se aplica: as ondas nunca serão detectadas pelo observador.

Consideremos agora a situação em que o observador permanece no mesmo sítio mas em que a fonte das ondas se move na sua direcção com uma velocidade  $v_f$ . Como cada *beep* seguinte é emitido de um ponto mais próximo do receptor do que o *beep* anterior, a distância entre as frentes da onda emitida na direcção do observador  $\lambda'$  é



fonte em movimento com velocidade  $v_f < v_{\text{onda}}$

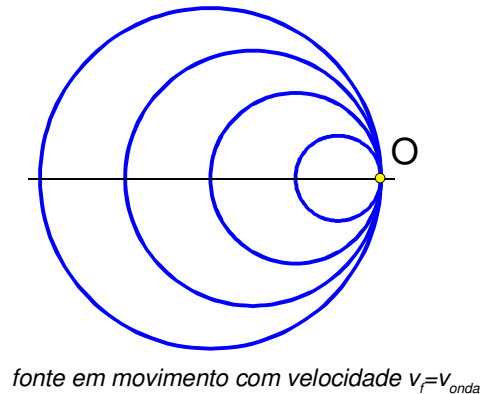
mais curta do que seria se o emissor estivesse em repouso (ver a figura):  $\lambda' = \lambda - v_f T$  (aqui  $v_f T$  corresponde ao deslocamento da fonte entre dois *beep*'s consecutivos). O intervalo de tempo  $T'$  entre os dois *beep*'s que mede o observador será igual a esta distância  $\lambda'$  dividida pela velocidade da onda:  $T' = \frac{\lambda - v_f T}{v}$ . Substituindo  $\lambda = vT$ ,

temos  $T' = T \frac{v - v_f}{v} = T \left( 1 - \frac{v_f}{v} \right)$  e para a frequência

$$f' = \frac{f}{1 - \frac{v_f}{v}}.$$

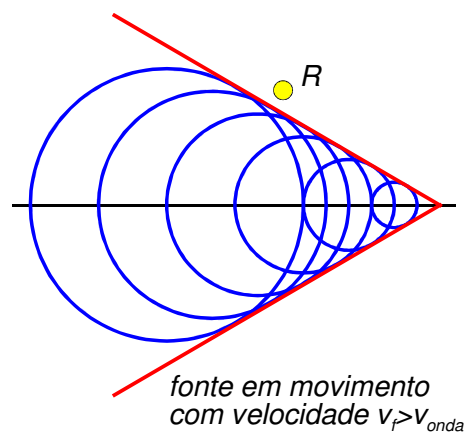
A situação que esta equação descreve é muito comum: é assim que varia a frequência do som da buzina de um carro que se aproxima, passa à nossa frente e depois se afasta. A variação da frequência é tanto maior quanto maior a velocidade do carro.

Quando a velocidade do emissor das ondas se aproxima da velocidade da onda  $v_f \rightarrow v$ , a frequência detectada tende para infinito. Esta situação corresponde ao que se mostra na figura ao lado: para  $v_f = v$ , todas as ondas emitidas pela fonte ao longo de todo o caminho chegam ao ponto  $O$  ao mesmo tempo, o que significa  $T' = 0$ . De facto, como  $v_f = v$



o emissor das ondas também deve estar no mesmo ponto  $O$ . O avião a voar à velocidade do som chegará ao observador no mesmo instante de tempo que o som emitido pelos seus motores... A acumulação das ondas sonoras numa região do espaço muito limitada faz com que se forme uma zona de alta compressão do ar em frente do avião a acelerar, que este tem que atravessar para ultrapassar a velocidade do som. Este fenómeno é conhecido como *barreira sonora*. Ao atravessar a barreira sonora o avião está sujeito a um grande stress mecânico e instabilidade que, obviamente, devem ser tomados em conta pelos construtores dos aparelhos supersónicos.

Com uma velocidade superior à da onda, o deslocamento da fonte durante um período é maior que a distância percorrida pela onda. A imagem das frentes da onda (dos *beep's*, no nosso exemplo) num dado instante de tempo demonstra-se na figura ao lado. Repare na superfície cónica formada pela sobreposição das frentes das ondas. Este cone representa uma *onda de choque* e dá origem a



um efeito específico conhecido sob o nome *sonic boom*, quando se trata de ondas sonoras. Um avião supersónico a sobrevoar um observador (ponto  $R$  na figura) com uma velocidade superior à do som, produz uma onda de choque que ao chegar ao observador resulta num som de estalo, por vezes parecido com uma explosão. O efeito

é, de certo modo, semelhante ao que se ouve quando se liga subitamente um som muito alto.

Ao combinar as duas equações para  $f'$ , podemos descrever a situação em que tanto a fonte como o receptor estão em movimento:

$$f' = f \cdot \frac{1 + v_R/v}{1 - v_f/v}.$$

Lembremo-nos que esta equação foi obtida considerando a fonte e o receptor a moverem-se um na direcção do outro. No caso quando um deles ou ambos se afastam, basta inverter o sinal da respectiva velocidade. Uma observação relevante que temos que fazer é que na derivação das equações acima só consideramos as situações quando o movimento ocorre ao longo da linha que junta o receptor e a fonte das ondas. No caso geral, para incluirmos as situações em que a direcção de um ou ambos vectores da velocidade não coincide com essa linha, temos que substituir  $v_R$  pelo produto  $v_R \cdot \cos \theta_R$  e  $v_f$  pelo  $v_f \cdot \cos \theta_f$ , onde os ângulos  $\theta_R$  e  $\theta_f$  são ângulos entre a recta, que junta a fonte com o receptor, e o vector da velocidade respectivo.

Há mais uma coisa que sobressai da última equação para  $f'$ . As velocidades  $v_R$  e  $v_f$  entram de uma maneira não simétrica. Isto pode parecer estranho dada a relatividade do movimento: não seria lógico esperar que só a velocidade relativa entre a fonte e o receptor aparecesse na equação? De facto, seria assim se a onda em causa não precisasse de um meio material para se propagar. No caso das ondas sonoras, que não se propagam no vazio, os efeitos devidos ao movimento do receptor e os devidos ao movimento da fonte são, na realidade, diferentes. O movimento do receptor não afecta as ondas a propagar-se no ar. A alteração da frequência é apenas uma consequência das condições em que esta está a ser medida. Por outro lado, quando o emissor se move, as ondas formadas no ar já não são iguais às emitidas pelo mesmo emissor mas imóvel: têm comprimento de onda diferente (no nosso exemplo,  $\lambda' < \lambda$ ). O importante aqui é a velocidade com que cada um se move em relação ao meio. Por essa razão as velocidades, quer do emissor quer do receptor, entram nas equações como  $v_f/v$  e  $v_R/v$ .

Já para a luz, que não precisa de um meio material para se propagar, as equações são diferentes. O efeito de Doppler para uma onda electromagnética descreve-se com a seguinte equação (o efeito de Doppler relativista, como se diz frequentemente)

$$f' = f \sqrt{\frac{1 + v/c}{1 - v/c}}.$$

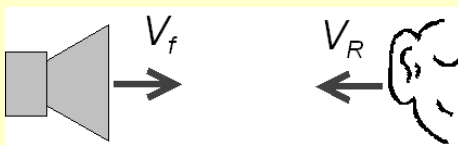
A velocidade  $v$  aqui é a velocidade do movimento relativo do emissor e do receptor e  $c$  é a velocidade da luz.

Como as frequências das ondas detectadas são uma função da velocidade do emissor ou/e receptor, o efeito de Doppler pode ser utilizado para as medir. É este o princípio de funcionamento de um radar que emite uma onda electromagnética de uma dada frequência e a compara com a frequência da onda reflectida por um objecto. O efeito de Doppler das ondas sonoras utiliza-se no diagnostico médico para medição da velocidade de fluxo sanguíneo. Na astronomia é conhecido um efeito chamado *desvio para o vermelho* (*red shift*) que consiste no comprimento de onda da luz dos objectos celestes detectado na Terra ser maior que o comprimento de onda da luz emitida por estes objectos. Do facto, quando existe um desvio para as frequências menores podemos concluir que as galáxias se afastam e, assim, concluir que o Universo está em expansão. Medindo a grandeza deste desvio em relação a um padrão (riscos atómicos – ver no capítulo sobre física quântica), podemos determinar as velocidades com que se afastam as galáxias.

#### Efeito de Doppler para onda sonora

$$f' = f \cdot \frac{1 + v_R/v}{1 - v_f/v}$$

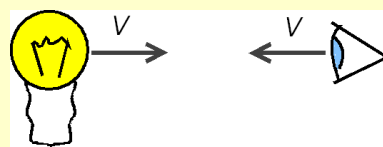
(corresponde à aproximação)



#### Efeito de Doppler para a luz

$$f' = f \sqrt{\frac{1 + v/c}{1 - v/c}}$$

(corresponde à aproximação)



## 6. Princípio de Huygens. Reflexão e Refracção das ondas.

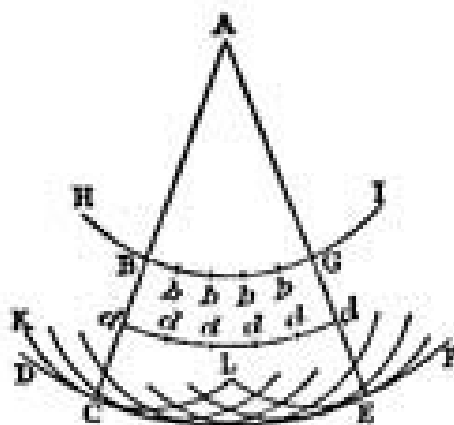
A questão que se coloca aqui é como é que a presença de obstáculos no espaço afecta a evolução de uma onda. Como podemos explicar, por exemplo, o padrão que produzem as ondas do mar ao contornar uma rocha?

Qualitativamente, a forma das ondas pode ser caracterizada como a forma geométrica da *frente da onda*. Frente da onda é o lugar geométrico de pontos no espaço em que a função de onda  $\Phi(x,t)$  tem o mesmo valor. Para uma onda harmónica, isto equivale a dizer que a frente da onda é o lugar geométrico de pontos com a mesma fase (i.e., o argumento do *seno* ou *co-seno*). O seguinte exemplo explica o conceito da frente da onda. Imaginemos que atiramos uma pedra à superfície de um lago. As ondas assim provocadas terão claramente a forma circular. Os círculos que se vêem na água são, de facto, as linhas que juntam às cristas das ondas. Diz-se, neste caso, que a onda tem uma frente circular. As ondas circulares iniciam-se no centro dos círculos e propagam-se em direcções radiais. Os círculos expandem-se no tempo mantendo a forma inalterável (se não houver obstáculos). Essas linhas, ao longo das quais as ondas se propagam, chamam-se *raios da onda* (na óptica, por exemplo, fala-se habitualmente em raios de luz). Os raios da onda são perpendiculares às frentes da onda.

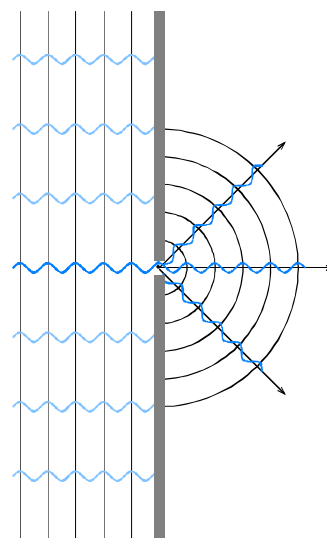
As ondas na superfície da água são ondas bi-dimensionais. No espaço de três dimensões, alguns exemplos de ondas com uma frente esférica são ondas de luz emitidas por uma lâmpada ou por uma estrela. As ondas de luz emitidas pelo Sol também têm simetria esférica. No entanto, devido à grande distância entre o Sol e a Terra, a curvatura da frente das ondas solares, quando estas chegam à Terra, é tão insignificante, que não se nota na grande maioria das experiências e observações à superfície da Terra (que têm, tipicamente, as dimensões da ordem de metros ou quilómetros, enquanto a distância da Terra ao Sol é de cerca de 150 milhões de quilómetros). Neste caso, as frentes da onda podem ser consideradas planas. A função da onda deixa de depender das três coordenadas espaciais ( $x, y, z$ ), como para uma onda esférica, e torna-se uma função de uma só coordenada  $x$  medida ao longo do sentido de propagação (i.e., ao longo do raio da onda) e do tempo. São essas ondas que estivemos a estudar nas secções anteriores.

Através das observações sabemos que uma onda circular conserva a sua forma a menos que encontre um obstáculo. Pergunta-se, se soubermos a forma de uma onda num certo instante de tempo, como podemos antecipar a sua forma no instante seguinte? Este problema foi resolvido em 1678 por um cientista holandês Christiaan Huygens que concluiu que cada partícula do meio, através do qual evolui a onda, transmite o movimento a todas as partículas à sua volta. Deste modo, cada partícula que está neste momento  $t$  na frente da onda pode ser considerada como uma fonte de ondas esféricas. A sobreposição de ondas esféricas emitidas por todas as partículas (em número infinito, em princípio) resulta numa onda que vamos observar no instante seguinte  $t + \Delta t$ .

A figura ao lado, atribuída ao próprio Huygens, ilustra esta ideia conhecida na física como *princípio de Huygens*. A curva  $HI$  representa a frente de uma onda esférica (circular, se for num plano), com o centro no ponto  $A$ , num certo instante. Para construir a frente da onda no instante seguinte, dividimos  $HI$  em pontos (pontos  $b$  na parte  $BG$  na figura) e consideramos cada ponto como emissor de ondas esféricas. Por exemplo, o arco  $KL$  representa a frente da onda emitida a partir do ponto  $B$ . A sobreposição destas ondas todas resulta numa esfera de raio  $AC$  cuja parte está representada pela curva  $DF$ .

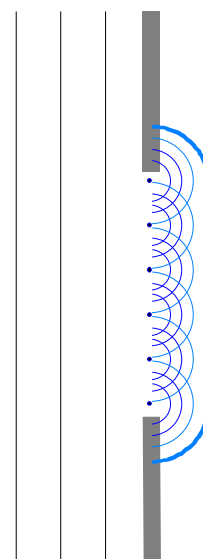


O princípio de Huygens permite prever o comportamento de uma onda quando esta encontra um obstáculo no seu caminho. A figura ao lado mostra a passagem de uma onda plana através de uma fenda estreita. De acordo com Huygens, a fenda pode ser considerada como uma fonte pontual de ondas secundárias, de modo que a onda atrás da fenda já não é plana, mas sim esférica. Para uma abertura mais larga, o que temos de fazer é dividi-la em pequenas partes e considerar cada uma como uma fonte pontual. A onda atrás do obstáculo resulta da sobreposição das ondas

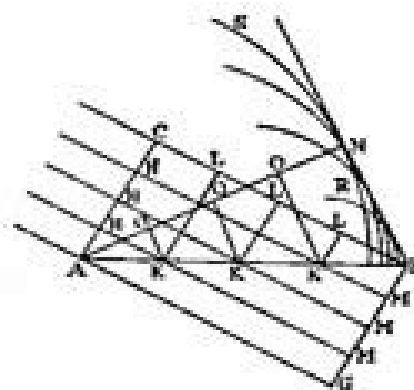


esféricas emitidas por todo o conjunto dessas fontes pontuais, como a figura a seguir mostra.

Huygens deduziu o seu princípio baseando-se apenas na construção geométrica, sem prova matemática rigorosa (e foi severamente criticado por isso). Nesta construção utilizou intuitivamente apenas uma parte da frente de cada onda – a parte que evolui para a frente (os semicírculos ou semiesferas). A prova matemática da validade do método veio mais tarde. Mostrou-se, em particular, que as partes das ondas emitidas para trás se anulam devido à sobreposição com a onda incidente (que continua a chegar, obviamente).



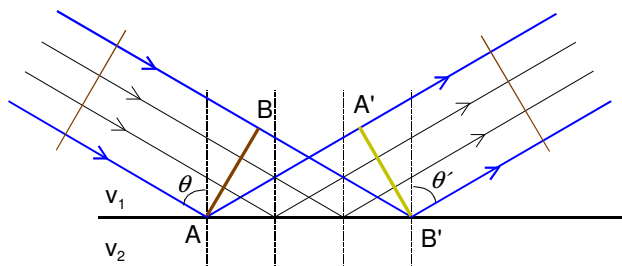
O princípio de Huygens também permite descrever os processos de reflexão e de refração das ondas numa superfície que separa dois meios. A figura representa uma onda plana com a frente AC (no instante  $t_0$ ) a incidir na superfície AB (desenho de Huygens). As diferentes partes da onda incidente chegam à superfície nos instantes diferentes. O primeiro raio chega ao ponto A em  $t = t_0$ ; o segundo chega ao ponto K em  $t = t_0 + \Delta t$ , etc. até  $t = t_0 + 4\Delta t$ , no desenho, em que a parte da frente, que no instante inicial estava no ponto C, chega ao ponto B. As linhas KL representam a evolução da frente da onda incidente em  $t_0 + \Delta t$ ,  $t_0 + 2\Delta t$  etc., enquanto que os arcos representam a posição das frentes das ondas esféricas emitidas nos pontos AKKKB no instante  $t_0 + 4\Delta t$ . A sobreposição destas ondas esféricas resulta numa onda plana. As linhas OK correspondem à evolução da frente da onda reflectida na mesma sequência de tempos. A frente da onda nos instantes intermédios  $t_0 + \Delta t$ ,  $t_0 + 2\Delta t$  e  $t_0 + 3\Delta t$  são OKL.



Nos desenhos seguintes, considera-se uma onda plana incidente sobre uma superfície de separação de dois meios, nos quais as velocidades de propagação das ondas são diferentes. Suponhamos que a velocidade da onda no meio 1 (parte de cima nos desenhos) é igual a  $v_1$  e no meio 2 (em baixo) é igual a  $v_2$ . O primeiro desenho representa a reflexão da onda incidente sobre a interface entre os dois meios.



Interessa-nos a relação entre os ângulos de incidência  $\theta$  e de reflexão  $\theta'$  (por convenção, os ângulos medem-se em relação à normal à superfície). As linhas  $AB$  e  $A'B'$  correspondem à



posição da frente da onda num instante  $t = t_0$  e algum tempo mais tarde, digamos  $t_0 + \Delta t$ . Ao analisarmos os triângulos  $\triangle ABB'$  e  $\triangle AA'B'$  reparamos que i) ambos os triângulos são triângulos rectângulos (a frente da onda é sempre perpendicular aos raios), ii) têm a hipotenusa  $AB'$  comum, e iii)  $AA' = BB'$  porque  $AA' = v_1 \Delta t$  e

$BB' = v_1 \Delta t$ , também. Portanto, os triângulos têm as hipotenusas e um dos dois catetos iguais. Pelo teorema de Pitágoras concluímos que os outros catetos também são iguais, ou seja  $\triangle ABB' = \triangle AA'B'$ . Daqui se conclui

O ângulo de reflexão é igual ao ângulo de incidência:  $\theta = \theta'$

que  $\angle A'AB' = \angle AB'B \Rightarrow \frac{\pi}{2} - \theta = \frac{\pi}{2} - \theta' \Rightarrow \theta = \theta'$ .

Repare, que o facto da onda evoluir sempre no mesmo meio, com a mesma velocidade, é essencial. A igualdade dos ângulos de incidência e de reflexão é uma consequência directa da igualdade das velocidades da onda antes e depois de ser reflectida. Caso contrário, as distâncias percorridas por diferentes partes da onda são diferentes e em consequência  $AA' \neq BB'$ . É precisamente isso que acontece na refacção das ondas quando as velocidades nos dois meios não são iguais. A figura seguinte representa esta situação. Considerando os triângulos rectângulos  $\triangle ABB'$  e  $\triangle AA'B'$ , concluímos que estes

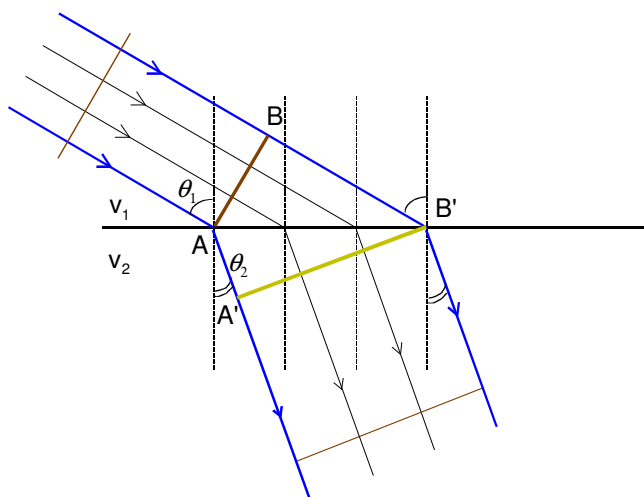
não são congruentes porque  $BB' = v_1 \Delta t$  mas  $AA' = v_2 \Delta t$ .

Podemos exprimir  $AA'$  e  $BB'$  também através dos ângulos  $\alpha$  e  $\beta$  e da hipotenusa  $AB'$  comum:

$$AA' = AB' \cdot \cos \beta \quad \text{e}$$

$$BB' = AB' \cdot \cos \alpha.$$

Então, a razão  $\frac{AA'}{BB'}$  é igual a  $\frac{\cos \beta}{\cos \alpha}$ , mas por



outro lado é igual a  $\frac{v_2}{v_1}$ . Daqui se conclui que  $\frac{\cos \beta}{\cos \alpha} = \frac{v_2}{v_1}$ . Como  $\beta = \frac{\pi}{2} - \theta_2$  e

$\alpha = \frac{\pi}{2} - \theta_1$ , temos

$$\frac{\sin \theta_2}{\sin \theta_1} = \frac{v_2}{v_1}.$$

Esta é a lei da refacção. O ângulo formado pelos raios da onda com a normal à superfície de separação dos dois meios é alterado quando a onda atravessa a superfície. A alteração é tanto maior quanto maior for a diferença entre as velocidades de propagação da onda nestes meios. No meio com uma velocidade menor, a onda aproxima-se da normal.

Os ângulos de refacção e de incidência relacionam-se através das velocidades da onda nos dois meios :

$$\frac{\sin \theta_2}{\sin \theta_1} = \frac{v_2}{v_1}$$

As leis de reflexão e de refacção acima derivadas aplicam-se às ondas de qualquer natureza, incluindo ondas de luz. As ondas de luz (como todas as ondas electromagnéticas) têm uma característica que as distingue das ondas de outro tipo, como bem sabemos: as ondas electromagnéticas não precisam de nenhum meio físico para se propagarem. A luz propaga-se mesmo no vácuo com uma velocidade finita<sup>1</sup>,  $c \approx 300\,000$  km/s. Nos meios, a velocidade da luz é menor. À razão da velocidade da luz no vácuo sobre a velocidade da luz num meio chama-se *índice de refacção* do meio  $n = \frac{c}{v}$ . Aplicado à luz, a lei de refacção é conhecido por *lei de Snell*

$$\frac{\sin \theta_2}{\sin \theta_1} = \frac{n_1}{n_2} \text{ ou, numa forma mais simétrica,}$$

$$n_1 \sin \theta_1 = n_2 \sin \theta_2.$$

Os índices de refacção dos materiais existentes variam entre 1 e 3,5. Os gases têm  $n$  muito próximos de 1, que é o índice de refacção para o vácuo. Por exemplo, para o ar em condições normais  $n = 1,0003$ ; para o vidro, tipicamente, entre 1,5 e 1,9. Como a velocidade da onda num meio depende, em geral, do comprimento de onda

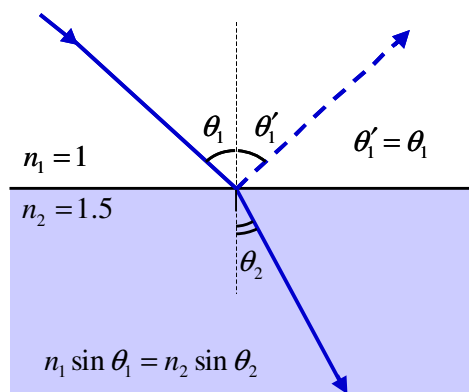
Definição do índice de refacção:  $n = \frac{c}{v}$

Lei de Snell  
 $n_1 \sin \theta_1 = n_2 \sin \theta_2$

<sup>1</sup> Curiosamente, já o Huygens no século XVII considerava que a luz se propagava com uma velocidade finita. Huygens introduziu a ideia sobre a natureza ondulatória da luz nos tempos quando dominava a teoria de Newton que considerava a luz como sendo um fluxo de partículas. Hoje sabemos que ambos tinham razão: todos os objectos da Natureza têm uma dupla identidade – são ondas-partículas.

(meios dispersivos – não esqueçamos!), o índice de refração da luz também tem esta dependência geral, aumentando para as ondas mais curtas. Isto resulta, por exemplo, na decomposição da luz branca em espectro de cores observada com um prisma de vidro.

Para nos lembrarmos mais facilmente da lei de Snell recomenda-se memorizar a imagem da trajetória de um raio de luz que entra do ar para o vidro: ao entrar no vidro o raio aproxima-se da normal.



Os processos de reflexão e de refração acompanham-se sempre um ao outro. Ao olhar para dentro de um armário com portas de vidro conseguimos ver os livros nas prateleiras, mas também vemos o reflexo da mobília à nossa volta. O brilho do reflexo depende do ângulo de incidência: o reflexo vê-se melhor para os ângulos rasantes. Em termos físicos é a questão da partilha da intensidade da luz incidente entre as duas componentes: componente reflectida e componente transmitida através da superfície (a intensidade da onda é proporcional à amplitude da onda ao quadrado). Se  $R$  for a fracção da intensidade da luz incidente que se reflecte, e  $T$  a fracção da intensidade da luz incidente que se transmite para dentro do material,  $R + T = 1$ . O mínimo de intensidade da luz reflectida atinge-se para a incidência normal (numa superfície ar-vidro, por exemplo, reflecte-se  $\approx 4\%$  da luz incidente). Com o aumento do ângulo de incidência, a fracção da luz incidente, que se reflecte, aumenta e chega aos 100% quando a luz incide a  $90^\circ$  (luz rasante).

Há duas coisas importantes a lembrar na aplicação das leis de reflexão e de refração:

1. os ângulos medem-se relativamente à normal à superfície de separação dos meios;
2. as leis em causa são locais, i.e. se existirem várias superfícies no caminho da luz, as leis aplicam-se a cada superfície, uma a uma, tendo em conta os pares de índices de refração dos meios, que são separados por cada superfície, e os ângulos em relação à normal a esta superfície.

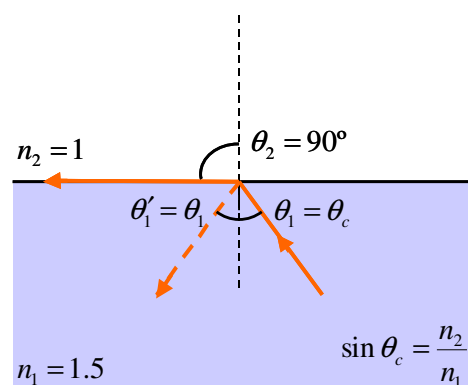
Mais uma observação: a trajectória da luz é reversível, i.e. se soubermos o caminho da luz num sentido, a luz mandada do ponto final no sentido contrário vai seguir ao longo dessa mesma linha.

Como já foi referido várias vezes, no meio com o índice de refacção maior os raios de luz aproximam-se à normal mais que no meio com  $n$  menor. Isto é uma consequência óbvia da lei de Snell  $n_1 \sin \theta_1 = n_2 \sin \theta_2$ : para  $n_1 > n_2$ ,  $\theta_1 < \theta_2$ . Portanto, numa situação em que a luz se aproxima da superfície do lado do meio com o índice maior, o ângulo  $\theta_2$  pode chegar aos  $\theta_2 = 90^\circ$  enquanto  $\theta_1 < 90^\circ$  (ver a figura). Ao ângulo  $\theta_1$ , em que isto acontece, chama-se

*ângulo crítico*. Colocando  $\theta_2 = 90^\circ$  na lei de Snell, chegamos à seguinte expressão para o ângulo crítico:  $\sin \theta_c = \frac{n_2}{n_1}$ , o que dá

$\sin \theta_c = \frac{1}{n_1}$ . No caso concreto dos índices

da figura, que corresponde à passagem da luz do vidro para o ar,  $\theta_c \approx 41,8^\circ$ .



A figura mostra também o raio reflectido (a tracejado) que, como já foi dito, existe sempre.

O que vai acontecer se aumentarmos  $\theta_1$  ainda mais, para além do ângulo crítico? Obviamente,  $\theta_2$  não pode aumentar para além de  $90^\circ$  (por definição, o raio refractado é aquele que é formado quando a luz passa de um meio para o outro). Matematicamente, para  $\theta_1 > \theta_c$  chegamos à seguinte contradição: como  $\sin \theta_1 > \sin \theta_c$

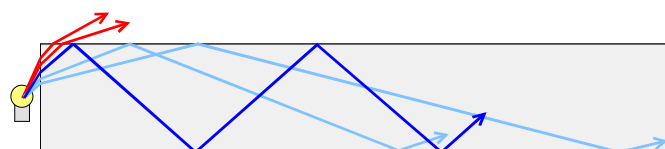
e  $\sin \theta_c = \frac{n_2}{n_1}$ , temos de acordo com a lei de Snell

$$\sin \theta_2 = \left( \frac{n_1}{n_2} \right) \cdot \sin \theta_1 > \left( \frac{n_1}{n_2} \right) \cdot \sin \theta_c = \left( \frac{n_1}{n_2} \right) \cdot \frac{n_1}{n_2} = 1$$

ou, seja,  $\sin \theta_2 > 1$ , o que não pode ser como é óbvio.

A interpretação física deste facto é que para os ângulos maiores que o ângulo crítico, o raio refractado não existe e só resta o raio reflectido. Como a reflexão é o único processo que existe nestas condições, toda a luz incidente é reflectida. Este fenómeno chama-se *reflexão total* (também, *reflexão interna*). A reflexão total tem

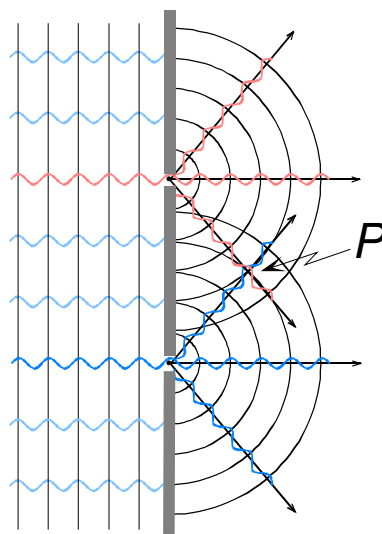
um papel muito importante para a tecnologia moderna. Assim funcionam algumas fibras ópticas. Um cabo de fibra óptica pode ter um comprimento de vários quilómetros mas o diâmetro de cada fibra é apenas da ordem de 0.1 mm. Está claro que a luz injectada de um lado da fibra sofre um número enorme de reflexões na superfície da fibra antes de chegar a outro lado da linha. Se assumirmos que em cada reflexão se perde-se 0.1% da luz, já depois de  $10^4$  reflexões (e pode haver muitíssimo mais) a intensidade do sinal luminoso constitui  $0.999^{10000} \sim 10^{-5}$  do sinal inicial. A existência da reflexão total muda tudo. Na reflexão total reflecte-se mesmo toda a luz, exactamente 100%, porque a reflexão é o único processo possível na fronteira entre os dois meios quando o ângulo da incidência é maior que o ângulo crítico.



## 7. Interferência e Difracção

### 7.1 Interferência em duas e três dimensões

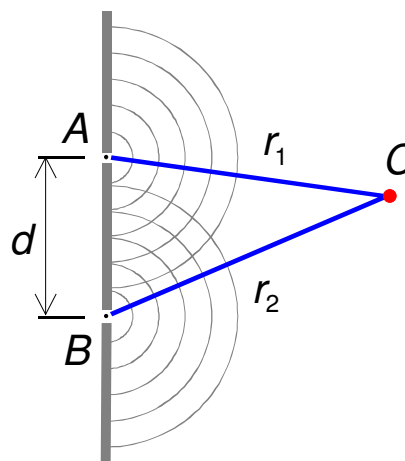
Já falámos sobre a interferência das ondas como resultado da sobreposição das ondas harmónicas planas. Tínhamos verificado que a soma de duas ondas harmónicas que se propagam no mesmo sentido não é igual à simples soma aritmética: com ondas,  $1 + 1$  nem sempre será igual a 2. Dependendo da diferença entre as fases das ondas, o resultado pode variar entre 0 e 2. Quando se somam as ondas com fases iguais, ou que diferem por um valor múltiplo de  $2\pi$ , a amplitude da onda resultante é o dobro da amplitude de cada onda (suponhamos aqui que as amplitudes das duas ondas a sobrepor-se são iguais). Mas se a diferença de fases for  $\pi + 2\pi n$  ( $n=0, \pm 1, \pm 2, \dots$ ) o resultado será catastrófico – as ondas aniquilam-se uma à outra. No primeiro caso



costuma-se dizer que temos interferência construtiva e, no segundo – interferência destrutiva (ver Secção 4.1 deste Capítulo).

A onda plana é uma onda unidimensional, descrita por uma função com dois argumentos  $\Phi(x,t)$ : uma coordenada espacial  $x$  e o tempo  $t$ . Interessa-nos agora a interferência das ondas em duas ou três dimensões: por exemplo, a interferência das duas ondas que se formam quando uma onda plana passa através de um diafragma com duas fendas estreitas, como se mostra na figura (suponhamos que o plano do diafragma é paralelo às frentes da onda incidente). O princípio de Huygens diz que as fendas podem ser consideradas como fontes pontuais das ondas secundárias, de modo que atrás do diafragma vamos ter duas ondas com frentes circulares. Estas duas fontes têm uma propriedade muito importante – oscilam com a mesma frequência e a mesma fase. Isto é porque ambas são induzidas pela mesma onda e, em qualquer instante de tempo, ambas as fontes estão na mesma frente da onda incidente.

De modo geral, o efeito de sobreposição das ondas num dado ponto do espaço é determinado pela diferença entre as fases das duas ondas com que estas chegam ao ponto, à semelhança da sobreposição das ondas planas. Por exemplo, é fácil verificar na figura acima que no ponto P, que pertence ao eixo de simetria entre as fendas, as duas ondas têm sempre as fases iguais e, portanto, as oscilações neste ponto efectuam-se com o dobro da amplitude em relação à amplitude que cada onda tem na sua origem (i.e. nos pontos A e B, respectivamente). Consideremos agora um ponto arbitrário C à direita do diafragma que fica à distância  $r_1$  da uma das fendas e  $r_2$  relativamente a outra (ver figura ao lado). As ondas emitidas a partir do ponto A propagam-se ao longo da linha AC e a função da onda pode ser escrita como  $\Phi_1(x,t) = A \cdot \sin(kx - \omega t)$  em que  $x$  é medido a partir do ponto A na direcção do ponto C. Aproveitamos aqui a simetria esférica da onda, que resulta em que basta apenas uma coordenada espacial – a coordenada radial – para caracterizar a onda em todo o espaço.



Para a onda emitida no ponto B, podemos escrever uma função semelhante  $\Phi_2(x',t) = A \cdot \sin(kx' - \omega t)$ , em que  $x'$  seria a coordenada medida ao longo da recta

BC a partir do ponto B. Nesta forma, a igualdade das fases das duas ondas nas suas origens está assegurada em qualquer instante de tempo. Realmente, como no ponto A  $x = 0$  e no ponto B  $x' = 0$ , temos ambas as fases (i.e. os argumentos dos *senos*) iguais a  $\varphi_1 = \varphi_2 = -\omega t$ . É importante de sublinhar que as funções  $\Phi_1(x, t)$  e  $\Phi_2(x', t)$  descrevem as duas ondas em todo o espaço à direita das fendas e em qualquer instante de tempo. No ponto C,  $x = r_1$  e  $x' = r_2$  e, portanto, as oscilações neste ponto descrevem-se com uma função  $\Phi_C(t) = \Phi_1(r_1, t) + \Phi_2(r_2, t)$ :

$$\Phi_C(t) = A \cdot \sin(kr_1 - \omega t) + A \cdot \sin(kr_2 - \omega t)$$

Recordando a regra de somar os senos, temos

$$\Phi_C(t) = 2A \cdot \sin\left(k \frac{r_1 + r_2}{2} - \omega t\right) \cdot \cos\left(k \frac{r_1 - r_2}{2}\right).$$

À semelhança da sobreposição de duas ondas harmónicas planas, reparamos que o argumento do *co-seno* não depende do tempo e, portanto, existem pontos no espaço em que o *co-seno* é sempre igual a 1 e que, assim, oscilam de acordo com a equação

$\Phi_C(t) = 2A \cdot \sin\left(k \frac{r_1 + r_2}{2} - \omega t\right)$ . Também existem pontos no espaço, para quais o *co-seno* é igual a 0 para qualquer  $t$ . Nestes pontos, não existem oscilações em nenhum instante de tempo:  $\Phi_C(t) = 0 \quad \forall t$ . No primeiro caso, temos interferência construtiva.

Esta acontece quando  $k \frac{r_1 - r_2}{2} = \pi n$  (com  $n = 0, \pm 1, \pm 2, \dots$ ). Substituindo  $k = \frac{2\pi}{\lambda}$

temos para estes pontos  $r_1 - r_2 = n\lambda$ . A

interferência destrutiva tem lugar em pontos para os

quais  $k \frac{r_1 - r_2}{2} = \frac{\pi}{2} + \pi n$  e, consequentemente,

$$r_1 - r_2 = \frac{\lambda}{2} + n\lambda.$$

A seguinte figura mostra o padrão de interferência em 2D (ondas na superfície de água) em que as bolas pretas são dois osciladores a oscilar com a mesma fase. Vêem-se na figura as linhas (linhas radiais cinzentas) ao longo das quais as duas ondas circulares se aniquilam uma à outra - zonas da interferência

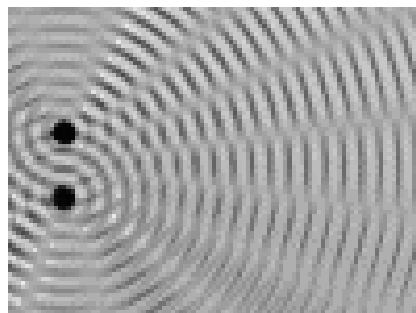
Interferência construtiva

$$r_1 - r_2 = n\lambda$$

Interferência destrutiva

$$r_1 - r_2 = \frac{\lambda}{2} + n\lambda$$

( $n = 0, \pm 1, \pm 2, \dots$ )

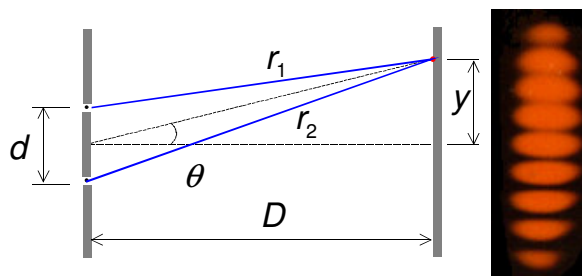


destrutiva. A água não oscila nestes pontos.

Um leitor atento deve ter reparado que ao somar as ondas assumimos silenciosamente que as amplitudes das ondas são constantes, não dependem da distância da fonte. Ora, isto é verdade para as ondas planas (ondas numa dimensão), mas para as ondas circulares ou esféricas, em duas ou três dimensões, não é assim. Primeiro, sabemos que as ondas na superfície de um lago induzidas por uma pedra diminuem em amplitude à medida que se afastam da sua origem. O mesmo deve acontecer com a amplitude da onda da luz (onda electromagnética) emitida pelo Sol. A razão para isto é a conservação da energia. A energia emitida da fonte por 1 segundo (potência) é finita. A mesma energia, que se emite pelo Sol, passa através da superfície de uma esfera imaginária à sua volta, seja ela de raio 58 milhões de quilómetros (distância do Sol ao Mercúrio) ou 150 milhões de quilómetros correspondentes ao raio da órbita da Terra. Como a área da superfície esférica aumenta com o raio como  $r^2$  ( $S_{esfera} = 4\pi r^2$ ), a energia transmitida pela onda de luz por unidade de área diminui como  $1/r^2$ . A energia da onda é proporcional ao quadrado da amplitude; daqui se conclui que a amplitude de uma onda esférica é inversamente proporcional à distância  $A(r) \propto 1/r$ . No entanto (voltando às fendas), quando se verifica que  $r \gg d$  (onde  $r$  é a distância do ponto de observação ao diafragma e  $d$  é a distância entre as fendas),  $r_1 \approx r_2$  e as amplitudes das duas ondas são muito próximas em qualquer ponto do espaço à direita do diafragma desde que este esteja longe das fendas.

## 7.2 A experiência com fenda dupla

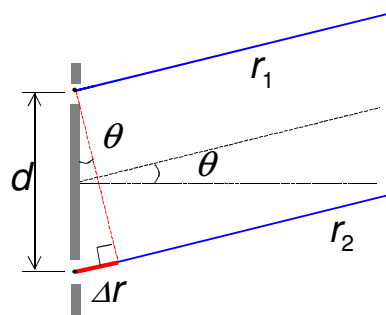
Agora já sabemos o suficiente para falarmos de uma experiência histórica, que também tem uma grande importância para a física moderna – uma experiência feita por um físico inglês, Thomas Young, em 1803, em que este mostrou claramente a natureza ondulatória da luz. A experiência do Young foi feita na época quando a ideia de luz, que dominava as mentes dos cientistas, era a teoria corpuscular de Newton que tratava a luz como um fluxo de





pequenas partículas. Thomas Young estava a observar a luz que passava através de um pequeno buraco numa cartolina, colocando um cartão, semelhante às cartas de jogar, em frente ao buraco ao longo do feixe de luz de forma a dividi-lo em dois. Com este equipamento muito simples Young mostrou a existência da interferência entre os dois raios de luz - o efeito típico das ondas. Em vez de observar num ecrã apenas duas manchas de luz separadas por uma zona escura no centro, como era natural esperar se a luz fosse um fluxo de pequenas partículas, observou uma imagem mais complexa – uma mancha clara no centro e ainda uma sequência de zonas claras e escuras que se estendiam para os dois lados. Posteriormente, a experiência foi melhorada e em vez do cartão passou utilizar-se um diafragma com duas fendas finas. Assim, a experiência de Young entrou na história da ciência como a experiência da fenda dupla.

O padrão da interferência num dado ponto, como já sabemos, é determinado pela diferença entre as distâncias percorridas pela luz emitida a partir de cada uma das fendas, i.e.  $r_2 - r_1$ . Considerando a geometria da experiência, verificamos que para as distâncias grandes, i.e.  $D \gg d$ , os dois raios de luz são praticamente



paralelos, então,  $\Delta r = r_2 - r_1$  é um cateto do triângulo rectângulo cuja hipotenusa é igual a  $d$  e o ângulo contraposto é igual a  $\theta$ . Deste triângulo temos  $\Delta r = d \sin \theta$ . Para que a interferência seja construtiva, tem que ser  $\Delta r = n\lambda$ . A interferência destrutiva observa-se quando  $\Delta r = \frac{\lambda}{2} + n\lambda$  ( $n = 0, \pm 1, \pm 2, \dots$ ). Daqui temos

$$\sin \theta_{\max} = n \frac{\lambda}{d},$$

onde  $\theta_{\max}$  são os ângulos em quais se observam os máximos da intensidade da luz, e

$$\sin \theta_{\min} = \left(n + \frac{1}{2}\right) \frac{\lambda}{d},$$

onde  $\theta_{\min}$  são os ângulos nos quais se observam os mínimos.

A posição dos máximos e dos mínimos no ecrã ( $y_{\max}$  e  $y_{\min}$ ) determina-se através dos ângulos

Posição dos máximos de luminosidade da imagem com fenda dupla

$$\sin \theta_{\max} = n \frac{\lambda}{d}$$

$\theta_{\max}$  e  $\theta_{\min}$ :  $y_{\max} = D \cdot \tan \theta_{\max}$  e  $y_{\min} = D \cdot \tan \theta_{\min}$ . Como  $D \gg d$ , os ângulos são pequenos de modo que  $\tan \theta \approx \sin \theta$  (porque  $\cos \theta \approx 1$ ) e, então

$$y_{\max} \approx n \frac{\lambda}{d} D$$

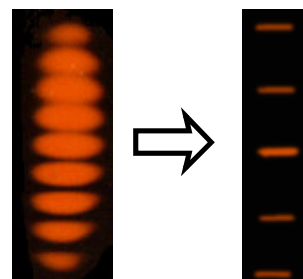
$$y_{\min} \approx \left(n + \frac{1}{2}\right) \frac{\lambda}{d} D \quad (n = 0, \pm 1, \pm 2, \dots).$$

Observamos que a distância entre os máximos é tanto maior quanto menor for a distância entre as fendas  $d$ .

A imagem obtida com uma fenda dupla mostra manchas de luz bastante largas. Mas à medida que aumentamos o número de fendas, mantendo a distância constante entre si, os máximos ficam cada vez mais estreitos e as zonas escuras mais largas. Isto deve-se ao facto de a imagem observada resultar da interferência de tantas ondas quantas fendas tem o diafragma. Os máximos de luminosidade formam-se em pontos em que todas as ondas têm a mesma fase. Contrariamente, o cancelamento das ondas acontece em pares e existem, portanto, muitas combinações possíveis que resultam em cancelamento (de facto, tantas quantas fendas há). Por exemplo, a 1ª onda cancela-se com a 2ª, a 3ª com a 4ª, etc., ou a 1ª com a 3ª, a 2ª com a 4ª, etc., ou a 1ª com a 4ª, a 2ª com a 5ª, etc, etc. Os instrumentos deste tipo chamam-se *redes de difracção* e utilizam-se muito para a análise dos espectros de luz emitida por materiais. A equação para a posição dos máximos de luminosidade é a mesma que para a fenda dupla:

$$y_{\max} = n \frac{\lambda}{d} D. \text{ Como } y_{\max} \text{ depende do comprimento de onda } \lambda, \text{ ao medir o intervalo}$$

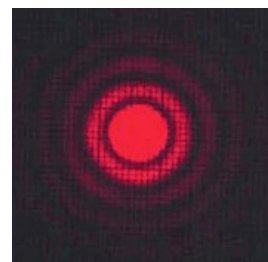
entre os máximos e sabendo a distância entre as fendas  $d$  e a distância entre a rede de difracção e o ecrã  $D$ , podemos determinar o comprimento de onda da luz incidente. É possível, deste modo, obter a informação precisa acerca da composição do material que emite a luz. A distância de separação entre os máximos é tanto maior quanto menor for a distância entre as fendas  $d$  e os máximos são mais nítidos quanto mais fendas tem a rede. As redes de difracção típicas têm entre 10 e 1000 fendas por milímetro.



## 7.3 Difracção

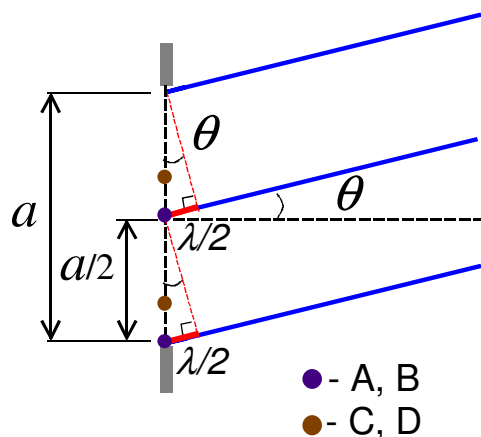
Entende-se o termo de *difracção* como um conjunto de fenómenos resultantes da propagação das ondas num meio em que existem não uniformidades bem pronunciadas, tais como orifícios ou fendas num ecrã, obstáculos, etc. Neste sentido, podemos dizer que a imagem observada com uma fenda dupla mostra a difracção da luz. No entanto, não podemos esquecer que todos os fenómenos deste tipo são, no fundo, o resultado da sobreposição das ondas, sejam estas apenas duas, ou em maior número, mas finito, ou em número infinito, até.

Ao falar-se sobre fenda dupla ou rede de difracção subentende-se que podemos ignorar o tamanho finito das fendas e considerar, em consequência disso, que cada fenda funciona como uma fonte pontual das ondas secundárias. Existem, no entanto, situações em que esta não é uma boa aproximação. Por exemplo, fazendo passar a luz por um orifício pequeno observa-se uma imagem complexa que não se consegue explicar considerando o orifício como uma fonte de ondas esféricas pontual (ver imagem ao lado). Para explicar este efeito, temos que tomar em conta a interferência das ondas emitidas por diferentes partes da abertura.



Consideremos uma abertura de largura  $a$  (uma fenda, para simplificar) e vamos procurar a equação para os mínimos de interferência aproveitando o facto de as ondas se cancelarem aos pares, como já verificámos quando falámos sobre redes de difracção. Existe alguma semelhança entre as duas situações, mas há também diferenças. Em ambos os casos há muitas fontes da luz, mas se no caso de uma rede de difracção as fontes são discretas e o seu número é, em princípio, finito, aqui temos um conjunto de fontes extensas e em número infinito, no limite.

Para começar, dividimos a abertura do diafragma em duas metades, como se mostra na figura. Se o ângulo  $\theta$  for tal, que o cateto contraposto (em triângulos a vermelho) é igual a  $\frac{\lambda}{2}$ , então para qualquer ponto-emissor da primeira metade da fenda existe um outro ponto-emissor na segunda

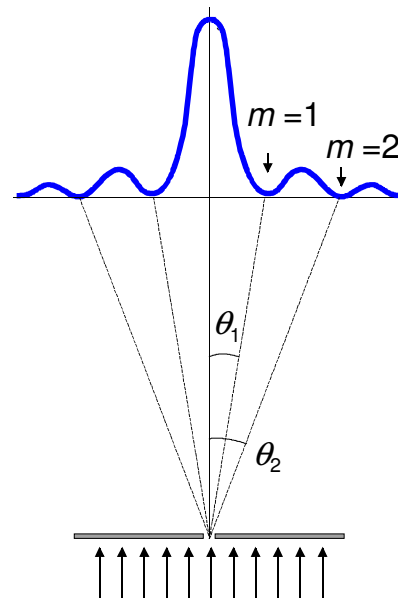


metade (a distância  $\frac{a}{2}$  do primeiro) tal, que a diferença entre os caminhos percorridos pelas duas ondas, quando essas chegarem ao ecrã, é igual a  $\frac{\lambda}{2}$ . Essas duas ondas cancelam-se, como é óbvio. Assim é, por exemplo, para os pares de pontos A e B, C e D, etc, na figura. O cateto e a hipotenusa de cada um dos triângulos em causa estão relacionados entre si como  $\frac{a}{2} \sin \theta_{\min} = \frac{\lambda}{2}$ , de modo que o ângulo correspondente ao primeiro mínimo é igual a  $\sin \theta_{\min} = \frac{\lambda}{a}$ . Os mínimos mais afastados do centro estão posicionados nos ângulos maiores, o que corresponde a uma distância menor entre as fontes cujas ondas se cancelam. Isto equivale, simplesmente, à divisão da fenda em partes mais pequenas, mas não esqueçamos que cada ponto tem que encontrar o seu par ou, seja, o número das partes tem que ser par – 2, 4, 6, 8, etc. A condição para observar o cancelamento das ondas é, então  $\frac{a}{2m} \sin \theta_{\min} = \frac{\lambda}{2}$ , onde  $\frac{a}{2m}$  é o comprimento de uma parte da fenda e  $m = \pm 1, \pm 2, \dots$ . Assim, a equação para a posição dos mínimos de difracção em relação ao máximo central é

$$\sin \theta_{\min} = m \frac{\lambda}{a} \quad (m = \pm 1, \pm 2, \dots).$$

O que nos chama a atenção nesta fórmula é o facto de a dimensão da abertura  $a$  estar no denominador. Parece um contra senso – ao diminuir a abertura do buraco a dimensão da imagem deste na parede aumenta... Então, será que alguém já alguma vez viu a imagem da janela na parede da cozinha a aumentar quando se fecham as persianas?! Claro que não e já vamos explicar porquê.

A descrição analítica das imagens de difracção obtidas com objectos diferentes é muito complicada. Por isso, os padrões de difracção obtêm-se muitas vezes por construções geométricas, que se baseiam no princípio de Huygens. Rigorosamente dito,



Largura do máximo central da difracção

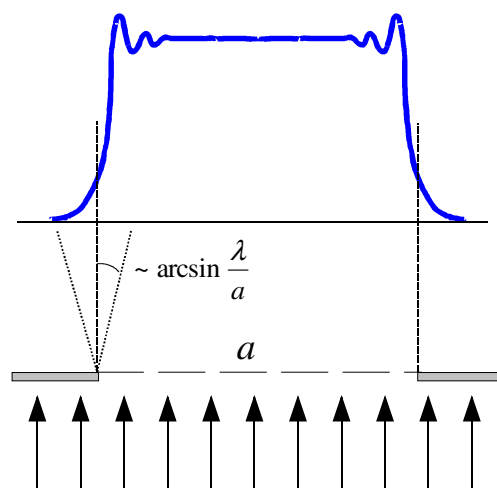
$$\sin \theta_1 = \frac{\lambda}{a}$$

a fórmula obtida para  $\theta_{\min}$  é válida só para as aberturas não muito grandes. Mesmo assim, ela permite-nos ter uma ideia simples, mas correcta sobre a física em causa. Como se vê nas imagens, a intensidade do máximo central é muito superior à intensidade dos que o acompanham. É este pico que se vê normalmente em condições habituais quando não se tomam medidas especiais. O ângulo  $\theta_1$ , que é  $\theta_{\min}$  para  $m = 1$ , caracteriza a largura deste pico ou, seja, a largura da imagem.

Consideremos as três seguintes situações.

1)  $a \gg \lambda$  (a abertura é grande relativamente o comprimento de onda da onda incidente). Neste caso,  $\sin \theta_1 \ll 1$  ou, seja, o ângulo é muito pequeno. Isto significa que os raios de luz (ou de outra onda qualquer) praticamente não se desviam do seu caminho inicial e a imagem, que produz a abertura no ecrã, é simplesmente a projecção geométrica da abertura. É isso

que observamos normalmente em casa: uma janela com persianas abertas produz na parede uma imagem com as mesmas dimensões, mas, ao fecharmos gradualmente as persianas, o tamanho da imagem diminui. O comprimento de onda da luz incidente na janela é muitas ordens de grandeza inferior às dimensões da janela. Esta é a razão por que o efeito da difracção não se nota. No entanto, a difracção da luz

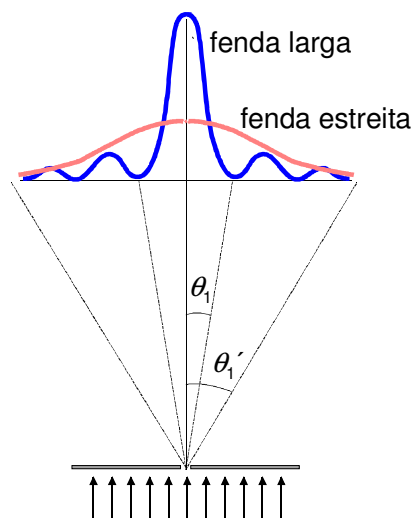


existe mesmo neste caso. O seu efeito revela-se perto das bordas da janela (por isso por vezes chama-se efeito das bordas). Podemos vê-lo, por vezes, como um pequeno arco-íris nas fronteiras das imagens (como o ângulo  $\theta_1$  depende do comprimento de onda, o desvio da luz de cor diferente também é diferente). O ângulo  $\sim \arcsin \frac{\lambda}{a}$

caracteriza a dimensão da zona em que a difracção é notável embora não tenha neste caso significado tão exacto como nós definimos acima para uma fenda estreita (por isso usamos o sinal “ordem de grandeza”  $\sim$ ). Ao diminuirmos a abertura da janela diminuimos a parte central da imagem que é uniforme. A parte central parece estar livre do efeito de difracção porque para cada pequeno fragmento de abertura, que dá um mínimo de difracção num dado ponto, existe um outro fragmento, cujo máximo

central coincide com o mínimo proveniente do primeiro fragmento. A parte da imagem perto das bordas é pouco afectada quando começamos a fechar a janela, até o tamanho da abertura diminuir e ser comparável ao comprimento da onda. De modo geral concluímos, que no caso  $\lambda \ll a$  a difracção tem pouco efeito na imagem.

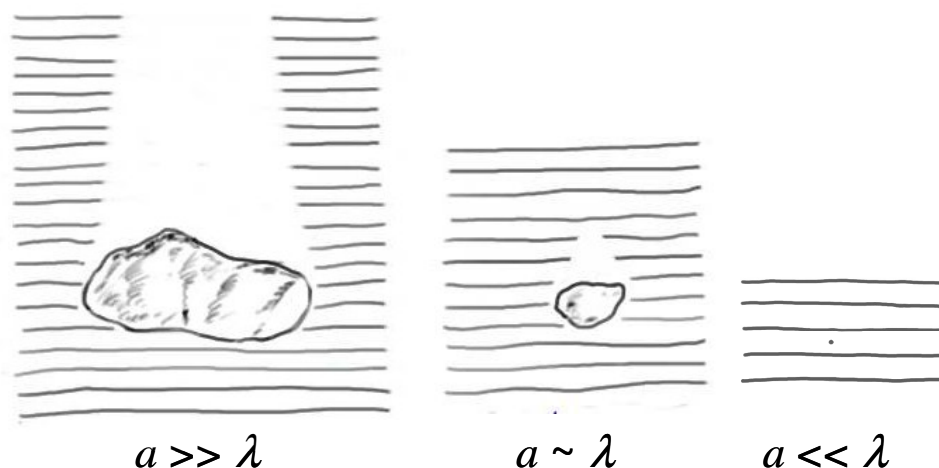
2) Caso  $a \sim \lambda$ . É precisamente o caso da fenda para o qual a equação  $\sin \theta_{\min} = m \frac{\lambda}{a}$  ( $m = \pm 1, \pm 2, \dots$ ) foi obtida. O ângulo  $\theta_{\min}$  indica a posição angular das zonas escuras na imagem da fenda que se observa num ecrã em relação ao máximo central. Para o primeiro mínimo da intensidade de iluminação do ecrã  $\sin \theta_1 = \frac{\lambda}{a}$ . Ao continuarmos a fechar a fenda (diminuir  $a$ ) a imagem desta fica mais larga (ver a figura) em vez de diminuir, como no caso anterior. Como o comprimento de onda da luz é da ordem de centenas de nanómetros (a luz verde, por exemplo, tem  $\lambda \approx 550 \text{ nm} = 5,5 \cdot 10^{-7} \text{ m}$ ), esta situação realiza-se só com as aberturas de algumas fracções de milímetro. O fenómeno de alargamento da imagem quando se fecha a fenda é muito importante não só por si (todos os instrumentos ópticos devem ter em conta a difracção), mas também para a física quântica, porque mostra como funciona o princípio de incerteza de Heizenberg do qual vamos falar no próximo capítulo. Concluimos, então, que o efeito da difracção é importante quando  $a \sim \lambda$ .



3)  $a < \lambda$ . Matematicamente esta situação não deve ser possível porque com estes valores de abertura  $\sin \theta_1 > 1$ . Na prática, claro, podemos fazer uma fenda destes (podemos, de facto, reduzi-la até às dimensões atómicas que são  $\sim 10^{-9} \text{ m}$ ). A interpretação física disto é que os objectos cuja dimensão é inferior ao comprimento de onda da onda incidente não afectam a propagação da onda no espaço. É como se estes não existissem. Um buraco destes não deixa passar a luz.

As mesmas conclusões gerais são também válidas no caso inverso, i.e. quando em vez de colocarmos no caminho da onda um obstáculo grande com uma abertura de tamanho  $a$ , colocamos um objecto com dimensões características  $a$ . Assim, a água do mar permanece calma atrás de uma rocha grande (mas, longe desta, oscila – é o efeito

da difracção nas bordas da rocha). Quando a largura da rocha é comparável ao comprimento de onda, o efeito da difracção é significativo: as ondas contornam o obstáculo, a zona de sombra é reduzida e há sempre alguma ondulação atrás da rocha. Finalmente, um objecto cuja dimensão é pequena em comparação com  $\lambda$  (um posto



para prender os barcos, por exemplo) não vai fazer diferença nenhuma – as ondas passam sem “o ver”.

Concluindo, o efeito de difracção torna-se muito importante quando o tamanho do objecto, com que a onda interage, é comparável ao comprimento de onda desta. A difracção impõe o limite de resolução dos instrumentos ópticos que é da ordem de  $1\ \mu\text{m}$ . Por isso é que os átomos não podem ser vistos mesmo com o melhor microscópio óptico (mas podemos vê-los com um feixe de electrões cujo comprimento de onda é menor que as dimensões dos átomos – veja-se o capítulo sobre mecânica quântica). Por isso é que não faz sentido fazer os pixeis dos sensores das máquinas fotográficas digitais, dos scanners e outros aparelhos para aquisição de imagens, mais pequenos do que alguns micrómetros (podemos verificar se as máquinas modernas já estão no limite de resolução ou ainda há espaço para melhorar). E por isso é que as fronteiras dos objectos nas imagens tiradas com luz visível nunca podem ser definidas com uma precisão melhor que  $\sim 1\ \mu\text{m}$ .

Finalmente, sublinhamos mais uma vez a importância da experiência com uma fenda com  $a \sim \lambda$  para percebermos melhor o princípio de incerteza de Heisenberg – um princípio fundamental do mundo quântico, que impõe um limite físico para a precisão com que as grandezas físicas podem ser medidas.

## Capítulo 6

# Mecânica Quântica

Neste capítulo vamos entrar no mundo da física quântica começando por falar das experiências e das ideias que levaram à nova física no início do século XX. O nascimento da teoria quântica não era pacífico. Muitos físicos consideravam-na uma aberração que destruía a harmonia a que a ciência tinha conseguido chegar no fim do século XIX no seu entendimento do mundo material. O mais revoltante, talvez, era o facto de a nova teoria acabar com a visão determinista do universo entregando-o, de certo modo, ao acaso, ao jogo imprevisível das probabilidades. “Deus não joga dados com o Universo” – foi mais ou menos assim que o grande Einstein, que contribuiu significativamente para o desenvolvimento da teoria quântica, exprimiu o seu desagrado com a aleatoriedade do mundo quântico. Na física clássica, o desenvolvimento de todos os processos era previsível e unicamente determinado. Ao saber a posição e a velocidade de um corpo num certo instante de tempo, podemos calcular a sua posição e a velocidade no instante seguinte, e assim por diante. A solução é única, basta saber as forças em qualquer ponto do universo e em qualquer instante de tempo. Faltava apenas completar essa informação para calcular o movimento de todos os corpos existentes, das partículas microscópicas às estrelas gigantes, a partir deste momento até ao futuro. A descoberta da dualidade da matéria, que representa as propriedades de uma onda e de uma partícula ao mesmo tempo, acabou com a ideia de que a posição e a velocidade podiam ser determinadas simultaneamente com exactidão e, consequentemente, com a possibilidade de os prever nos instantes mais próximos. Este é o conteúdo do princípio de incerteza de Heisenberg do que vamos falar na Secção 2.

De acordo com a teoria quântica, um processo, como nós o observamos, é apenas a realização de uma das possibilidades entre muitas. Mas seria errado pensar no mundo quântico como num caos absoluto. Este mundo tem as suas leis e as probabilidades de realização desta ou aquela hipótese estão sujeitas a essas leis. Por exemplo, a posição de um electrão no átomo de hidrogénio num certo instante de tempo não está definida, mas a probabilidade de o encontrar numa dada região do espaço ao efectuarmos uma



medição é perfeitamente calculável. A mecânica quântica oferece-nos um formalismo para calcular esta probabilidade através da solução da equação de Schrödinger (Secção 3). As soluções desta equação, as funções de onda  $\Psi(x,t)$ , contêm toda a informação necessária para esse efeito.

A nova física descobre novos efeitos, inimagináveis no mundo clássico. O efeito de túnel que permite, vulgarmente dito, passar através de uma parede é um deles. Na Secção 3.6 vamos mostrar como isso é possível. Finalmente, mas já no Capítulo seguinte, vamos discutir algumas questões fundamentais da física quântica relacionadas com a possibilidade de realização de uma nova forma de processamento da informação e de computação - informática quântica.

## 1. Como tudo começou

Tudo começou com uma pequena “nuvem” (usando as palavras de lord Kelvin) no horizonte brilhante que assombrava o edifício sólido, lógico e consistente da física, que parecia ser capaz de explicar todos os fenómenos conhecidos na altura. Esta nuvem, que inicialmente parecia insignificante e apenas embaraçosa, era a impossibilidade de descrever teoricamente o espectro de ondas electromagnéticas emitidas por um corpo aquecido como, por exemplo, por um prego posto nas brasas que a certa temperatura começa emitir a luz de cor vermelha – radiação do corpo negro, como se costuma chamar a este tipo de fenómenos. Para obter uma descrição satisfatória Max Plank, um físico alemão, foi forçado a romper com a ideia de continuidade dos processos na natureza e introduzir um elemento descontínuo, discreto – um quantum de energia (uma ideia que não era, aliás, absolutamente alheia à física – lembremo-nos do espectro de frequências com que vibra uma corda fixa em ambas as extremidades). Outras descobertas surgiram e, pouco a pouco, tornou-se claro que estávamos perante uma nova física que preenchia uma lacuna onde a física clássica deixava de ser válida.

Uma boa descrição destas descobertas pode ser encontrada no livro *Introdução à Física* (autores Jorge Dias de Deus, Mário Pimenta, Ana Noronha, Teresa Peña, Pedro Brogueiro – 2ª edição, McGraw-Hill, 2000). Não vamos repeti-la. Representamos apenas um diagrama resumido e referimos as respectivas secções do livro.

Energia é quantizada $\Delta E = h\nu$	<b>Max Plank</b> Radiação do corpo negro
A luz é constituída por fotões com energia $E = h\nu$ <b>Onda → partícula</b>	<b>Albert Einstein</b> Efeito fotoelétrico
A onda electromagnética tem momento linear $p = \frac{h}{\lambda}$	<b>Arthur H. Compton</b> Efeito de Compton
As partículas têm comprimento de onda $\lambda = \frac{h}{p}$ <b>Partícula → Onda</b>	<b>Louis de Broglie</b> Dualidade da matéria
A estrutura dos espectros atómicos explicada com a base nas ideias quânticas	<b>Niels Bohr</b> Modelo do átomo

### **1.1 A radiação do corpo negro (Max Plank)**

(DD, Capítulo 7, Secção 7.1.1, pp.501-504; até “Façamos um pequeno exercício de Física Estatística”)

### **1.2 Efeito fotoelétrico (Albert Einstein)**

(DD, Capítulo 7, Secção 7.1.2, pp.509-512)

### **1.3 Efeito de Compton (Arthur H. Compton)**

(DD, Capítulo 7, Secção 7.1.3, pp.512-513)

### **1.4 Dualidade onda-partícula (Louis de Broglie)**

(DD, Capítulo 7, Secção 7.1.4, pp.514-516)

### **1.5 Modelo de átomo de Bohr (Niels Bohr)**

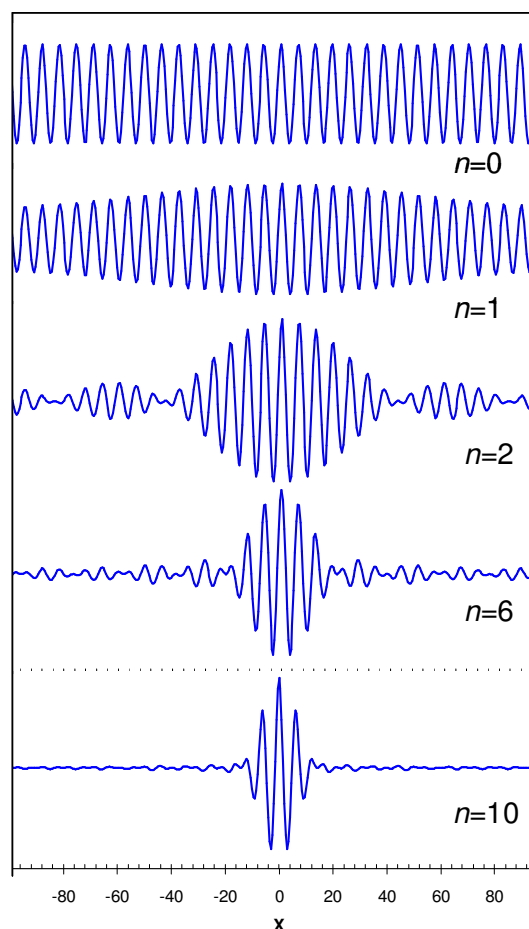
(DD, Capítulo 7, Secções 7.1.5 e 7.2, pp.517-525)

## 2. O princípio da incerteza

Assim que o princípio da dualidade da matéria é enunciado, surge uma pergunta. Se todos os objectos na Natureza têm propriedades de uma partícula e de uma onda simultaneamente, como é que nós podemos determinar, por exemplo, uma coisa tão simples e tão importante para a descrição de qualquer fenómeno físico, como a posição do corpo? Para uma partícula (i.e., uma partícula mesmo, sem qualquer propriedade ondulatória), não temos a mínima dúvida que podemos, em princípio, determinar a posição desta no espaço em qualquer instante de tempo. Mas para uma onda isto já não parece ser possível – a onda está, de facto, por todo o lado. Bem, podemos colocar esta pergunta retórica numa aula de física, mas quando tomamos o pequeno almoço não temos dúvida nenhuma que o pacote de leite está em cima da mesa da cozinha, e não “por todo o lado”, como as propriedades ondulatórias obrigariam a pensar...

Esta contradição não tem solução enquanto continuarmos a pensar em termos de onda harmónica e monocromática (i.e., com uma única frequência). Uma onda dessas, realmente, não tem nem início nem fim (caso contrário, não podia ser monocromática como já discutimos na Secção 4.3). Precisamos de fazer uma onda mais compacta, limitá-la no espaço de alguma maneira. Um impulso, talvez, fosse uma representação mais adequada.

Quando falámos no efeito de batimentos na Secção 4.3, verificámos que ao adicionarmos a uma onda harmónica uma outra onda com frequência ligeiramente diferente, uma onda puramente senoidal transformava-se numa onda pulsada. Se continuarmos a adicionar mais ondas com frequências próximas (mas diferentes!) podemos fazer com que os impulsos fiquem cada vez mais estreitos e mais espaçados entre



si. É precisamente assim que conseguimos construir um impulso quadrado isolado (ver a caixa no fim da Secção 4.3). Como neste caso estamos interessados na localização no espaço e não no tempo, como no caso dos batimentos, temos que somar as ondas com comprimentos de onda diferentes (e não frequências). A figura seguinte mostra o resultado da sobreposição de  $n$  ondas, com números de onda  $k = \frac{2\pi}{\lambda}$  diferentes; à onda-base corresponde  $n = 0$  na figura:

$$\Phi(x, t = \text{const}) = A_0 \cos(kx) + \sum_{i=1}^n A_i \{ \cos[(k - i\Delta k)x] + \cos[(k + i\Delta k)x] \}$$

em que  $k=1$  e  $\Delta k = 0.03$  (números relativos, apenas para ver o efeito).

A partir deste exemplo se vê que quanto mais ondas com diferentes  $k$  adicionarmos, melhor se consegue localizar a onda e melhor se conseguem conciliar os dois conceitos – conceito de onda e conceito de partícula. O preço a pagar pela melhor localização é que agora já não temos uma onda com o número de onda  $k$  bem definido, mas sim um conjunto de ondas (um *pacote* de ondas) em que os números de onda estão compreendidos num intervalo  $k \pm n\Delta k$ . Por outras palavras, isto significa que quanto mais precisão pretendemos ter na localização espacial, maior incerteza temos no conhecimento do número de onda (ou no comprimento de onda porque  $k = \frac{2\pi}{\lambda}$ ).

Os dois mundos – o mundo de ondas e o mundo de partículas – estão interligados através da relação de De Broglie  $\lambda = \frac{h}{p}$ . Com umas transformações algébricas simples podemos verificar que o número de onda é directamente proporcional ao momento linear  $p = \hbar k$  (em que  $\hbar = \frac{h}{2\pi}$  também é a constante de Plank). Aqui chegamos a uma conclusão de extrema importância para toda a física moderna: quando melhor for o nosso conhecimento da posição do objecto, pior será definido o momento linear (ou a velocidade) deste. Este facto foi descoberto por um físico alemão Werner Heisenberg e é conhecido hoje como o *princípio da incerteza de Heisenberg*. Este princípio afirma que é impossível conhecer em simultâneo e com exactidão a posição de uma partícula e a sua velocidade.

Em termos formais o princípio da incerteza exprime-se frequentemente como

$$\Delta x \cdot \Delta p_x \geq \hbar$$

onde  $\Delta x$  e  $\Delta p_x$  são incertezas em  $x$  e em componente  $x$  do momento linear  $p$ , respectivamente (em termos ainda mais rigorosos deveríamos escrever  $\sigma_x \sigma_p \geq \frac{\hbar}{2}$ , onde  $\sigma_x$  e  $\sigma_p$  têm significado do desvio padrão da respectiva grandeza).

Nas considerações anteriores, o princípio da incerteza surgiu como uma propriedade matemática (para quem já estiver habituado à análise de Fourier, podemos adiantar que é assim mesmo – isto é uma das propriedades da transformada de Fourier!). Para percebermos melhor como o princípio da incerteza se revela na prática, consideremos a seguinte experiência, também já por nós conhecida: a passagem das partículas (electrões, por hipótese) através de uma fenda de largura  $a$ . Como os electrões também possuem propriedades ondulatórias, deve existir o efeito de difracção (ver Secção 7.3). A difracção resulta em que os electrões não serão projectados no ecrã geometricamente, as vão sofrer um desvio da sua trajectória inicial e produzir uma imagem semelhante à produzida por uma onda de luz (ver a figura). A largura angular do máximo central é caracterizada por um ângulo  $\theta_m$ , relacionado com o comprimento de onda da radiação incidente  $\lambda$  (para os electrões  $\lambda = \frac{h}{p}$ ) e com a dimensão da fenda, através da expressão  $\sin \theta_m = \frac{\lambda}{a}$ , como é sabido (Secção 7.3). Esta equação mostra claramente, que quanto melhor conhecermos a posição  $z$  dos electrões que produzem a imagem, i.e. quanto menor for a abertura da fenda  $a$ , mais larga será a imagem ou, em termos mais físicos, maior será a indefinição no componente  $z$  do vector de velocidade dos electrões que passaram a fenda, ou ainda, no momento linear  $\Delta p_z$ .

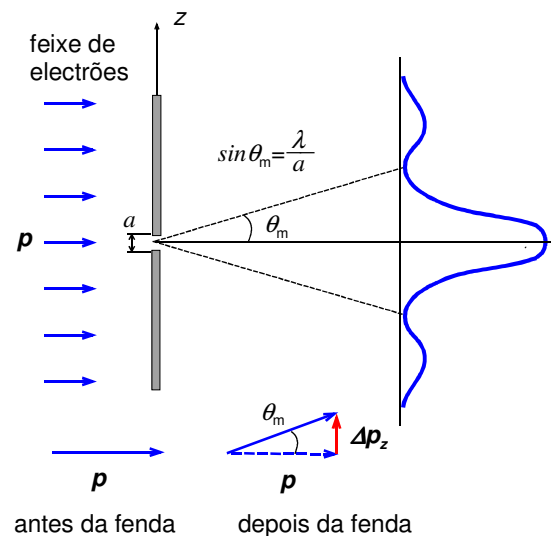
Podemos facilmente relacionar  $\Delta p_z$  com o momento inicial e o ângulo  $\theta_m$ :

$$\Delta p_z = p \sin \theta_m \quad \text{Com } p = \frac{h}{\lambda} \text{ e } \sin \theta_m = \frac{\lambda}{a}$$

temos  $\Delta p_z = \frac{h}{a}$ . Como a precisão com que

podemos conhecer a coordenada  $z$  das partículas não pode exceder  $a$ , chegamos à expressão equivalente à do princípio da

$$\text{incerteza } \Delta p_z = \frac{h}{\Delta z}.$$



O princípio da incerteza estabelece um limite físico da precisão com que as grandezas físicas podem ser medidas. Este limite não tem nada a ver com as técnicas e instrumentos de medição utilizados e que, obviamente, podem sempre ser melhorados. O princípio da incerteza reflecte as propriedades básicas da Natureza e não é ultrapassável, qualquer que sejam as tecnologias de medição disponíveis. Este princípio é uma consequência directa da dualidade da matéria onda-partícula.

O princípio da incerteza teve um impacto dramático na nossa visão do mundo físico. A ideia determinista que dominava a ciência no início do Século XX e que se baseava no facto de, de acordo com a mecânica Newtoniana, o conhecimento exacto da posição e da velocidade dos corpos num dado instante de tempo ser suficiente para prever o futuro desenvolvimento do sistema. Mas, como afirma o princípio de Heisenberg, são precisamente esses dois parâmetros – posição e velocidade – que não podem ser medidos com exactidão, ao mesmo tempo. Se nem sequer estes podem ser conhecidos com a exactidão, o futuro também não pode ser conhecido ao certo.

### 3. A equação de Schrödinger

Eis a equação famosa que Ervin Schrödinger, um físico austríaco, compôs em 1926 e que tem na física quântica um papel semelhante à segunda lei de Newton na mecânica clássica:

$$\left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + U(x) \right] \Psi(x,t) = i\hbar \frac{\partial}{\partial t} \Psi(x,t)$$

Aqui  $i = \sqrt{-1}$ ,  $\hbar = \frac{h}{2\pi}$ ,  $h = 6,63 \cdot 10^{-34}$  J·s é a constante de Plank,  $m$  é a massa da

partícula,  $U(x)$  - a energia potencial e a função  $\Psi(x,t)$  é a função de onda da partícula. A função  $\Psi(x,t)$  caracteriza a distribuição de probabilidades de encontrar a partícula numa certa região do espaço. É precisamente o significado desta função que está na origem do desagrado de muitos físicos que gostariam de continuar a viver num mundo que segue um caminho determinado. De acordo com a definição da  $\Psi(x,t)$ , a probabilidade, ao efectuar uma medição, de se encontrar a partícula num intervalo do espaço  $(x, x+dx)$  é  $dP(x,t) = |\Psi(x,t)|^2 dx$ . Para uma região de  $x \in (a,b)$ , a respectiva

probabilidade é  $P(t) = \int_a^b |\Psi(x,t)|^2 dx$ . Como a probabilidade de se medir a partícula

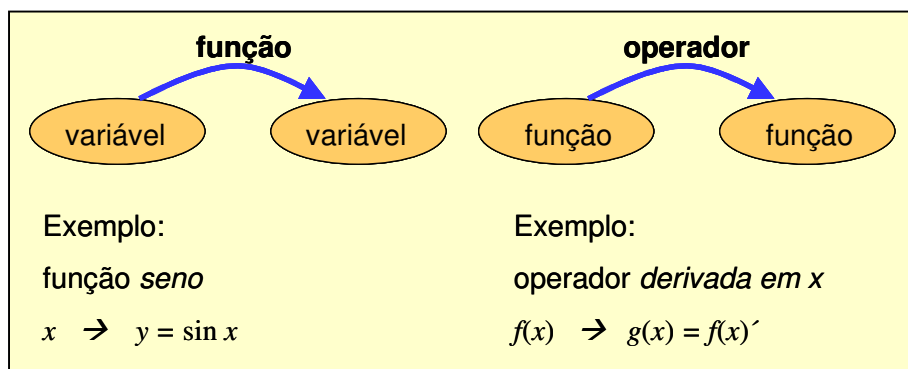
algures é igual a 1 (se esta existe, claro), o integral do quadrado do módulo da função de onda da partícula sobre todo o espaço deve ser igual a 1:  $\int_{-\infty}^{\infty} |\Psi(x,t)|^2 dx = 1$ . Esta condição chama-se normalização da função de onda.

Antes de continuarmos, convém fazer uma observação matemática acerca da forma como a equação está escrita. Em primeiro lugar, pode parecer estranho a seguinte escrita:  $\frac{\partial^2}{\partial x^2}$  e  $\frac{\partial}{\partial t}$ . Faz lembrar as derivadas parciais de segunda e de primeira ordem,

respectivamente, mas falta a função... Isto não é por acaso. A nova física necessita de um novo aparelho matemático. No formalismo da mecânica quântica, as entidades matemáticas que correspondem às grandezas físicas tais como energia, momento linear, momento angular, etc. já não são funções mas sim operadores que actuam sobre a função de onda (ver a caixa). Assim, o operador  $i\hbar \frac{\partial}{\partial t}$  do lado direito da equação de

Schrödinger é o operador da energia. Na prática, aquilo que está escrito do lado direito, é apenas  $i\hbar \frac{\partial \Psi(x,t)}{\partial t}$ . O termo  $-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$  do lado esquerdo corresponde à energia cinética e  $U(x)$  é o operador da energia potencial que sendo aplicado à função de onda resulta num simples produto  $U(x) \cdot \Psi(x,t)$ . Deste modo, a equação de Schrödinger na forma matemática mais habitual (mas fisicamente, talvez, menos transparente) é

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x,t)}{\partial x^2} + U(x) \cdot \Psi(x,t) = i\hbar \frac{\partial \Psi(x,t)}{\partial t}.$$



### 3.1 Análise da equação de Schrödinger

O caminho pelo qual Schrödinger chegou à sua equação é um assunto que, em si, merece uma investigação. Mas pelo menos duas coisas são claras: a futura equação deveria conservar a energia e qualquer função que descrevesse uma onda deveria satisfazê-la. Primeiro porque em todo o conjunto de experiências que existiam na altura (e que existem até agora) não se encontrava nenhuma razão para duvidar que a energia se conservava em todos os processos conhecidos. A segunda condição é a sequência de dualidade da matéria que, de facto, está na base da toda a mecânica quântica. (Aliás, o primeiro nome que foi dado à nova física era *wave mechanics* - mecânica das ondas.)

Para melhor compreensão da equação de Schrödinger e da forma como ela funciona, vamos fazer um pequeno exercício e provar que uma onda é, realmente, uma solução da equação e que a lei de conservação da energia está satisfeita.

Consideremos uma onda plana (i.e., unidimensional) a propagar-se no sentido  $x$  positivo. A equação é complexa e, portanto, a forma complexa para a função de onda seria mais adequada:

$$\Psi(x,t) = Ae^{i(kx - \omega t)}.$$

Lembrando que a definição do número de onda é  $k = \frac{2\pi}{\lambda}$  e o comprimento de onda de

uma partícula (comprimento de onda de De Broglie) é  $\lambda = \frac{h}{p}$ , temos  $k = \frac{2\pi}{h} p = \frac{p}{\hbar}$ .

Por outro lado, a frequência da onda está relacionada com a energia da partícula como

$\omega = \frac{E}{\hbar}$ . Portanto, podemos reescrever a expressão para a função de onda da seguinte

maneira

$$\Psi(x,t) = Ae^{\frac{i}{\hbar}(px - Et)}.$$

Agora, só falta substituir  $\Psi(x,t)$  na equação de Schrödinger por esta expressão e verificar se a equação se tornará uma igualdade. Ao calcularmos as derivadas, temos

$$\frac{\partial \Psi(x,t)}{\partial t} = \frac{\partial}{\partial t} \left( Ae^{\frac{i}{\hbar}(px - Et)} \right) = -A \cdot \frac{iE}{\hbar} \cdot e^{\frac{i}{\hbar}(px - Et)}$$

$$\frac{\partial^2 \Psi(x,t)}{\partial x^2} = \frac{\partial}{\partial x} \left( \frac{\partial \Psi(x,t)}{\partial x} \right) = \frac{\partial}{\partial x} \left( A \cdot \frac{ip}{\hbar} \cdot e^{\frac{i}{\hbar}(px - Et)} \right) = A \cdot \left( \frac{ip}{\hbar} \right)^2 \cdot e^{\frac{i}{\hbar}(px - Et)}$$



Substituindo as derivadas parciais na equação pelos valores calculados, chegamos a

$$-\frac{\hbar^2}{2m} \cdot \left( -\frac{p^2}{\hbar^2} \cdot A e^{\frac{i}{\hbar}(px-Et)} \right) + U(x) \cdot A e^{\frac{i}{\hbar}(px-Et)} = i\hbar \cdot \left( -i \frac{E}{\hbar} \cdot A e^{\frac{i}{\hbar}(px-Et)} \right)$$

ou, fazendo multiplicação e curtando  $A e^{\frac{i}{\hbar}(px-Et)}$ ,

$$\frac{p^2}{2m} + U(x) = E,$$

o que é nada mais nada menos que a lei de conservação da energia  $T + U = E$  porque

$$\frac{p^2}{2m} = \frac{(mv)^2}{2m} = \frac{mv^2}{2} = T \text{ é a energia cinética, i.e., a equação tornou-se uma igualdade}$$

(uma tautologia) o que prova que a onda plana é uma solução da equação de Schrödinger (mas também há outras soluções).

Daqui torna-se claro o significado de cada termo na equação de Schrödinger:

$\left[ \underbrace{-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}}_{\text{Energia cinética}} + \underbrace{U(x)}_{\text{Energia potencial}} \right] \Psi(x,t) = i\hbar \underbrace{\frac{\partial}{\partial t}}_{\text{Energia total}} \Psi(x,t)$	
Operador da energia cinética	$\hat{T} = -\frac{\hbar^2}{2m} \cdot \frac{\partial^2}{\partial x^2}$
Operador da energia potencial	$\hat{U} = U(x)$
Operador da energia total	$\hat{E} = i\hbar \cdot \frac{\partial}{\partial t}$
Operador do momento linear	$\hat{P} = i\hbar \cdot \frac{\partial}{\partial x}$

De certo modo, a equação de Schrödinger tem na mecânica quântica um papel semelhante ao da segunda lei de Newton na mecânica clássica. Da mesma maneira que a segunda lei de Newton descreve todos os fenómenos mecânicos no mundo clássico (basta conhecer as forças que actuam no sistema e resolver a equação diferencial como fizemos, por exemplo, ao descrevermos as oscilações harmónicas), a equação de Schrödinger descreve todos os fenómenos quânticos desde que a velocidade seja pequena comparativamente com a velocidade da luz (i.e., fenómenos não relativistas). De mesma maneira, como toda a grande variedade de fenómenos clássicos está contida

na variedade das forças, a função  $U(x)$  – a energia potencial do sistema – tem dentro de si todo o mundo quântico.

### 3.2 A equação de Schrödinger independente do tempo

Existem muitas situações práticas em que a dependência temporal da função de onda não nos interessa. Por exemplo, quando estudamos o átomo de hidrogénio não nos interessa como varia no tempo a função de onda do electrão porque experimentalmente temos acesso apenas à distribuição da probabilidade média de o encontrar numa ou noutra região do espaço.

Na função  $\Psi(x,t)$  para a onda plana, as variáveis  $x$  e  $t$  separam-se facilmente:

$$\Psi(x,t) = A e^{\frac{i}{\hbar}(px-Et)} = A \cdot e^{\frac{i}{\hbar}px} \cdot e^{-\frac{i}{\hbar}Et}$$

de modo que o tempo entra na função através do termo multiplicativo  $\exp\left(-\frac{i}{\hbar}Et\right)$ .

Generalizando, afirma-se que qualquer função  $\Psi(x,t)$  pode ser desdobrada em produto do termo espacial e do termo dependente do tempo

$$\Psi(x,t) = \psi(x) \cdot e^{-\frac{i}{\hbar}Et}.$$

Se substituirmos  $\Psi(x,t)$  na equação de Schrödinger por este produto, temos

$$-\frac{\hbar^2}{2m} \cdot \frac{\partial^2}{\partial x^2} \left( \psi(x) \cdot e^{-\frac{i}{\hbar}Et} \right) + U(x) \cdot \psi(x) \cdot e^{-\frac{i}{\hbar}Et} = i\hbar \frac{\partial}{\partial t} \left( \psi(x) \cdot e^{-\frac{i}{\hbar}Et} \right)$$

$$-\frac{\hbar^2}{2m} \cdot e^{-\frac{i}{\hbar}Et} \cdot \frac{d^2 \psi(x)}{dx^2} + U(x) \cdot \psi(x) \cdot e^{-\frac{i}{\hbar}Et} = i\hbar \left( -\frac{i}{\hbar}E \right) \cdot \psi(x) \cdot e^{-\frac{i}{\hbar}Et}$$

e, depois de dividirmos ambas as partes da equação por  $e^{-\frac{i}{\hbar}Et}$  chegamos à equação de Schrödinger *independente do tempo*

$$\left[ -\frac{\hbar^2}{2m} \cdot \frac{d^2}{dx^2} + U(x) \right] \psi(x) = E \psi(x).$$

(Repare, que a derivada parcial em ordem a  $x$  foi substituída pela derivada completa porque a função de onda agora só tem uma única variável,  $x$ .) As soluções da equação de Schrödinger independente do tempo chamam-se *soluções estacionárias*. Eles

correspondem aos estados estacionários, i.e. os que não se alteram com o tempo. O sistema pode permanecer num estado destes durante um tempo infinito.

A condição de normalização da função  $\psi(x)$ , obviamente, mantém-se:

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1.$$

A equação de Schrödinger independente do tempo pode ser escrita na forma

$$\hat{H}\psi(x) = E\psi(x)$$

em que o operador  $\hat{H}$  é definido como  $\hat{H} = -\frac{\hbar^2}{2m} \cdot \frac{d^2}{dx^2} + U(x)$ . Do ponto de vista matemático, a equação de Schrödinger nesta forma representa um problema sobre valores e funções próprias: procura-se um conjunto de funções  $\psi_i(x)$  e de valores  $E_i$  tais que a acção do operador  $\hat{H}$  sobre cada uma dessas funções resulte numa simples multiplicação da função pelo respectivo valor  $E_i$ :  $\hat{H}\psi_i(x) = E_i \cdot \psi_i(x)$ . Este problema é bem conhecido na álgebra linear: para uma dada matriz  $A$  procuram-se os pares de vectores e de valores  $\{\vec{x}_i, a_i\}$  tais que  $A\vec{x}_i = a_i \cdot \vec{x}_i$ . A importância deste facto consiste na solução da equação de Schrödinger independente do tempo se tornar muito mais fácil do que a solução da equação de Schrödinger dependente do tempo, porque a álgebra desenvolveu um conjunto de receitas para este tipo de problemas.

Nas secções seguintes vamos resolver a equação de Schrödinger para alguns casos mais simples, mas importantes, para percebermos como funciona a mecânica quântica.

### **3.3 Soluções da equação de Schrödinger com $U=\text{const}$**

Como já foi referido, toda a variedade de fenómenos quânticos está contida na função de energia potencial  $U(x)$ . As funções diferentes correspondem às diferentes classes de fenómenos. Parece muito simples – basta resolver a equação para cada tipo de função e temos todo o mundo nas mãos. No entanto, não é assim tão fácil. Na realidade, existem poucas funções  $U(x)$  que tornam a solução exacta possível. E mesmo com a utilização de métodos numéricos nem sempre é possível resolvê-la para os sistemas mais complexos. Como se diz, o diabo está nos detalhes...

O caso mais simples é, obviamente, quando  $U(x) = U_0 = \text{const}$ . A equação de Schrödinger será neste caso

$$\left[ -\frac{\hbar^2}{2m} \cdot \frac{d^2}{dx^2} + U_0 \right] \psi(x) = E \psi(x).$$

Depois de fazermos algumas transformações simples, temos

$$\begin{aligned} -\frac{\hbar^2}{2m} \cdot \frac{d^2 \psi(x)}{dx^2} + U_0 \psi(x) &= E \psi(x) \\ -\frac{\hbar^2}{2m} \cdot \frac{d^2 \psi(x)}{dx^2} &= (E - U_0) \psi(x) \\ \frac{d^2 \psi(x)}{dx^2} &= -k^2 \psi(x), \end{aligned}$$

em que  $k^2 = \frac{2m(E - U_0)}{\hbar^2}$  é uma constante em relação a  $x$ . Esta equação, do ponto de vista matemático, é idêntica à que já encontrámos quando falámos das oscilações harmónicas simples,  $\frac{d^2 x(t)}{dt^2} = -k^2 x(t)$  (ver Capítulo 3), com uma diferença: a variável agora é  $x$  em vez do  $t$  e a função é  $\psi(x)$  em vez da  $x(t)$ . De resto é a mesma equação: procura-se uma função cuja segunda derivada seja igual ao simétrico da própria função. As funções que satisfazem a equação, como já sabemos, são  $\sin(kx)$ ,  $\cos(kx)$ ,  $e^{ikx}$  e  $e^{-ikx}$  como, também, qualquer combinação linear destas (de facto, das quatro funções só duas são independentes porque como *seno* tanto *co-seno* podem ser expressas através dos exponenciais  $\sin a = \frac{e^{ia} - e^{-ia}}{2i}$ ,  $\cos a = \frac{e^{ia} + e^{-ia}}{2}$  e vice versa,  $e^{ia} = \cos a + i \cdot \sin a$  e  $e^{-ia} = \cos a - i \cdot \sin a$ ).

Convém utilizarmos a forma exponencial, assim, a solução geral pode ser escrita na forma

$$\psi(x) = A e^{ikx} + B e^{-ikx}$$

onde  $A$  e  $B$  são constantes de integração (como a equação de Schrödinger é uma equação diferencial de segunda ordem, a solução desta passa necessariamente por duas integrações, cada uma das quais adiciona uma constante). Para determinar as constantes, precisamos de mais duas equações que, neste caso, se chamam *condições de fronteira* – um análogo das condições iniciais no caso das oscilações. Neste caso particular, com a energia potencial constante em todo o espaço, não existem fronteiras e, logo, as constantes  $A$  e  $B$  são indefinidas. Do ponto de vista da física, isto não nos deve preocupar muito porque a nossa suposição inicial é pouco realista – as fronteiras

existem sempre e, portanto,  $U$  não pode ser constante em todo o espaço. Nas secções seguintes aproximamo-nos mais da realidade, mas para já vamos tirar proveito da facilidade com que a equação de Schrödinger se resolveu e tentar obter algumas conclusões úteis.

A função de onda completa  $\Psi(x,t)$  obtém-se multiplicando  $\psi(x)$  por  $e^{\frac{i}{\hbar}Et}$  ou por  $e^{-i\omega t}$  (porque  $\frac{E}{\hbar} = \omega$ ):

$$\Psi(x,t) = \psi(x) \cdot e^{-i\omega t} = (Ae^{ikx} + Be^{-ikx}) \cdot e^{-i\omega t},$$

donde,

$$\Psi(x,t) = Ae^{i(kx-\omega t)} + Be^{-i(kx+\omega t)}.$$

É fácil de perceber que realidade física corresponde a esta função: o primeiro termo descreve uma onda harmónica plana que se desenvolve no sentido positivo de  $x$  enquanto o segundo termo corresponde a uma onda harmónica plana a propagar-se em sentido oposto. Ambas as ondas representam soluções da equação de Schrödinger com  $U = \text{const.}$  Cabe-nos a nós escolher aquela que se adapta à situação real; ou até podemos escolher as duas que sobrepostas resultam numa onda estacionária (A tem que ser igual a  $B$ , neste caso).

Pode-se perguntar (e com todo o direito) qual é a interpretação deste resultado em termos de partículas porque a dualidade onda-partícula da matéria permite-nos utilizar ambos os conceitos ao mesmo tempo. Para tal, voltemos à definição do número de onda  $k$ . Ao resolver a equação de Schrödinger, definimos  $k$  como  $k^2 = \frac{2m(E-U_0)}{\hbar^2}$ . A partir

desta definição e utilizando o facto de que  $k = \frac{2\pi}{\lambda}$  e  $\lambda = \frac{h}{p}$ , em que  $p$  é o momento

linear, chegamos a  $k = \frac{2\pi}{h} p = \frac{p}{\hbar}$ . Continuando,

$$\frac{p^2}{\hbar^2} = \frac{2m(E-U_0)}{\hbar^2} \Rightarrow \frac{p^2}{2m} = E - U_0$$

e, finalmente,  $E = \frac{p^2}{2m} + U_0$  em que  $\frac{p^2}{2m} = T$  é a energia cinética, como é óbvio. Uma

vez que  $U_0 = \text{const.}$ , também a energia cinética o é (assumimos, naturalmente, que o sistema é conservativo). Esta situação corresponde a uma *partícula livre* que se move com uma velocidade constante. “Livre” quer dizer que a partícula não está sujeita a

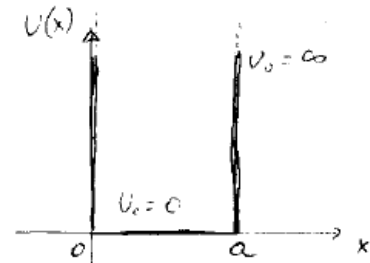
nenhuma força - caso contrário, a velocidade não seria constante. Isto está de acordo com a relação da mecânica clássica entre força e energia potencial, de que já temos conhecimento,  $F(x) = -\frac{dU(x)}{dx}$ : a derivada de uma constante é igual a 0.

Outra observação importante é que não existe nenhuma condição que limite os valores da energia  $E$  que uma partícula livre pode ter. O conjunto dos valores da energia é infinito e contínuo. Não será assim para uma partícula cujo movimento está limitado no espaço, que vamos considerar já a seguir.

### 3.4 Uma partícula numa caixa infinita

Consideremos uma partícula num campo com a energia potencial descrita pela seguinte expressão

$$U(x) = \begin{cases} \infty, & \text{para } x < 0 \\ 0, & \text{para } 0 \leq x \leq a \\ \infty, & \text{para } x > a \end{cases}$$



(partícula confinada numa caixa de largura  $a$  com as paredes infinitamente altas – ver a figura). Como em cada das três regiões a energia potencial é constante (um valor infinito também é uma constante), a solução da equação de Schrödinger é idêntica à solução obtida na secção anterior:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

com constante  $k$  definida como  $k = \frac{1}{\hbar} \sqrt{2m(E - U_0)}$ , em que  $U_0$  é igual ou a 0 ou a infinito, neste caso.

Dentro da caixa,  $U_0=0$  e, portanto,  $k = \frac{1}{\hbar} \sqrt{2mE}$ . Fora dela,  $U_0$  é infinitamente grande e positivo, de modo que  $2m(E - U_0) < 0$ , ou seja  $k$  é uma grandeza imaginária.

Podemos explicitar este facto escrevendo  $k = \frac{i}{\hbar} \sqrt{2m(U_0 - E)}$  em que a raiz quadrada já tem um valor real. Ao introduzir um novo parâmetro  $\alpha = \frac{1}{\hbar} \sqrt{2m(U_0 - E)}$  ( $\alpha$  é real) temos  $k = i\alpha$  e

$$\psi(x) = Ae^{-\alpha x} + Be^{\alpha x}.$$

Como  $\alpha = +\infty$ , claro está que o primeiro termo desta equação é igual a 0 qualquer que seja a constante  $A$  e, em relação ao segundo termo, podemos afirmar que  $B=0$  porque função de onda  $\psi(x)$  determina a densidade de probabilidade de encontrar a partícula num certo local do espaço e a probabilidade não pode ter um valor infinito.

Resumindo, a solução completa da equação de Schrödinger para uma partícula dentro de uma caixa com as paredes infinitas é

$$\psi(x) = \begin{cases} 0, & \text{para } x < 0 \\ Ae^{ikx} + Be^{-ikx}, & \text{para } 0 \leq x \leq a \\ 0, & \text{para } x > a \end{cases}$$

Falta, no entanto, determinar as constantes de integração  $A$  e  $B$ . Estas podem ser obtidas tomando em conta que a probabilidade não se pode alterar abruptamente de um ponto do espaço para outro, i.e.,  $\psi(x)$  deve ser contínua em função de  $x$  (de facto, a primeira derivada  $\psi'(x)$  também tem que ser contínua,). Isto significa, que

$$\begin{cases} \psi(0) = 0 \\ \psi(a) = 0 \end{cases}.$$

Da primeira equação concluímos que  $A = -B$  e, então

$$\psi(x) = Ae^{ikx} - Ae^{-ikx} = 2iA \cdot \frac{e^{ikx} - e^{-ikx}}{2i} = A_1 \sin kx$$

(aqui  $A_1$  é uma nova constante  $A_1 = 2iA$ ). O facto de  $\psi(a) = 0$  significa que

$$A_1 \sin ka = 0 \Rightarrow ka = \pi n \ (n = 0, \pm 1, \pm 2, \dots) \Rightarrow k = \frac{\pi n}{a}.$$

Lembrando que definimos  $k$  como  $k^2 = \frac{2mE}{\hbar^2}$ , podemos determinar a energia da partícula na caixa

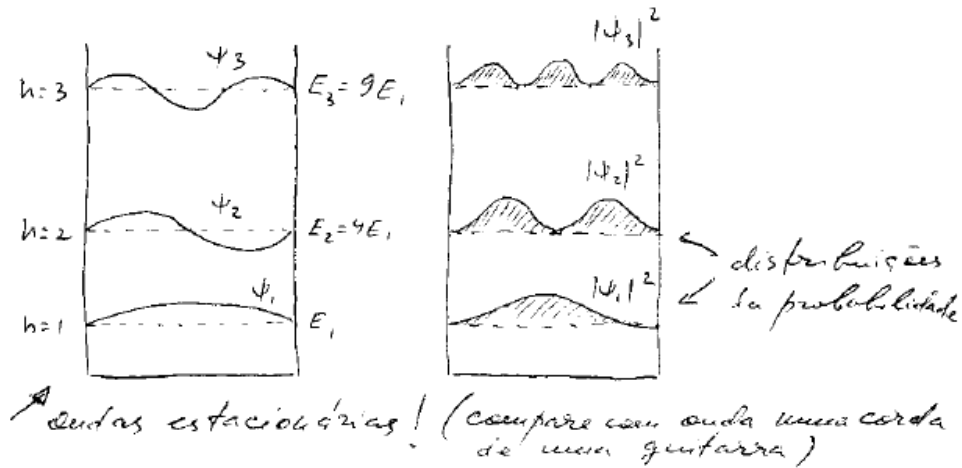
$$E = \frac{\hbar^2 k^2}{2m} = \frac{\pi^2 \hbar^2}{2ma^2} n^2.$$

Resumindo, a solução da equação de Schrödinger para uma partícula na caixa (i.e., conjunto de funções próprias e de valores próprios) é

$$\{\psi_n, E_n\} = \left\{ A_1 \sin\left(\frac{\pi nx}{a}\right), \frac{\pi^2 \hbar^2}{2ma^2} n^2 \right\}.$$

Ao definir  $E_1 = \frac{\pi^2 \hbar^2}{2ma^2}$ , temos o espectro de energias disponíveis para o sistema  $E_1, 4E_1,$

$9E_1, 16E_1,$  etc. e as funções de onda correspondentes  $\psi_1(x) = A_1 \sin \frac{\pi x}{a}$ ,



$\psi_2(x) = A_1 \sin \frac{2\pi x}{a}$ ,  $\psi_3(x) = A_1 \sin \frac{3\pi x}{a}$  etc. A figura seguinte mostra essas funções de

onda e as densidades de probabilidade respectivas  $|\psi_n(x)|^2$ . Como se vê, as funções de onda são ondas estacionárias, como se fossem vibrações de uma corda de guitarra (ver Capítulo 6), com uma frequência fundamental no nível  $E_1$  e harmónicas de respectiva ordem nos níveis de energia superiores.

### 3.5 Uma partícula numa caixa finita

O exemplo de uma partícula numa caixa com paredes infinitas, considerada na secção anterior é uma abstracção, como é óbvio. Nas situações reais, a energia potencial não pode ter valor infinito. No entanto, esta simplificação nos permitiu, de uma maneira simples e exacta, resolver a equação de Schrödinger e obter o conjunto das funções próprios  $\psi_n(x)$  e dos valores próprios,  $E_n$  e demonstrar a quantização da energia da partícula confinada a uma região do espaço. Para nos aproximarmos a situações mais realistas, vamos agora considerar um sistema com uma energia potencial parecida com a anterior, mas finita em todo o espaço:

$$U(x) = \begin{cases} U_0, & \text{para } x < 0 \\ 0, & \text{para } 0 \leq x \leq a \\ U_0, & \text{para } x > a \end{cases}$$



onde  $0 < U_0 < \infty$  (ver a figura). A solução da equação de Schrödinger para esta função  $U(x)$ , de facto, já foi obtida por nós na secção anterior. Dentro da caixa, esta é

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

com constante  $k$  definida como  $k = \frac{1}{\hbar} \sqrt{2m(E - U_0)}$ .

Fora da caixa, há duas situações possíveis:  $E > U_0$  e  $E < U_0$ . Interessa-nos a segunda situação porque só neste caso o movimento da partícula está confinado entre  $x = 0$  e  $x = a$ , o que pode ser visto facilmente tratando o problema classicamente (ver Capítulo 3). Neste caso, como no caso anterior,  $k$  tem um valor imaginário. A substituição  $k = i\alpha$ , com  $\alpha = \frac{1}{\hbar} \sqrt{2m(U_0 - E)}$ , leva-nos a uma solução do tipo

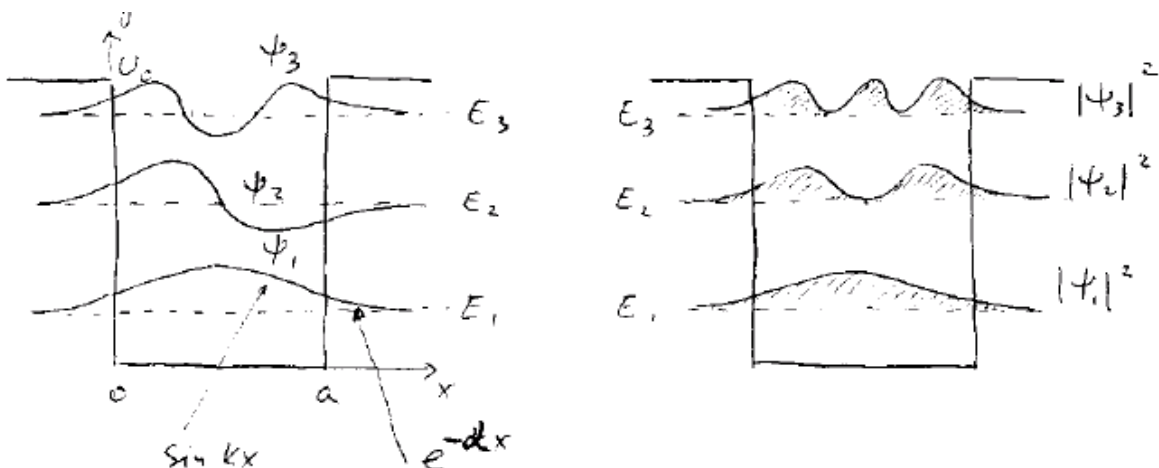
$$\psi(x) = Ae^{-\alpha x} + Be^{\alpha x}$$

na qual  $A = 0$ , para  $x < 0$ , e  $B = 0$ , para  $x > a$  (caso contrário,  $\psi(x) \rightarrow \infty$  com  $x \rightarrow -\infty$  e  $x \rightarrow \infty$ , respectivamente). Juntando as soluções numa única expressão, temos

$$\psi(x) = \begin{cases} C_1 e^{\alpha x}, & \text{para } x < 0 \\ Ae^{ikx} + Be^{-ikx}, & \text{para } 0 \leq x \leq a \\ C_2 e^{-\alpha x}, & \text{para } x > a \end{cases}$$

Foram introduzidas duas constantes novas  $C_1$  e  $C_2$  nas soluções fora da caixa porque, à partida, estas não são necessariamente iguais às constantes  $A$  e  $B$  dentro dessa. A relação entre todas as constantes pode ser obtida da condição de continuidade da função de onda em  $x = 0$  e  $x = a$ . Nós vamos concentrar a nossa atenção na análise qualitativa da solução obtida.

Vê-se que o comportamento da função de onda dentro da caixa é semelhante ao do caso anterior: é uma função oscilatória. No entanto, fora dos limites da caixa a função



de onda já não é igual a 0, como com energia potencial infinita, mas é representada por um exponencial decrescente. A continuidade da função de onda nas fronteiras requer que em  $x = 0$  e  $x = a$  esta seja diferente do 0. A figura mostra qualitativamente o comportamento das três primeiras funções de onda, tendo como ponto de partida as funções obtidas na secção anterior. Ao lado destes, está esboçado o quadrado do módulo de cada uma dessas funções que representa a densidade de probabilidade de detectar a partícula num dado região de espaço em função do  $x$ . O mais notável aqui é o facto de a probabilidade de encontrar a partícula dentro das paredes não ser exactamente 0, como seria de esperar do ponto de visto clássico (e como era no caso anterior). Lembremos, que física clássica proíbe tais situações porque como para  $x < 0$ , tanto para  $x > a$ , a energia mecânica  $E < U$  o que exige que a energia cinética da partícula  $T = \frac{mv^2}{2}$  seja negativa (ver Capítulo 3).

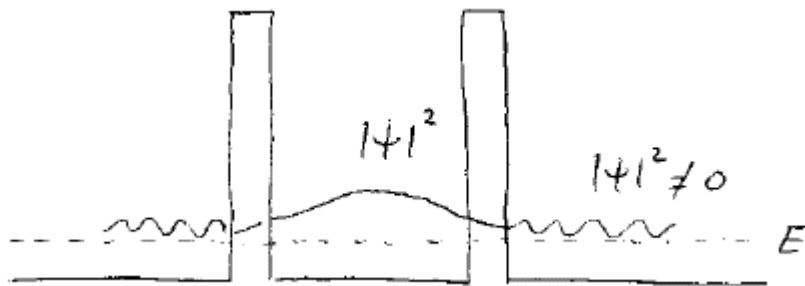
A penetração em regiões do espaço classicamente proibidas é uma das mais importantes propriedades dos objectos quânticos.

### 3.6 Efeito de túnel

Nas secções anteriores, começámos por considerar uma energia potencial que modela uma espécie da caixa de paredes infinitamente altas. A solução da equação de Schrödinger resultou em funções próprias (funções de onda) e valores próprios da energia do sistema. Assim, descobrimos que a energia do sistema é quantizada. Depois, baixámos a altura das paredes até um valor finito e descobrimos que a função de onda penetra as paredes da caixa. A probabilidade de penetração a uma certa profundidade decresce rapidamente com a profundidade, mas como uma função exponencial só tende para 0 assintoticamente, existe uma probabilidade não nula de encontrar a partícula dentro da parede a qualquer distância da sua superfície.

Vamos agora fazer as paredes da caixa não só finitas em altura, mas também em espessura (ver a figura). A função de onda terá um carácter oscilatório dentro da caixa, como antes. Nas regiões classicamente proibidas (i.e. onde  $E < U$ ), a função de onda descreve-se com um exponencial decrescente. Como este nunca chega ao 0, a função de onda da partícula tem um valor não nulo mesmo na fronteira exterior da parede. Isso obriga-nos a concluir que nas regiões exteriores às paredes a função de onda não pode ser igual a zero (caso contrário, havia uma descontinuidade na fronteira entre a parede e

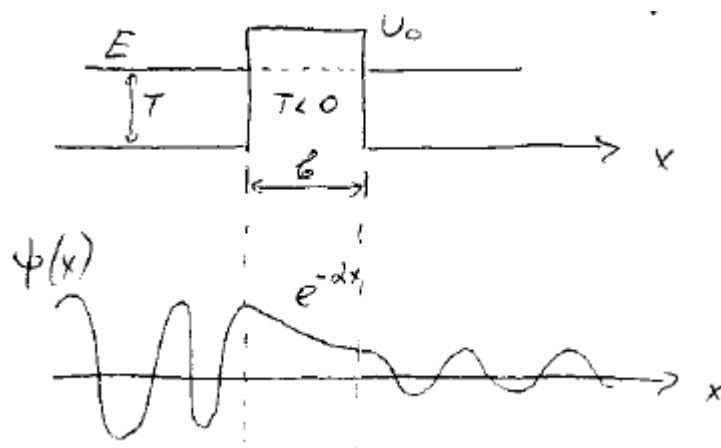
a região exterior). Como as regiões exteriores são semelhantes à região interior da caixa no que toca a relação entre a energia mecânica  $E$  e a energia potencial  $U$  ( $E > U$  em ambas),  $k$  tem um valor real, o expoente da  $e$  na função de onda é imaginário e, então, a função de onda é uma função oscilatória, como a figura mostra. Há, no entanto, uma diferença entre a região interior e as regiões exteriores. Na região interior existem duas fronteiras que limitam o movimento. Pensando classicamente – esses são os dois pontos de retorno em quais a partícula inverte a sua marcha. Isto resulta em que o movimento dentro da caixa tenha carácter oscilatório (ver Capítulo 3). Nas regiões exteriores só há um ponto de retorno de modo que se a partícula saiu da caixa para o lado direito, por exemplo, vai continuar a mover-se no mesmo sentido sem qualquer limite. A energia potencial nesta zona é constante; a força, então, é nula; consequentemente a partícula move-se como uma partícula livre. Em termos ondulatórios, uma partícula livre descreve-se com uma onda plana.



Matematicamente, dentro da caixa a função de onda tem que satisfazer duas condições de fronteira o que resulta em ondas estacionárias (ver Capítulo 5). Fora, a onda estacionária não se forma porque só há uma condição de fronteira. Além disso, a onda estacionária é o resultado de sobreposição de duas ondas a propagarem-se em sentidos contrários. Isto verifica-se dentro da caixa mas não fora desta.

Isto significa, que uma partícula inicialmente confinada numa dada região do espaço, onde, classicamente, devia permanecer durante um tempo infinito tem alguma probabilidade de ser encontrada fora dessa região, ou seja, de fugir da caixa. O efeito aqui considerado é conhecido como *efeito de túnel*. É um efeito puramente quântico sem qualquer analogia no mundo clássico.

A figura seguinte mostra a penetração de uma partícula, que incide numa barreira potencial de altura  $U_0$  e largura  $b$ , através dessa barreira. A probabilidade deste processo pode ser calculada sendo esta aproximadamente igual a  $P \approx e^{-2\alpha b}$  com  $\alpha = \frac{1}{\hbar} \sqrt{2m(U_0 - E)}$ , como antes. O comportamento da função de onda da partícula nestas circunstâncias também está qualitativamente esboçado. Lembremos, que é o



quadrado do módulo da função de onda que dá a probabilidade de descobrir a partícula numa dada região do espaço. Este é proporcional ao quadrado da amplitude da função de onda o que explica porque a amplitude das oscilações da  $\psi(x)$  é maior do lado esquerdo.

A existência do efeito de túnel foi confirmada experimentalmente muitas vezes. É este efeito que explica um dos tipos de decaimento radioactivo dos núcleos atómicos – decaimento  $\alpha$  – e também descreve a dinâmica do processo de cisão nuclear. O efeito de túnel utiliza-se também num microscópio especial, sensível até a escala atómica, em que electrões são emitidos por uma agulha atravessando depois uma barreira de potencial. Outro exemplo prático é um díodo de efeito de túnel – um componente electrónico cujo princípio de funcionamento se baseia neste efeito quântico. Neste caso, o efeito de túnel resulta na existência de uma região de tensões eléctricas com resistividade negativa (i.e., ao aumentar-se a tensão aplicada, a corrente eléctrica diminui em vez de aumentar também, como seria normal).

## Capítulo 7

# Informação Quântica

### 1. O gato de Schrödinger

Imaginemos uma caixa fechada e não transparente dentro da qual se coloca um gato. Também, se coloca um frasco com um veneno instantâneo com uma tampa, inicialmente fechada, mas que se pode abrir aleatoriamente em qualquer instante de tempo. É importante que o mecanismo de abrir o frasco seja um mecanismo aleatório, i.e. sobre qual nós não temos influência nenhuma. Um bom exemplo deste tipo de processo é o decaimento radioactivo de um núcleo: sabemos que mais cedo ou mais tarde o núcleo vai decair, mas não temos nenhuma maneira de prever em que instante acontecerá isto. Colocamos isso tudo na caixa e fechamo-la.

Algum tempo depois, perguntamos a nós próprios: como está o gato? Está vivo ou está morto? Se o frasco com veneno se abriu, o gato morreu. Se o frasco continua a estar fechado, o gato vive. Este raciocínio, aliás, não adianta nada porque não podemos saber se o frasco já abriu ou não. Então, ansiosos por saber, abrimos a caixa e olhamos para dentro, i.e., efectuamos uma medição do estado do sistema, falando em termos físicos. E então sim, saberemos com toda a certeza. Mas... estando na caixa fechada, o gato estava vivo ou estava morto?

Esta experiência conceitual foi sugerida por Ervin Schrödinger para mostrar uma importante propriedade da natureza, como esta é vista pela mecânica quântica. A descrição quântica de um sistema permite determinar os estados em que o sistema quântico pode estar (estados permitidos) e em que não pode estar (estados proibidos), mas não é capaz de dizer em qual destes estados o sistema realmente está num dado momento. A equação de Schrödinger dá-nos um conjunto de soluções (finito ou infinito) – pares dos valores próprios e das funções próprios  $\{E_i, \Psi_i(r, t)\}$  – mas não diz qual destes se realiza num dado instante. Apenas permite calcular a probabilidade para

cada um deles, igual a  $|\Psi_i(r,t)|^2$ . Na nossa situação imaginária, o sistema “gato na caixa”, tal como um sistema quântico, tem dois estados permitidos: estado “*lgato vivo*” e o estado “*lgato morto*” (os símbolos  $|\dots\rangle$  utilizam-se na mecânica quântica para designar um estado do sistema). Na impossibilidade de determinar, sem fazer uma medição, em qual destes dois estados está o sistema, não temos nada melhor do que supor que o sistema está numa sobreposição dos dois estados

$$|\text{gato na caixa}\rangle = \alpha_1 |\text{gato vivo}\rangle + \alpha_2 |\text{gato morto}\rangle$$

ou seja, o gato não está nem vivo, nem morto ( $\alpha_1$  e  $\alpha_2$  são duas constantes). Formalmente, esta representação justifica-se pelo facto de que, como sabemos, se  $\Psi_1$  e  $\Psi_2$  forem as soluções da equação de Schrödinger, uma combinação linear destas  $\Psi = \alpha_1 \Psi_1 + \alpha_2 \Psi_2$  é também uma solução (i.e., também é um estado permitido).

Repare, que estamos a falar sobre os estados de um sistema isolado. Claro, que uma medição sobre o sistema deve resultar num valor concreto. Ao abrir a caixa, podemos dizer sem a mínima dúvida qual é o estado do gato: vivo ou morto. Mas antes disso, não. Obviamente, o caso do gato é apenas uma ilustração, algo exagerada, de um fenómeno físico e, portanto, não está livre de inconsistências. Por exemplo, ao descobrir o gato vivo quando abrimos a caixa somos levados logicamente à conclusão de que este esteve sempre vivo, mesmo antes de se abrir a caixa (a menos que acreditemos na ressurreição dos gatos). Mas para um sistema mais simples, por exemplo um átomo que pode transitar espontaneamente de um estado para outro e vice versa, o facto de medir o sistema no estado  $|a\rangle$  ou  $|b\rangle$  não significa de maneira nenhuma que o sistema estava neste estado antes da medição. Lembremos, que uma medição significa necessariamente uma interacção com o sistema, ou seja, alteração do estado do sistema.

A impossibilidade de saber em que estado, entre uma gama de estados possíveis, está o sistema leva-nos à seguinte conclusão pouco habitual: *o sistema quântico está simultaneamente em todos os estados permitidos e só no acto de medir se concretiza num estado determinado*.

Que isto não é tão absurdo como pode parecer à primeira vista, pode-se constatar com a experiência com uma fenda dupla (mais uma vez!). Ao explicar o resultado desta experiência no Capítulo 5, aplicámos o princípio de Huygens às fendas, que permitiu considerá-los como duas fontes pontuais da luz coerentes. Cada um destas duas fontes emite ondas esféricas e a imagem, que se observa num ecrã, é o resultado da sobreposição dessas duas ondas da luz. Repare nas frases “duas fontes”, “duas ondas”.

Na altura, para explicar o efeito, não precisámos de considerar as ondas da luz como um feixe de partículas – os fótons (as partículas da dualidade onda-partícula no caso da luz). Ao pensar na luz como uma corrente contínua de fótons, de facto, não se altera muito a nossa imagem do fenómeno: alguns fótons passam através da uma das fendas, outras passam através da outra abertura e, ao chegar ao ecrã, interferem uns com outros como ondas (neste ponto precisamos voltar à imagem da luz como sendo ondas – o efeito de interferência só existe para ondas).

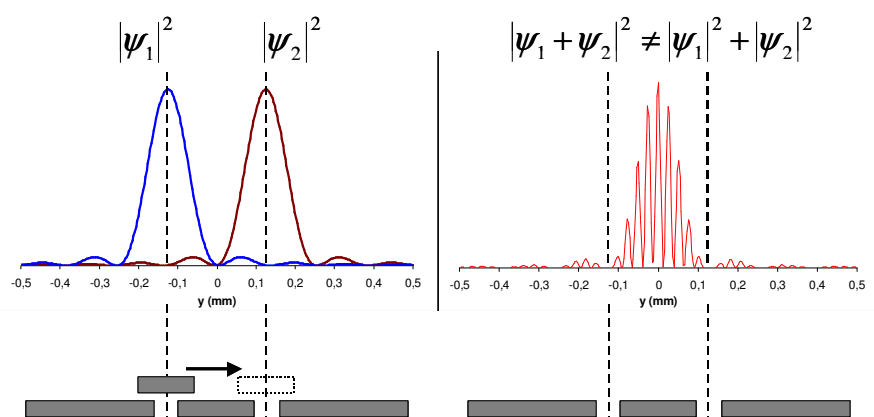
Agora vamos levar a experiência ao limite dos fluxos extremamente pequenos. Emitimos os fótons um a um, de tal modo que o intervalo de tempo entre eles é muito superior ao tempo necessário para um fóton atravessar todo o arranjo experimental e desaparecer no ecrã (i.e. ser absorvido pelo material do ecrã). Assim, já não podemos pensar no resultado como sendo a interferência entre os fótons que passaram pela primeira fenda e os que passaram pela outra. Claro, que já não se vê a imagem no ecrã (só temos um fóton de cada vez!), mas podemos substituir o ecrã por um detector de fótons com capacidade de medir as coordenadas do cada fóton. Ao registar um fóton, o detector envia um sinal com as coordenadas para um computador que os guarda na memória. Depois de acumular um número suficiente de eventos, podemos fazer um histograma da distribuição dos eventos ao longo da coordenada perpendicular às fendas. Traduzindo o histograma para a terminologia habitual (zonas luminosas, zonas escuras), podemos ver que existem regiões com mais eventos, correspondentes a uma iluminação elevada, e zonas com poucos eventos, correspondentes aos mínimos de interferência.

Estas experiências foram feitas em diferentes configurações, mas o resultado foi sempre mesmo: a imagem no plano de detecção obtida com fótons únicos era igual ao padrão de interferência observado com grandes fluxos de fótons, i.e. um máximo de luminosidade entre as fendas, seguido pelas zonas escuras, uma de cada lado, seguidas pelas bandas luminosas, bandas escuras etc. O efeito de interferência está claramente presente, mas, se só temos um fóton de cada vez, este interfere com quê? A resposta é que cada fóton passa simultaneamente por ambas as fendas como se fosse uma onda plana habitual e, então, o padrão de interferência no ecrã é o resultado de sobreposição do fóton com ele próprio. Se fosse eu a ler este texto, perguntava-me se o autor estava a espera de que eu acreditasse nisso...

No fim das contas, só a experiência pode apurar a verdade. Então, será que podemos verificar se um fóton realmente passa por ambas as fendas ao mesmo tempo, ou passa ora por uma, ora por outra, como deve ser num mundo *mais razoável*? Não há

nada mais fácil: fechamos aleatoriamente uma das fendas de modo a obrigar cada fóton a passar só por uma abertura. O resultado dessa experiência é representado na figura. O gráfico à esquerda representa as distribuições das coordenadas dos fótons (luminosidade, se fosse num ecrã e com um fluxo grande) medidas com uma fenda fechada<sup>1</sup>. À direita está o padrão medido com ambas as fendas abertas. Claramente, a soma das duas distribuições do lado esquerdo não dá a da direita.

Podemos fazer as seguintes conclusões desta observação. Primeiro, temos que excluir a hipótese de fóton passar só por uma fenda, quando temos as duas fendas abertas. A única possibilidade de explicar o padrão de interferência é admitir que o fóton passa através das ambas as fendas ao mesmo tempo. Em termos quânticos podemos dizer, que o fóton está em sobreposição de dois estados: estado  $|1\rangle$  com função de onda  $\Psi_1$ , que corresponde à passagem através da fenda número 1, e do estado



$|2\rangle$  com função de onda  $\Psi_2$ , que corresponde à passagem pela fenda número 2. A segunda observação é que uma tentativa de medir a trajectória do fóton (é isso que tentamos fazer ao fechar uma das fendas) obriga o fóton assumir definitivamente um dos dois desses estados.

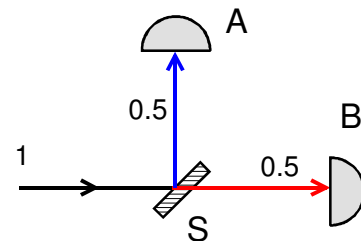
O padrão da imagem obtida numa ou noutra experiência pode ser descrita pelo quadrado do módulo da respectiva função da onda  $|\Psi_i(r)|^2$  porque, como já é sabido, o significado físico da função da onda na física quântica é: se uma partícula está no estado quântico  $|i\rangle$ , o quadrado de módulo da respectiva função da onda é a probabilidade de observar a partícula numa dada região de espaço  $P(r, r + \Delta r) = |\Psi_i(r)|^2 \Delta r$ . Deste modo,

<sup>1</sup> Os gráficos são calculados para as fendas de 0,05 mm de largura, 0,25 mm uma da outra, com o detector a 10 cm de distância das fendas;  $\lambda=650$  nm (luz vermelha).



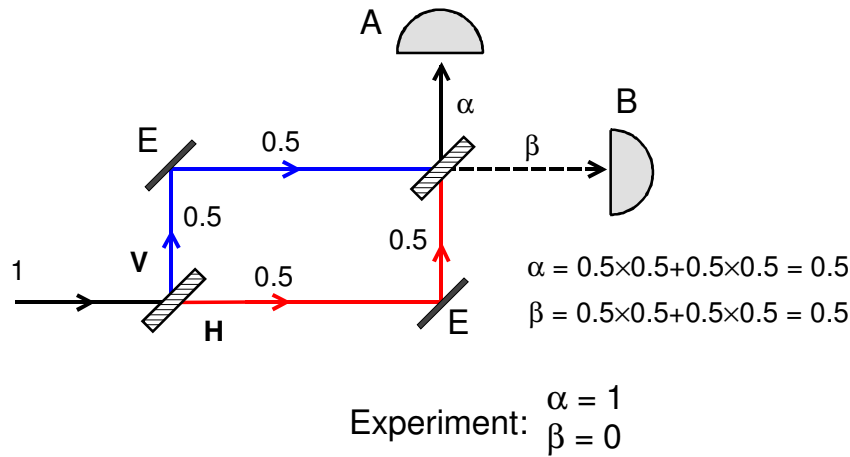
percebe-se porque razão a soma das duas imagens obtidas com uma das fendas fechada não é igual ao resultado medido com ambas as fendas abertas: o padrão de interferência no último caso descreve-se com  $|\Psi_1(r) + \Psi_2(r)|^2$  que não é igual à soma do  $|\Psi_1(r)|^2$  e  $|\Psi_2(r)|^2$ , correspondendo cada uma à experiência com uma das fendas fechada.

Mais uma experiência que aponta no mesmo sentido é a experiência em que se utiliza um divisor de feixe de fótons – *beam splitter*. No caso mais simples, um *beam splitter* é apenas uma placa de vidro semi-transparente que deixa passar exactamente 50% da luz incidente e reflecte a outra metade. Se utilizarmos um dispositivo destes (*S* na figura), os detectores *A* e *B* registrarão um número igual de fótons. A interpretação deste resultado é intuitiva. Pensando em termos de ondas, podemos dizer que 50% da energia da onda é reflectida do *beam splitter* e detectada pelo sensor *A*. Os outros 50% da energia passam através do dispositivo e produzem sinal no detector *B*. Uma coisa semelhante acontece quando a luz incide sobre um vidro normal: uma fracção dele é reflectida e o resto passa para dentro. Ao pensar na luz em termos de fluxo de fótons, podemos imaginar que 50% dos fótons incidentes são reflectidos da placa e 50% dos fótons atravessam-na. E se o fluxo for reduzido até termos apenas um fóton de cada vez? Um fóton único não se divide em dois. Ao chegar ao *beam splitter* este tem duas hipóteses: ou passar ou ser reflectido. Nenhuma dessas hipóteses pode ser excluída – caso contrário não poderemos explicar como é que se observam sinais em ambos os detectores quando utilizamos um feixe da luz de uma intensidade apreciável. Como aprendemos com a mecânica quântica, para uma partícula só podemos pensar em termos probabilísticos, o que neste caso quer dizer que o fóton tem 50% de probabilidade de passar e 50% de probabilidade de ser reflectido. Então, em metade dos casos vamos observar sinais no detector *A* (o detector *B* não detecta nada), e em metade dos casos o sinal vai aparecer só no detector *B*.



A surpresa vem quando tentamos dividir o feixe em dois, fazendo estes passar pelos caminhos diferentes, e juntamo-los outra vez com um arranjo experimental semelhante ao exposto na figura. Seguindo a lógica anterior, podemos raciocinar que no primeiro *beam splitter* um fóton tem 50% de probabilidade de passar e, então tomar caminho horizontal (*H*), e 50% de probabilidade de ser reflectido e seguir caminho vertical (*V*). Os dois espelhos *E* servem apenas para mudar a direcção e fazer convergir

ambos os caminhos no segundo *beam splitter*. Este segundo divisor do feixe é simétrico em todos os sentidos e, qualquer que seja a direcção de chegada dos fotões, estes tem 50% probabilidade de passar e 50% probabilidade de serem reflectidos. Assim os



caminhos *H* e *V* juntam-se (os dois feixes são misturados) e são formados dois caminhos possíveis (dois feixes emergentes) com detectores de fotões *A* e *B* no fim do percurso. Uma conta simples para a probabilidade de medir um fotão com detector *A* ou com detector *B*, dá

$$\alpha = 0,5 \times 0,5 + 0,5 \times 0,5 = 0,5$$

$$\beta = 0,5 \times 0,5 + 0,5 \times 0,5 = 0,5$$

respectivamente.

Porém, isto não acontece. A experiência com fotões únicos mostra que o detector *A* conta sempre e o detector *B* nunca. Mas se fecharmos um dos caminhos (não faz diferença qual) com um obstáculo, cada detector começa a contar o número igual de fotões, i.e. 25% cada um... Outra vez temos que reconhecer que a única explicação possível é admitir que cada fotão passa simultaneamente pelos ambos os caminhos (quando abertos, claro), ou seja, encontra-se em sobreposição dos dois estados possíveis: o estado  $|H\rangle$  e o estado  $|V\rangle$ .

## 2. Bits e qubits

A tecnologia moderna dos computadores, cada vez mais compactos, começou a exigir que as leis quânticas sejam tomadas em conta há muito tempo. Apesar disso, o processamento da informação nos computadores continua a ser feito de acordo com as regras clássicas. A quantidade mínima da informação continua a ser um bit, que pode

assumir dois valores: 0 ou 1. Na sequência disso, o portador físico de um bit, qualquer que seja a natureza física dele, deve poder existir em dois estados físicos bem distintos. Nos computadores antigos, para definir o endereço inicial do programa de arranque, utilizaram-se interruptores comuns com duas posições: um interruptor ligado introduzia à memória um bit com valor 1; o interruptor desligado dava bit 0. Nos circuitos lógicos um 1 pode corresponder a uma certa tensão eléctrica enquanto a ausência da tensão (tensão 0 V) pode significar um 0. Uma célula de memória RAM pode ser um simples condensador. Um integrado – uma caneta de memória – com uma capacidade de 512 Mbytes tem  $512 \times 8$  milhões desses condensadores. Um condensador carregado é interpretado como bit 1 e o descarregado – como bit 0 (ou vice versa, o que de facto não tem importância nenhuma). Embora a tecnologia tenha evoluído, o método de guardar e processar a informação não se alterou: a célula ou está no estado 1, ou está no estado 0.

Um sistema quântico, como já vimos anteriormente, encontra-se, geralmente, numa sobreposição de todos os estados permitidos para o sistema pelas leis quânticas. Deste modo, um sistema com dois estados, que podem ser designados por estado  $|0\rangle$  e estado  $|1\rangle$ , está sempre num estado misto  $\alpha|0\rangle + \beta|1\rangle$  em que os coeficientes  $\alpha$  e  $\beta$  caracterizam a probabilidade com que o sistema pode ser encontrada num determinado estado quando se efectua uma medição sobre este sistema. A probabilidade de medir o sistema no estado  $|0\rangle$  é igual a  $\alpha^2$  e a probabilidade de medi-la no estado  $|1\rangle$  é igual a  $\beta^2$ . Pode acontecer que um desses coeficientes seja desprezável, i.e.  $\approx 0$ ; o outro, então deve ser igual a 1, o que significa que as medições sobre o sistema resultarão sempre no mesmo estado, nomeadamente naquele cujo coeficiente é igual a 1. Esta situação corresponde a um bit clássico no estado 0, se  $\alpha=1$  e  $\beta=0$ , e no estado 1, se  $\alpha=0$  e  $\beta=1$ .

Em princípio, qualquer sistema quântico com dois estados pode ser usado como um suporte físico para um bit quântico – *qubit*. Por exemplo, um átomo no estado fundamental e no estado excitado; um fóton polarizado na vertical e um fóton polarizado na horizontal; uma partícula a propagar-se numa direcção e uma partícula a propagar-se na outra etc. Para percebermos o que os bits quânticos trazem de novo em relação aos bits clássicos, consideremos um sistema de dois qubits. Cada qubit pode encontrar-se num de dois estados,  $|0\rangle$  ou  $|1\rangle$ . Os dois juntos podem formar um sistema com quatro estados possíveis  $|0\rangle_1|0\rangle_2$ ,  $|0\rangle_1|1\rangle_2$ ,  $|1\rangle_1|0\rangle_2$  e  $|1\rangle_1|1\rangle_2$ . Aqui o índice em baixo indica o

número do qubit e a escrita  $|a\rangle_1|b\rangle_2$  corresponde ao produto das duas funções de onda: da função de onda que descreve o qubit nº1 no estado “a” e da função de onda correspondente ao qubit nº2 no estado “b”. Lembremos, que fisicamente a função de onda (no sentido quântico) está relacionada com a probabilidade de encontrar o sistema num dado estado (nomeadamente, o quadrado do módulo desta) e, portanto, a probabilidade, que ao efectuar uma medida do sistema de dois qubits, encontraremos qubit nº1 no estado “a” e qubit nº2 no estado “b” é igual ao produto das duas funções de onda (corresponde a uma operação lógica AND). Podemos utilizar esses quatro estados para codificar 4 bits de informação: 00, 01, 10 e 11, de mesma maneira como se faz num computador clássico. Até agora, não se vê nenhuma diferença entre os bits e qubits. Mas se lembrarmos, que um sistema quântico pode estar numa sobreposição dos estados e não necessariamente num estado certo como os bits clássicos, vêm logo propriedades novas.

Imaginemos que em vez de quatro estados tradicionais, i.e. com os valores bem definidos do cada bit (qubit), utilizássemos quatro combinações lineares destas, como, por exemplo,

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}} \{ |0\rangle_1 \cdot |1\rangle_2 + |1\rangle_1 \cdot |0\rangle_2 \}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}} \{ |0\rangle_1 \cdot |1\rangle_2 - |1\rangle_1 \cdot |0\rangle_2 \}$$

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} \{ |0\rangle_1 \cdot |0\rangle_2 + |1\rangle_1 \cdot |1\rangle_2 \}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}} \{ |0\rangle_1 \cdot |0\rangle_2 - |1\rangle_1 \cdot |1\rangle_2 \}$$

(estas quatro combinações são designados por *estados de Bell*; outras combinações também são possíveis). O que salta à vista nestes estados combinados, é que o estado de cada qubit não está definido em nenhum deles. Mais uma vez, estas expressões podem ser lidas como expressões lógicas com operações AND e OR. Por exemplo, o estado  $|\Psi^+\rangle$  é a seguinte combinação dos estados  $|0\rangle_1$ ,  $|1\rangle_1$ ,  $|0\rangle_2$  e  $|1\rangle_2$ :

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}} \{ (|0\rangle_1 \text{ AND } |1\rangle_2) \text{ OR } (|1\rangle_1 \text{ AND } |0\rangle_2) \}$$

(o coeficiente  $\frac{1}{\sqrt{2}}$  é apenas normalização da função de onda: o quadrado de módulo da função de onda deve ser igual a 1 o que significa que a probabilidade de medir o sistema em qualquer estado deve ser igual a 1). Por isso, ao efectuar uma medição do sistema no estado  $|\Psi^+\rangle$  podemos medir o bit nº1 como no estado  $|0\rangle$  tanto no estado  $|1\rangle$  com probabilidades iguais (porque ambos os estados estão presentes na função de onda  $|\Psi^+\rangle$  com pesos iguais). O mesmo é verdade para o bit nº2. No entanto, se por hipótese, a medição deu  $|0\rangle$  para o bit nº1, o bit nº2 tem que obrigatoriamente estar no estado  $|1\rangle$ , porque no estado  $|\Psi^+\rangle$  há combinação  $|0\rangle_1 \cdot |1\rangle_2$  e não existe combinação  $|0\rangle_1 \cdot |0\rangle_2$ . Da mesma maneira, se o bit nº1 estiver no estado  $|1\rangle$  o bit nº2 estará no estado  $|0\rangle$ .

Esta é a propriedade mais importante dos estados de Bell: os bits não são independentes, mas sim correlacionados. Esta correlação não tem analogia no mundo clássico e é mais forte do que qualquer outro tipo de correlações. Os estados correlacionados deste modo designam-se em inglês por *entangled states* (estados entrelaçados), e o fenómeno em geral designa-se por *entanglement*. Como já tínhamos notado, os valores de cada bit nos estados de Bell não estão definidos. Os bits assumam um valor concreto apenas no instante em que se efectua uma medição do sistema. Nesse momento, o estado  $|\Psi^+\rangle$  “aparece” perante dos olhos do experimentador ou como estado  $|0\rangle_1 \cdot |1\rangle_2$ , ou como estado  $|1\rangle_1 \cdot |0\rangle_2$ , i.e. já com valores de cada bit bem definidos. Imaginemos agora, que nós medimos apenas um dos dois bits do estado  $|\Psi^+\rangle$ , por exemplo o bit nº1. Neste instante sabemos o valor deste -  $|0\rangle_1$ , por hipótese. Mas se o bit nº1 for  $|0\rangle_1$ , o bit nº2 só poderá estar no estado  $|1\rangle_2$ . Isto quer dizer que no instante em que medimos bit nº1, o bit nº2 assume um valor definido apesar de não ter efectuado nenhuma operação (medição) sobre ele. Esta é a propriedade mais espantosa dos *entangled states*.

Antes de continuarmos, convinha responder a uma pergunta muito natural nestas circunstâncias: será que os *entangled states* são uma realidade física ou apenas uma fantasia matemática? Se esses estados são reais, deve haver uma maneira de os produzir no laboratório. De facto, a existência destes estados foi confirmada experimentalmente com vários sistemas. Um dos métodos de os criar é irradiar um cristal não linear com

uma luz ultravioleta. Neste caso, existe uma pequena (mas finita) probabilidade de conversão do fóton incidente em dois fótons de energia mais baixa (comprimento de onda maior), emitidos com um certo ângulo entre si e polarizados nos planos perpendiculares. As polarizações dos fótons não são definidas, mas mesmo assim sendo, estão correlacionadas: se a polarização de um dos fótons for medida a coincidir com o eixo  $x$  (por hipótese) o outro fóton seria obrigatoriamente polarizado ao longo do eixo  $y$ . Estes dois fótons podem ser desviados por reflectores e enviados por caminhos diferentes.

Olhando a situação de um outro ângulo, podemos questionar: como é que o segundo bit “sabe” que alguém tinha efectuado uma medição do primeiro bit? Fisicamente, os dois bits são dois sistemas distintos (dois átomos ou dois fótons, por exemplo) e até podem estar separados no espaço por uma distância grande. “Saber”, normalmente, implica a existência de uma interacção entre eles. Ora, esta interacção não está relacionada com nenhuma das quatro interacções fundamentais das quais falámos no Capítulo 3. É um fenómeno diferente, puramente quântico e instantâneo cujo “propagação” no espaço não está limitada pela velocidade da luz.

### 3. Criptografia quântica

Como um exemplo de aplicação de novos conceitos na informática consideremos a encriptação quântica – uma aplicação que parece estar mais perto à implementação prática do que outras.

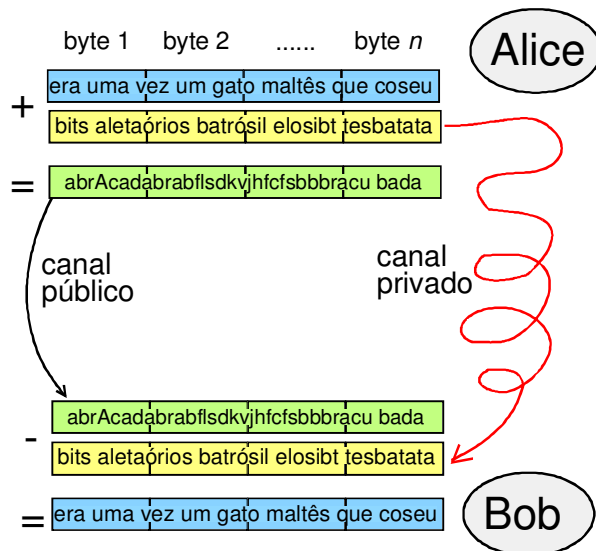
#### Métodos tradicionais de encriptação

Os métodos tradicionais mais utilizados baseiam-se na encriptação da mensagem a ser transmitida com uma chave sem a qual não é possível (leia-se: muito difícil) decodificar a informação recebida para obter a mensagem original. Consideremos dois tipos de encriptação: a encriptação com uma chave privada (mais segura, mas pouco prática) e encriptação com uma chave pública (mais prática, mas menos segura).

#### Encriptação com chave privada.

A chave privada é uma informação mantida em segredo e acessível apenas a duas pessoas: ao remetente e ao receptor. A seguinte figura mostra o princípio de funcionamento deste método. Imaginemos que a Alice pretende enviar a Bob (estas são as personagens tradicionais na teoria de encriptação) uma mensagem com  $n$  bytes de

comprimento. Produz-se, então, uma outra sequência de bytes do mesmo comprimento, chamada chave, e adiciona-se à mensagem, bit a bit. Se a sequência dos bits da chave for aleatória, o resultado desta soma será também uma sequência aleatória, i.e. um conjunto de símbolos sem qualquer sentido. É esta a mensagem que será enviada ao destinatário por um canal público, por exemplo



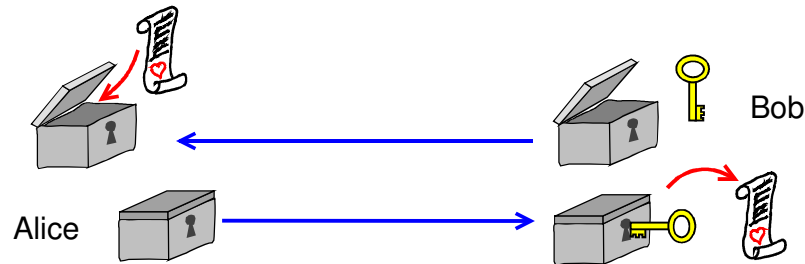
através da Internet. Bob, ao receber a mensagem distorcida (ou uma outra pessoa que o interceptou), não pode fazer sentido desta a menos que possua a chave. A chave é enviada ao Bob e este subtrai-a à mensagem recebida, obtendo a mensagem original. A grande inconveniência deste método é que a chave não pode ser enviada pelo canal público porque o risco de interceptação das ambas as comunicações é grande. Temos que a enviar por um canal de comunicação privado, por um serviço de estafeta especial, por exemplo.

A encriptação está baseada na aleatoriedade dos bits da chave e, portanto, o método é seguro desde que a chave seja utilizada apenas uma vez. Na segunda utilização esta propriedade perde-se e quando mais vezes a chave for usada, maior será a probabilidade de alguém conseguir reconstruí-la.

### Encriptação com chave pública

Este método, como o nome sugere, utiliza a encriptação com uma chave que não precisa de ser guardada em segredo. Deste modo, a mensagem encriptada e a chave podem ser enviadas por uma linha de comunicação aberta. Obviamente, um elemento secreto deve estar presente algures e uma chave privada existe de facto. No entanto, esta não é enviada pelas linhas de comunicação, sendo apenas utilizada para produzir a chave pública pelo destinatário que depois a comunica ao remetente (no primeiro método, ambas as pessoas têm que possuir a chave secreta para que o sistema funcione). Mostra-se na figura seguinte uma ilustração deste método. Bob está a espera de uma mensagem da Alice. Então, Bob compra uma caixa com fechadura e manda a caixa aberta a Alice, mas guarda a chave da fechadura. Alice recebe a caixa, escreve a sua

mensagem em texto aberto num papel, coloca o papel na caixa e fecha-a. A partir deste momento, a mensagem está protegida contra todos, incluindo a própria autora, e pode ser enviada abertamente ao Bob que, possuindo a chave da caixa, pode abri-la e ler a mensagem. Nesta história, algo vulgarizada, a chave pública é a caixa aberta, a mensagem encriptada é o texto colocado dentro da caixa fechada, e a chave privada é a



chave da fechadura da caixa. Repare, que a chave privada foi guardada no mesmo sítio e todas as comunicações se efectuam pelo canal público, o que constitui uma grande vantagem deste método.

Nas comunicações electrónicas, para a encriptação das mensagens utiliza-se um algoritmo matemático “com sentido único” (*one way function*, em Inglês), i.e. um algoritmo que permite facilmente fazer a encriptação da mensagem, mas torna muito difícil o cálculo no sentido contrário, mesmo sabendo a chave pública. Com a chave privada, no entanto, a descodificação torna-se possível. A função  $\sin(x)$ , embora não seja exactamente uma *one way function*, dá uma ideia sobre tal função: é fácil calcular  $f(x)=\sin(x)$  dado o  $x$ , mas não é possível determinar unicamente  $x$  sabendo  $f(x)$ . Na prática, usa-se um procedimento chamado factorização em números primos. Este baseia-se no teorema fundamental da aritmética que diz que para cada número inteiro e positivo existe uma única combinação dos números primos cujo produto é igual a este número (exemplos:  $20=1\times 20=1\times 2\times 10=\underline{1\times 2\times 2\times 5}$ ;  $77=1\times 77=\underline{1\times 7\times 11}$ ;  $1383=\underline{1\times 3\times 461}$ ). A representação de um número como um produto de números primos designa-se por factorização. É extremamente fácil calcular o produto de dois números primos (e obter a chave pública), mas para quem sabe apenas o produto, é muito mais difícil factorizá-lo (i.e. determinar quais são os números primos que o compõem). Claro, que aqui falamos em números muito maiores do que nos exemplos em cima. Mostra-se, que a dificuldade do processo de factorização de um número cresce exponencialmente com o número de dígitos que este tem. Por exemplo, em 2005, a factorização de um número com 200 dígitos decimais (663 bits) em dois números primos levou 500 anos do melhor CPU (para mostrar isto utilizaram-se muitos processadores em paralelo, como é óbvio). No



ano 2006, a encriptação com chave pública foi considerada suficientemente segura desde que a chave tenha pelo menos 1024 bits. Com a chave privada (um dos dois números primos), a descodificação torna-se muito simples.

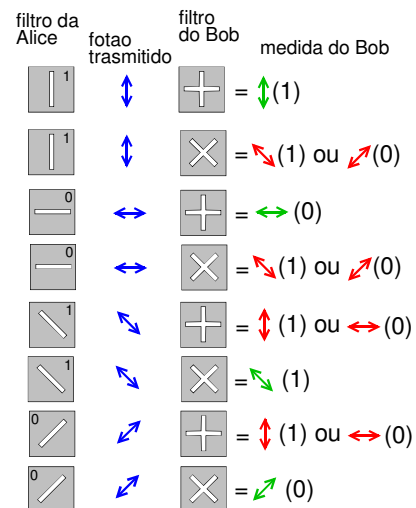
A grande vantagem deste método é que não precisamos enviar a chave por um canal secreto, como no primeiro caso, e isto torna-o muito mais prático. Nas comunicações via Internet utiliza-se muito este sistema. No entanto, como a chave pública está disponível para todos que a queiram ouvir, não existe garantia absoluta de que a mensagem encriptada (também aberta a toda a gente) não seja descodificada. O facto de nenhum algoritmo eficiente de quebrar a chave ter sido descoberto até agora, não quer dizer que este não exista de todo. Esta descoberta pode acontecer em qualquer momento (e até já pode ter acontecido, mas mantida em segredo!).

### **Encriptação quântica.**

A computação quântica afecta a segurança nas comunicações de duas maneiras. Por um lado, a muito maior rapidez com que um computador quântico poderá factorizar os números grandes vai imediatamente tornar obsoletos os algoritmos de encriptação com chave pública. Mas por outro lado, permitirá utilizar um método absolutamente novo (e, pelos vistos, absolutamente seguro) com que poderá ser transmitida a chave privada voltando, assim, ao primeiro método de encriptação - encriptação com chave privada. Além disso, há mais uma vantagem na encriptação quântica. Como qualquer medição de um sistema quântico resulta necessariamente em alteração do estado do sistema, temos um mecanismo de controlo para saber se a mensagem já foi lida por alguém durante a transmissão ou não.

O facto de um sistema quântico permanecer, geralmente, como uma sobreposição de todos os estados permitidos para o sistema (estes podem ser obtidos através de resolução da equação de Schrödinger), implica que o sistema pode ser encontrado (medido) por um observador em qualquer um destes estados, cada um com uma dada probabilidade. Se o destinatário (ou malfeitor) medir o sistema no estado, digamos  $|a\rangle$ , isto não quer dizer que este tenha sido realmente enviado neste estado. Isto significa apenas que o estado  $|a\rangle$  faz parte do conjunto dos estados permitidos para o sistema. Este facto, em conjunto com o facto de que uma medição altera o estado do sistema, permite efectuar a transmissão da chave pelo canal público sem risco de perder a privacidade da informação encriptada com este tipo de chave. Um exemplo concreto vai-nos ajudar de perceber como isto funciona.

Imaginemos que Alice envia ao Bob a chave privada em forma de fótons polarizados, mas polarizados aleatoriamente. Por exemplo, Alice pode usar quatro filtros polarizadores (ver a figura). Quando quer transmitir ao Bob um bit com valor igual a 1, escolhe ao acaso ou o polarizador vertical, ou o polarizador aos  $-45^\circ$ . Quando quer transmitir bit 0, escolhe ou o polarizador horizontal, ou o polarizador aos  $45^\circ$ . O Bob tem ao seu dispor apenas dois filtros: um filtro  $+$  que deixa passar aos



fótons polarizados na horizontal ou na vertical (e deste modo pode transmitir os bits 0 e 1), e o outro  $\times$ , que deixa passar aos fótons polarizados aos  $\pm 45^\circ$  (este também inclui bits 0 e 1). Além disso, Bob dispõe de dois detectores de fótons para distinguir entre os fótons com polarizações diferentes que passaram os seus filtros.

Ao transmitir a informação, Alice regista os filtros que utiliza para transmissão de cada bit (embora a sequência destes seja aleatória); Bob regista também os seus filtros e ainda o resultado das medições. Depois, Bob informa Alice quais foram os filtros que utilizou para registar cada bit, mas não os resultados das medidas. Alice responde, quais desses filtros são compatíveis com os dela e quais não são, mas não diz (obviamente!) quais foram as polarizações que usou. Os filtros compatíveis são aqueles que resultam numa informação correcta com 100% de probabilidade. Por exemplo, se Alice usou a polarização vertical, e na recepção foi utilizado o filtro  $+$ , Bob necessariamente mede 1 (ver a figura) e quando Alice confirma que este era o filtro correcto, Bob sabe com uma certeza absoluta que o bit transmitido é 1. Se para o mesmo fóton Bob utilizasse filtro  $\times$ , o resultado seria ou 1 (i.e., o fóton passa o filtro) ou 0 (não passa), com uma probabilidade de 50% para cada das duas possibilidades. Portanto, o facto de medir 1 não quer dizer que foi transmitido mesmo 1. Neste caso, Alice diz ao Bob que os seus filtros são incompatíveis e o bit respectivo é ignorado porque traz informação ambígua (na figura, os resultados que trazem informação certa estão a verde e os ambíguos são marcados com cor vermelha).

Repare, que todas as comunicações ocorrem através de uma linha pública. Mas, não há perigo se esta estiver sob escuta de terceiros. Por exemplo, Bob diz abertamente que para medição do bit número  $n$  utilizou o filtro  $+$  e Alice responde que este está

correcto. Quem está a escutar a conversa não pode concluir nada porque o filtro  $\oplus$  deixa passar tanto fótons polarizados na vertical (correspondem a 1), como fótons polarizados na horizontal (0) com probabilidades iguais. O mesmo também se aplica ao filtro  $\otimes$ .

Depois de verificar deste modo todos os bits recebidos, Bob fica apenas com uma parte da informação transmitida pela Alice (50% em média), mas sabe que tem informação correcta (i.e. aquela que foi realmente transmitida pela Alice). Alice, por seu lado, sabe qual parte da informação foi recebida correctamente. Ambos possuem agora a mesma sequência de bits e podem utilizá-la como chave na sua comunicação, que a partir deste momento se faz de modo tradicional, com chave privada.

Neste esquema, a comunicação quântica serviu para gerar a chave privada através da linha pública. Esta parte funciona só com fótons únicos, mas a partir do momento em que a chave está criada, a transmissão pode ser feita como se faz hoje, i.e com centenas e milhares de fótons. A mesma fibra óptica, em princípio, pode ser utilizada para ambas as partes da comunicação.

Pode-se perguntar: se o malfeitor interceptar a mensagem na fase de geração da chave e depois a reemitir outra vez, como é que vamos saber se a mensagem foi lida? De facto, o malfeitor tem aqui um problema: detectou uma sequência de bits mas não sabe quais destes estão certos porque não sabe qual foi a sequência dos filtros que Alice utilizou na transmissão. Então, ao reemitir os bits, não só a sequência destes estará errada, mas também a sequência dos filtros que o malfeitor utilizar não será a mesma e, portanto, o resultado das medições do Bob será diferente do caso da transmissão directa.

Resumindo, podemos designar os seguintes aspectos da física quântica utilizados na encriptação quântica:

1. Se para um sistema quântico existe uma série dos estados permitidos, qualquer sobreposição destes estados é permitida, também.
2. Uma medição de um sistema quântico, que se encontra numa sobreposição dos estados, pode resultar em qualquer destes estados com uma certa probabilidade (por outras palavras, o sistema assume um certo estado no instante em que se efectua a medição).
3. Uma medição de um sistema quântico resulta na alteração do estado do sistema.