

MECÂNICA QUÂNTICA

A. F. R. de Toledo Piza

Versão A(2K+1)

Prefácio.

Este livro foi concebido, e assim utilizado em diversas versões preliminares, como base para o curso de Mecânica Quântica, com duração de dois semestres, oferecido regularmente a estudantes de pós graduação do Instituto de Física da Universidade de São Paulo. A forma sob a qual o seu conteúdo é apresentado aqui resultou de um processo evolutivo paralelo à repetida ministração desse curso em duas épocas distintas, a primeira no final da década de 70 e a segunda cobrindo a última virada de século, que assinala ao mesmo tempo o centenário do trabalho de Planck que levou finalmente à teoria quântica, na forma em que ela é entendida hoje. O texto não exige, a rigor, familiaridade prévia da parte dos estudantes com a mecânica quântica, além do que é parte integrante de currículos normais de graduação em física e objeto de disciplinas usualmente identificadas como “Física Moderna”. O objetivo dos primeiros três capítulos é, de fato, construir o arcabouço formal da teoria quântica em termos tanto quanto possível gerais e definitivos, embora desimpedidos da discussão mais pormenorizada de sutilezas técnicas que possam obstruir a percepção do conjunto pelo estudante. Os três capítulos seguintes tratam do desenvolvimento das técnicas de aproximação básicas utilizadas nas diversas aplicações da teoria, e de sua ilustração sobretudo aos problemas clássicos da estrutura atômica. Esses seis capítulos que integram a primeira parte do texto correspondem ao primeiro semestre do curso de pós graduação ministrado no Instituto de Física da USP.

Elementos de natureza histórica são utilizados para dar perspectiva e realismo aos assuntos tratados, sem preocupações maiores, no entanto, com questões de rigor histórico. Observações apresentadas em caracteres reduzidos em diversas alturas do texto assinalam a eventual necessidade ou cabimento de recursos técnicos adicionais, sem que a sua omissão deva prejudicar a compreensão do texto básico.

Os cinco capítulos que integram a segunda parte do texto correspondem ao segundo semestre do curso. Eles tratam de tópicos também de uso geral porém mais dependentes de um maior aparato técnico, como sistemas envolvendo partículas idênticas e o seu tratamento em termos de técnicas de “segunda quantização”, o tratamento perturbativo de processos eletromagnéticos no contexto atômico utilizando uma versão quantizada do campo eletromagnético livre, e a teoria quântica formal do espalhamento de potencial. O último capítulo consiste numa introdução à

mecânica quântica relativística, que se orienta no final para a obtenção do espectro discreto de uma partícula de Dirac em um campo externo Coulombiano. Essa orientação visa, em particular, evitar superposição com a parte introdutória de cursos oferecidos sobre a teoria quântica de campos, que tratam a equação de Dirac em termos de segunda quantização.

Cada um dos capítulos é acompanhado de uma coleção de problemas que variam desde simples aplicações de resultados e procedimentos discutidos no texto até a elaboração de alguns tópicos mais específicos, desde que com interesse e aplicação geral. As referências, seja as incluídas ao longo do texto, seja as de caráter geral listadas ao final de cada um dos capítulos, não pretendem ser exaustivas em qualquer sentido, mas apenas subsidiar o contacto tanto com os trabalhos originais como com apresentações alternativas ou diversamente concebidas dos tópicos tratados.

O prolongado processo de elaboração do texto aqui apresentado se beneficiou enormemente de contribuições e comentários de um grande número de pessoas, tanto estudantes como colegas de profissão, várias delas tendo inclusive transitado da primeira para a segunda dessas situações.

São Paulo, 23 de fevereiro de 2001.

A. F. R. de Toledo Piza.

Sumário

I	Parte 1	1
1	Dificuldades com as teorias clássicas e a solução quântica.	3
1.1	Radiação de corpo negro: contexto clássico e a hipótese de Planck.	5
1.2	Trabalhos de Einstein e Bohr.	11
1.2.1	A “velha Mecânica Quântica” do átomo de hidrogênio: frequências.	13
1.2.2	Intensidades.	17
1.2.3	Einstein, os princípios de Bohr e a radiação de corpo negro.	17
1.3	A nova mecânica de Heisenberg.	20
1.3.1	A teoria de Heisenberg passada a limpo: as matrizes de Born e Jordan. .	26
1.3.2	O oscilador harmônico unidimensional segundo a mecânica das matrizes.	30
1.4	Opções determinantes e decorrências da teoria de Heisenberg, Born e Jordan. . .	33
2	Estrutura algébrica da teoria de Heisenberg. Postulados interpretativos.	41
2.1	Estrutura algébrica da teoria de Heisenberg: discussão heurística.	42
2.2	Estrutura formal e interpretação da teoria quântica.	46
2.2.1	Espaços de Hilbert.	47
2.2.2	Operadores lineares.	51
2.2.3	Espectro de operadores lineares e a interpretação da Mecânica Quântica.	57
2.2.4	Compatibilidade de observáveis e relações de incerteza.	68
2.2.5	Muitos graus de liberdade, espaços produto e sistemas compostos.	74
2.3	Avatares da dinâmica.	77
2.3.1	O operador de evolução e as descrições de Heisenberg, de Schrödinger e “de interação”.	78
2.3.2	Extensão para hamiltonianos que dependem explicitamente do tempo. . .	82

2.3.3	Relações de incerteza energia-tempo.	85
2.4	A “redução do pacote de ondas”.	86
2.4.1	Medidas reprodutíveis, o “efeito Zeno” e a redução do pacote de ondas. . .	90
3	Sistemas quânticos simples.	107
3.1	O oscilador harmônico.	107
3.1.1	Espectro e autovetores do hamiltoniano do oscilador.	108
3.1.2	Dinâmica do oscilador forçado.	112
3.2	Sistemas de dois níveis.	115
3.2.1	Estados estacionários e o operador de evolução.	117
3.2.2	“Tunelamento” e interferência de estados estacionários.	120
3.3	A representação de Schrödinger.	123
3.3.1	Propriedades gerais da dinâmica de uma partícula em um potencial externo.	127
3.3.2	Funções de onda estacionárias para o oscilador harmônico.	131
3.3.3	Comportamento quântico de uma partícula livre.	136
3.3.4	Partícula sujeita a potencial externo em uma dimensão.	152
4	Métodos de aproximação.	171
4.1	Métodos perturbativos.	173
4.1.1	Perturbação de estados não degenerados em espectros discretos.	174
4.1.2	Perturbação de um estado degenerado ou quase-degenerado.	186
4.1.3	Perturbação da equação de movimento.	192
4.1.4	Transições num espectro contínuo ou quase-contínuo: a “regra áurea” de Fermi.	202
4.2	Formulação variacional do problema de autovalores para espectros discretos. . .	212
4.3	Aproximação semi-clássica para a função de onda.	215
4.3.1	A aproximação de Wentzel, Kramers e Brillouin (WKB)	218
4.3.2	Tratamento de regiões perigosas: fórmulas de conexão.	223
4.3.3	Uso das fórmulas de conexão: penetração de barreiras de potencial e regras de quantização de Bohr.	231
5	Momento angular, potenciais centrais.	247
5.1	Momento angular na dinâmica de uma partícula sujeita a um potencial central.	250
5.1.1	Autovalores e autovetores do momento angular.	252
5.1.2	Partícula livre com momento angular bem definido.	254

5.1.3	Poço de potencial quadrado em três dimensões.	258
5.1.4	Espectro contínuo de um potencial central de curto alcance. Defasagens.	261
5.1.5	Condições de contorno, amplitudes e seção de choque de espalhamento.	263
5.1.6	Espalhamento a muito baixas energias. Expansão de alcance efetivo.	268
5.2	Problema de dois corpos com forças centrais. O átomo de hidrogênio.	272
5.2.1	Correlações de dois corpos e raios de estados ligados.	278
5.2.2	Propriedades empíricas do espectro do átomo de hidrogênio e insuficiência do modelo simples.	279
6	Estrutura fina e campos externos no átomo de hidrogênio.	285
6.1	Realização explícita das variáveis de spin. Álgebra das matrizes de Pauli.	286
6.1.1	Direção de polarização do vetor de estado geral de spin 1/2.	289
6.1.2	Matrizes de Pauli como base de operadores no espaço de spin.	291
6.2	Estados de uma partícula com spin 1/2.	292
6.2.1	Momento angular de uma partícula com spin 1/2. Adição de momentos angulares.	294
6.3	Momentos magnéticos e interações dependentes do spin.	300
6.3.1	Interação spin-órbita e estrutura fina.	300
6.3.2	Interação spin-spin e estrutura hiperfina.	309
6.4	Campos eletromagnéticos externos. Efeitos Stark e Zeeman no hidrogênio.	313
6.4.1	Efeito Stark no hidrogênio.	315
6.4.2	Efeito Zeeman no hidrogênio.	331
II	Parte 2	347
7	Sistemas de partículas idênticas.	349
7.1	Férmions e bósons.	353
7.1.1	Sistemas de muitas partículas idênticas.	356
7.1.2	Exemplos simples. Indistinguibilidade quântica e clássica.	359
7.2	A tabela periódica. Aproximações autoconsistentes para a estrutura atômica.	368
7.2.1	Potenciais médios auto-consistentes. Aproximação de Thomas-Fermi.	372
7.2.2	Aproximações de Hartree e de Hartree-Fock.	376
7.3	Métodos de “segunda quantização” para sistemas de muitas partículas idênticas.	382

7.3.1	Estados no espaço de Fock e funções de onda de muitas partículas idênticas.	386
7.3.2	Variáveis dinâmicas no espaço de Fock.	392
7.3.3	Exemplos de aplicação: gás de Fermi livre e com forças de dois corpos.	399
8	Interação da matéria com a radiação.	415
8.1	O campo de radiação livre.	415
8.1.1	Quantização do campo de radiação livre.	418
8.1.2	Energia, momento e momento angular do campo livre.	421
8.2	Interação com um sistema de partículas carregadas.	429
8.2.1	Espalhamento Thomson.	432
8.2.2	Radiação de freiamento (Bremsstrahlung).	443
8.2.3	Transições radiativas em sistemas ligados.	451
8.2.4	Análise multipolar das transições radiativas.	461
9	Transformações, simetrias e leis de conservação.	469
9.1	Grupos de transformações.	470
9.1.1	Grupos contínuos, geradores.	471
9.1.2	Representações.	472
9.2	Translações espaciais e momento.	477
9.3	Rotações, momento angular e spin.	479
9.3.1	Spin.	483
9.3.2	Representação de rotações gerais em termos de ângulos de Euler.	488
9.3.3	Matrizes de rotação, tensores esféricos, teorema de Wigner-Eckart.	491
9.4	Isospin.	498
9.5	Paridade e paridade intrínseca.	507
9.6	Invariança por translações temporais e a lei dinâmica de evolução.	510
9.7	Inversão temporal.	513
9.7.1	Sistemas com spin.	517
9.7.2	Invariança sob inversão temporal.	519
10	Espalhamento por um potencial de curto alcance.	529
10.1	A equação de Lippmann-Schwinger.	530
10.1.1	A expansão de Born.	532
10.2	Descrição dependente do tempo de um processo de espalhamento por um potencial.	534

10.2.1	Escalas típicas em processos de espalhamento.	535
10.2.2	Tratamento dependente do tempo do espalhamento de um pacote de ondas por um potencial de curto alcance.	539
10.2.3	Conservação de norma e o teorema óptico.	547
10.3	Estados de espalhamento no espaço de momentos.	549
10.4	Equação de Lippmann-Schwinger formal e operador de transição.	555
10.4.1	A “fórmula de dois potenciais” de Gell-Mann e Goldberger.	560
10.5	Operador de espalhamento S . Propriedades formais.	563
10.6	Diagonalização do operador de espalhamento para potenciais invariantes sob rotações. Ondas parciais e defasagens.	566
10.6.1	Ingredientes cinemáticos.	567
10.6.2	Transformação de S e $T(k)$ para a base esférica.	570
10.6.3	Reconstrução da amplitude de espalhamento e da forma assintótica da função de onda.	571
11	Equações relativísticas.	579
11.1	Sumário da cinemática da relatividade restrita.	580
11.1.1	O eletromagnetismo de Maxwell como exemplo de “boa” teoria.	585
11.2	A equação de Klein-Gordon.	587
11.3	A equação de Dirac.	592
11.3.1	Forma hamiltoniana e soluções da equação de Dirac livre.	594
11.3.2	Propriedades algébricas das matrizes gama.	601
11.3.3	Transformações de Lorentz e rotações.	602
11.3.4	Inversão espacial.	609
11.3.5	Momento angular na equação de Dirac livre.	610
11.4	Propriedades eletromagnéticas de uma partícula de Dirac carregada.	612
11.5	Estados ligados de um elétron de Dirac em um campo coulombiano externo, atrativo.	616
11.5.1	Simetrias e funções angulares.	617
11.5.2	Equações radiais acopladas.	619
11.5.3	Soluções para o espectro discreto.	622

- II -

Meu duvidar é da realidade sensível aparente - talvez só um escamoteio das percepções. Porém, procuro cumprir. Deveres de fundamento a vida, empírico modo, ensina: disciplina e paciência.

...

- Tem-se que redigir um abreviado de tudo.

...

Tudo se finge, primeiro; germina autêntico é depois. Um escrito, será que basta? Meu duvidar é uma petição de mais certeza.

(J. Guimarães Rosa, 4^o prefácio de Tutaméia).

Parte I

Capítulo 1

Dificuldades com as teorias clássicas e a solução quântica.

Como aparentemente voltou a ocorrer no fim do século XX, a física do fim do século XIX chegou a sugerir aos pesquisadores da época a iminência de uma grande síntese (dizemos hoje uma “teoria final”) que seria organizada, daquela vez, em torno da mecânica de Newton e da eletrodinâmica de Maxwell. O que efetivamente se viu a seguir, no entanto, foi um dos processos mais extensos e radicais de renovação conceitual de toda a história da física. De fato, esse processo não se limitou aos mecanismos dinâmicos envolvidos nas teorias clássicas, mas exigiu uma revisão da própria base *cinemática* envolvida na sua formulação. Em outras palavras, não só a estruturação dos fenômenos proposta pelas teorias clássicas se mostrou inadequada, mas até mesmo os termos usados na própria identificação e caracterização dos fenômenos tiveram que ser revistos. Não há então exagero algum em descrever o que aconteceu durante as duas ou três primeiras décadas do século XX como uma redefinição bastante radical daquilo que se entende por realidade física.

Esse processo radical de reconstrução da realidade pode ser visto como tendo se dado ao longo de duas frentes, inicialmente independentes. Por um lado, a teoria especial da relatividade reformulou a imagem do espaço-tempo que tinha sido admitida explicitamente pelo menos desde a física de Galileu. Em termos históricos, embora esquemáticos, pode-se dizer que essa reformulação foi necessária para salvar o princípio de relatividade, formulado por Galileu em termos do espaço-tempo “clássico”, de certas complicações inerentes à descrição de fenômenos eletrodinâmicos. Além da re-avaliação conceitual geral, ligada a ela, que atingiu também a eletrodinâmica de Maxwell, ela acarretou, para a mecânica de Newton, mudanças essenciais na própria estrutura da teoria. Essas mudanças alteram profundamente, em particular, resultados quantitativos no caso de velocidades altas (próximas da velocidade da luz).

A mecânica quântica, objeto deste curso, foi a segunda frente de revisão das teorias clássicas.

Como será visto adiante, ela envolveu também uma redefinição da cinemática que, sob vários aspectos, foi ainda mais surpreendente que a da teoria especial da relatividade. De fato, esta última envolveu basicamente uma modificação das propriedades geométricas admitidas anteriormente para o espaço-tempo, que se traduz, por exemplo, na relativização do conceito de simultaneidade, devido à relação existente entre as escalas de tempo de dois referenciais que se movem um com relação ao outro. A teoria quântica, por outro lado, atinge até mesmo o sentido de objetividade que fora emprestado aos sistemas físicos pelas descrições clássicas, e que foi integralmente mantido uma vez feitas as alterações prescritas pela teoria especial da relatividade. Na teoria quântica, o que pode ser chamado “natureza objetiva” de um sistema físico passa a transcender o resultado de possíveis observações feitas sobre ele. A descrição quântica é de tal natureza que dá margem, em particular, a que um mesmo sistema tenha comportamentos “contraditórios” em situações observacionais diferentes, como são, por exemplo, manifestações de caráter ondulatório e manifestações de caráter corpuscular. Dessa forma pode-se dizer que a teoria quântica implica numa relativização da “natureza” do sistema físico, através da transformação do caráter de suas manifestações com uma mudança de dispositivo de observação.

Não é possível, portanto, “deduzir” a mecânica quântica de noções ou esquemas anteriores, nem tão pouco construí-la como alguma forma de extensão de teorias clássicas. O tema deste primeiro capítulo deve ser visto, portanto, como uma espécie de paráfrase do complicado e geralmente ilógico processo de invenção de um esquema teórico novo. Apesar de uma certa abundância de referências históricas, ele não pretende ser história. Em particular, do ponto de vista da história, é inadmissível a omissão das ideias e resultados devidos a de Broglie e Schrödinger que levaram à construção alternativa da mecânica quântica que levou ao que foi inicialmente chamado mecânica ondulatória, por exemplo. No entanto, o caminho seguido paralela e independentemente por Heisenberg parece se prestar melhor ao propósito de enfatizar o ponto essencial em que se dá a ruptura radical com os esquemas clássicos, além de mostrar muito claramente o que se ganha (e o que eventualmente se perde) com uma tal ruptura. Um panorama historicamente mais isento pode ser encontrado em referências listadas no fim do capítulo, ou mesmo, e preferencialmente, nos trabalhos originais citados nessas referências ou ao longo do texto. O propósito das referências históricas usadas neste capítulo é na realidade tratar de indicar *como* e *porque* posições teóricas, tão desconfortáveis de um certo ponto de vista, vieram a se consolidar da forma que o fizeram. Para adiantar uma resposta talvez genérica demais mas pelo menos posta em poucas palavras, a razão para isso é que elas dão conta, e dão conta de uma forma extremamente coerente, de uma vasta classe de fenômenos que parecem de outro modo incompreensíveis. Ainda mais, é precisamente devido a esse enorme poder sintetizador que o muito que a própria mecânica quântica possa ter ainda hoje de “incompreensível” pareça antes um desafio que uma inconveniência.

1.1 Radiação de corpo negro: contexto clássico e a hipótese de Planck.

Em 1900, supondo que a emissão de energia sob a forma de radiação eletromagnética de frequência ν por um corpo negro estivesse associada a porções discretas (“quanta”) de magnitude $h\nu$ (isto é, de magnitude proporcional à frequência), Max Planck obteve para a distribuição espectral da radiação de corpo negro a expressão

$$\rho_\nu = \frac{\alpha \nu^3}{\exp \frac{h\nu}{kT} - 1}, \quad (1.1)$$

onde $\rho_\nu d\nu$ é a energia da radiação, por unidade de volume, com frequência compreendida entre ν e $\nu + d\nu$ quando o corpo negro está à temperatura T , que concorda extremamente bem com resultados experimentais para uma escolha conveniente do valor da constante de proporcionalidade h .

A hipótese crucial de Planck contradiz frontalmente o que se poderia esperar a partir da mecânica e da eletrodinâmica clássicas. No entanto, as consequências dessas teorias clássicas contradizem frontalmente fatos empíricos (no caso, os resultados experimentais para a distribuição espectral ρ_ν), e a natureza da contradição foi suficientemente grave para conduzir à invenção de uma estrutura teórica nova. O domínio de ação de tendências conservadoras teve que se retrair, nesse processo, à exigência de que essa nova estrutura se acomodasse à condição de conduzir a resultados de alguma forma equivalentes aos das teorias clássicas nos contextos em que estes se tivessem mostrado adequados.

O trabalho de Planck não realizou esse programa. Ele apenas teve o mérito de fornecer o germe inicial que resultou decisivo para a elaboração da nova teoria. Esta só emergiu de uma forma razoavelmente coerente dos trabalhos de Heisenberg e de Schrödinger, mais de duas décadas depois do passo inicial dado por Planck.

Da teoria eletromagnética e de princípios gerais da termodinâmica decorrem dois resultados gerais para a radiação eletromagnética contida numa cavidade, mantida à temperatura T :

a) A energia por unidade de volume depende apenas de T e é proporcional a T^4 (lei de Stefan-Boltzman).

b) A distribuição espectral da radiação nessas condições deve ter a forma

$$\rho_\nu = \nu^3 f\left(\frac{\nu}{T}\right),$$

um resultado conhecido como lei de Wien. Ele pode ser obtido de considerações termodinâmicas gerais e do efeito Doppler que altera a frequência da radiação num processo adiabático de compressão ou expansão. A forma da função $f\left(\frac{\nu}{T}\right)$ não é fixada por essas considerações gerais,

mas pode ser obtida a partir de um modelo para a interação entre a radiação e a cavidade; a generalidade do argumento termodinâmico mostra, no entanto, que a forma de $f\left(\frac{\nu}{T}\right)$ não deve depender do particular modelo considerado. Deve-se esperar, então, que o resultado obtido para um modelo específico qualquer tenha validade geral.

Este último resultado consiste, na realidade, da combinação de duas leis obtidas por Wien em 1893, utilizando comprimentos de onda em vez de frequências. Comparando a radiação antes e depois de um processo de compressão (ou de expansão) adiabática, Wien mostrou que a) $\lambda T = \lambda' T'$ e b) $I(\lambda, T)/T^5 = I(\lambda', T')/T'^5$ donde se pode concluir que

$$I(\lambda, T) = T^5 F(\lambda T),$$

sendo λ o comprimento de onda da radiação, $I(\lambda, T)d\lambda$ a energia emitida pelo corpo negro no intervalo de frequências entre λ e $\lambda + d\lambda$ e $F(\lambda T)$ uma função universal. A importância desses resultados vem de que eles correspondem à *previsão*, experimentalmente verificável *independentemente do conhecimento da forma de $F(\lambda T)$* , de uma propriedade de escalonamento das densidades espectrais correspondentes a diferentes temperaturas. De fato, essa propriedade pode ser verificada graficando a quantidade experimentalmente mensurável $I(\lambda, T)/T^5$ como função de λT para diferentes temperaturas. Sendo ela correta, os resultados correspondentes a diferentes temperaturas caem sobre uma mesma curva, que fornece então uma determinação “experimental” de $F(\lambda T)$. É claro que uma análise equivalente pode ser feita em termos de temperatura e frequência para obter experimentalmente a forma de $f\left(\frac{\nu}{T}\right)$.

Um ponto que pode ser curioso observar é o de que a combinação λT não é, evidentemente, adimensional, o que torna algo embaraçosa a sua aparição como argumento de uma função supostamente *universal*. No entanto, o embaraço desaparece caso existam constantes, também *universais*, que permitam exprimir as dimensões de λT (isto é, que permitam escrever uma expressão adimensional, envolvendo λT além de constantes universais). Hoje isso efetivamente é feito escrevendo $\lambda T \times k/hc = kT/h\nu$, sendo k a constante de Boltzmann, h a constante de Planck e c a velocidade da luz, estas últimas evidentemente inexistentes como constantes universais à época destes acontecimentos. Com o benefício dos desenvolvimentos posteriores, podemos então hoje reconhecer, já na lei de Wien, uma indicação da existência de uma escala *universal* para quantidades com dimensão de ação.

Um modelo simples para a interação de radiação de frequência ν com a cavidade consiste em representar a cavidade, nessa interação, por um oscilador harmônico carregado (carga e e massa m), de frequência ν . A energia radiada por segundo é dada pela expressão conhecida como fórmula de Larmor¹, que no caso do oscilador leva a

¹A referência original (J. J. Larmor, *Phil. Mag.* 1897, pag. 512) é citada em A. Sommerfeld, *Electrodynamics*, Academic Press, N.Y. 1952, pag. 151. Esta fórmula é deduzida na maioria dos textos sobre eletrodinâmica clássica, e.g. J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, N.Y. 1975 (2nd. edition).

$$\delta\epsilon_{rad} = \frac{2}{3} \frac{e^2}{c^3} \langle \ddot{q}^2 \rangle = \frac{2}{3} \frac{e^2}{mc^3} (2\pi\nu)^2 \langle \epsilon \rangle$$

onde q e ϵ representam, respectivamente, a coordenada do oscilador e a sua energia, e os colchetes indicam valores médios. Por outro lado, a taxa média de absorção de energia pelo oscilador, considerando o campo elétrico da radiação de frequência ν que age sobre a carga e , é²

$$\delta\epsilon_{abs} = \frac{\pi e^2}{3m} \rho_\nu.$$

Se a radiação de frequência ν está em equilíbrio com o oscilador $\delta\epsilon_{rad} = \delta\epsilon_{abs}$, o que dá para a densidade espectral ρ_ν

$$\rho_\nu = \frac{8\pi\nu^2}{c^3} \langle \epsilon \rangle.$$

Deve-se notar que essa expressão independe de características específicas do oscilador (como carga e massa), envolvendo *apenas* a sua energia média. Tratando-se de um sistema (oscilador mais radiação) em equilíbrio térmico à temperatura T , é fácil relacionar $\langle \epsilon \rangle$ com T através de procedimentos usuais da mecânica estatística. Classicamente, uma coleção de osciladores se distribui em energia ϵ , à temperatura T , com uma densidade de probabilidade de Boltzmann³

$$p(\epsilon) = \frac{1}{Z} e^{-\frac{\epsilon}{kT}}$$

com a constante de normalização

$$Z = \int_0^\infty e^{-\frac{\epsilon}{kT}} d\epsilon = kT$$

de modo que

$$\langle \epsilon \rangle = \int_0^\infty \epsilon p(\epsilon) d\epsilon = -\frac{1}{Z} \frac{d}{d(\frac{1}{kT})} \int_0^\infty e^{-\frac{\epsilon}{kT}} d\epsilon = kT.$$

O que se obtém dessa forma para a distribuição espectral da radiação é portanto

²v. e.g. J. D. Jackson, *op. cit.*

³Para uma introdução a esta fórmula e às suas inúmeras aplicações, inclusive à teoria quântica, v. e.g. R. P. Feynman, R. B. Leighton e M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley Publ. Co., muitas edições, Lectures #40-42.

$$\rho_\nu = \frac{8\pi\nu^2}{c^3} kT$$

que corresponde, em termos da lei de Wien, à função

$$f\left(\frac{\nu}{T}\right) = \frac{8\pi k}{c^3} \frac{T}{\nu}.$$

Esse é o resultado clássico para ρ_ν , conhecido como lei de Rayleigh-Jeans. Ele descreve corretamente o comportamento observado da distribuição espectral da radiação de corpo negro no limite de baixas frequências ($\nu \rightarrow 0$), mas a proporcionalidade ao quadrado da frequência é claramente catastrófica no limite de frequências altas, no qual a densidade espectral observada tende, de fato, a zero, o que em particular é necessário até mesmo para que a potência total irradiada a uma temperatura T dada (que é proporcional à densidade espectral ρ_ν integrada sobre ν) seja finita. Trata-se ainda de uma catástrofe com características embaraçosas de “robustez”, no sentido de que ela aparece como consequência de resultados bastante gerais da física clássica (mais especificamente da mecânica, eletrodinâmica e termodinâmica). Desse modo, não parece razoável pensar em evitá-la sem questionar de alguma forma esses mesmos resultados.

A proposta de Planck consistiu em tomar liberdades com relação à mecânica do oscilador harmônico. Como é bem sabido, a energia de um tal oscilador está associada ao quadrado da amplitude das oscilações, podendo assumir portanto qualquer valor positivo (ou zero), o que é efetivamente incorporado no cálculo da energia média do oscilador, $\langle \epsilon \rangle$. O resultado obtido por Planck, no entanto, corresponde a modificar o cálculo anterior através da suposição de que a energia de um oscilador de frequência ν assume *apenas* valores que são múltiplos inteiros de uma quantidade elementar (ou “quantum”) ϵ_ν , que pode depender da frequência do oscilador. Dessa forma, os valores possíveis da energia de um oscilador de frequência ν são $n\epsilon_\nu$. Todos os demais ingredientes do cálculo da distribuição espectral são mantidos sem qualquer outra alteração. A distribuição de probabilidades das diferentes energias à temperatura T é então agora discreta e dada por

$$P_n = \frac{1}{Z} e^{-\frac{n\epsilon_\nu}{kT}}$$

com a constante de normalização obtida agora em termos da soma de uma série geométrica

$$Z = \sum_{n=0}^{\infty} e^{-\frac{n\epsilon_\nu}{kT}} = \frac{1}{1 - \exp\left(-\frac{\epsilon_\nu}{kT}\right)}$$

de modo que a energia média aparece como

$$\langle \epsilon \rangle = \frac{1}{Z} \sum_{n=0}^{\infty} n \epsilon_{\nu} e^{-\frac{n \epsilon_{\nu}}{kT}} = -\frac{1}{Z} \frac{d}{d \left(\frac{1}{kT} \right)} \sum_{n=0}^{\infty} e^{-\frac{n \epsilon_{\nu}}{kT}} = \frac{\epsilon_{\nu}}{\exp \left(\frac{\epsilon_{\nu}}{kT} \right) - 1}.$$

Usando esse resultado, os mesmos passos que levavam à lei de Rayleigh-Jeans dão agora

$$\rho_{\nu} = \frac{8\pi\nu^2}{c^3} \frac{\epsilon_{\nu}}{\exp \left(\frac{\epsilon_{\nu}}{kT} \right) - 1}. \quad (1.2)$$

Para que esse resultado satisfaça a lei de Wien, é preciso que o “quantum” ϵ_{ν} seja proporcional à frequência, isto é, $\epsilon_{\nu} = h\nu$. Dessa forma, de fato, a densidade espectral assume a forma

$$\rho_{\nu} = \frac{8\pi h \nu^3}{c^3} \frac{1}{\exp \left(\frac{h\nu}{kT} \right) - 1}$$

que corresponde à expressão (1.1) com $\alpha = 8\pi h/c^3$. Em termos da lei de Wien, esta densidade espectral corresponde à função

$$f \left(\frac{\nu}{T} \right) = \frac{8\pi h}{c^3} \frac{1}{\exp \left(\frac{h\nu}{kT} \right) - 1}.$$

Para frequências baixas (no sentido de que $h\nu \ll kT$) recupera-se a lei de Rayleigh-Jeans através de uma expansão em $h\nu/kT$. Por outro lado, para frequências altas (isto é, $h\nu \gg kT$), a densidade espectral se comporta como

$$\rho_{\nu} \rightarrow \frac{8\pi h \nu^3}{c^3} e^{-\frac{h\nu}{kT}} = \alpha \nu^3 e^{-\frac{h\nu}{kT}} \quad (h\nu \gg kT), \quad (1.3)$$

que tende devidamente a zero, reproduzindo uma expressão proposta em 1896 por Wien que se mostrara adequada no limite de frequências *altas*, embora falhando no extremo de frequências baixas.

A expressão obtida por Planck, portanto, “interpola” com sucesso entre a distribuição de Rayleigh-Jeans (no extremo de frequências baixas, onde ela funciona) e a expressão de Wien, empiricamente adequada para frequências altas. Mas ela é obtida lançando mão de uma hipótese que efetivamente introduz vínculos entre a energia e a frequência de um oscilador harmônico que simplesmente *não cabem* no contexto da teoria clássica: os valores $nh\nu$ da energia de um oscilador de frequência ν , sendo h uma constante com dimensões de ação (energia \times tempo), são “especiais”. Vale a pena notar que uma vinculação entre energia e frequência na realidade

aparece também no trabalho de 1896 de Wien (*Ann. Physik* **58**, 662), sob a forma da suposição de que a frequência da radiação emitida por um oscilador fosse proporcional à sua energia cinética. Devido a essa e outras hipóteses envolvidas na sua dedução, a fórmula de Wien não chegou a ser vista como teoricamente bem fundamentada, sendo freqüentemente citada como “empírica”, em eventual homenagem à sua porção de sucesso. A fórmula de Planck, por outro lado, não só consegue sucesso total na descrição da distribuição espectral da radiação de corpo negro, mas o faz com apoio substancial em bases quase tão robustas quanto as da fórmula de Rayleigh-Jeans, o desvio correndo completamente por conta da hipótese dos “quanta”. Essa hipótese tem ainda a característica importante de manter contacto com o resultado “completamente clássico” (de Rayleigh-Jeans) através de um processo bem definido de limite em que $h \rightarrow 0$. De fato, a fórmula de Rayleigh-Jeans é recuperada para $h\nu \ll kT$ o que, para qualquer temperatura não nula, vale *sempre* nesse limite.

O relato acima incorre em um número enorme de licenças com relação ao possa ser a verdade histórica do ocorrido, e para perceber isso basta notar algumas datas: a lei de Rayleigh-Jeans foi obtida por Rayleigh em 1900 (*Phil. Mag.* **49**, 539) mas teve sua robustez mostrada apenas “tardamente” por Jeans em 1909 (*Phil. Mag.* **17**, 229,273; *ibid* **18**, 209). Os trabalhos de Planck, por outro lado, apareceram em *Verh. Deut. Physik. Ges.*, **2**, 237 (1900) e em *Ann. Physik* **4**, 553 (1901). Consistentemente com o contexto daí resultante, as hipóteses efetivamente usadas por Planck parecem ter procurado ser tanto quanto possível conservadoras, não chegando em particular a discretizar completamente a energia dos osciladores harmônicos. Elas foram na realidade as seguintes: 1) Um oscilador carregado *absorve* energia *continuamente* da forma prescrita pela eletrodinâmica; 2) Um oscilador *só pode radiar* quando a sua energia é um múltiplo inteiro de ϵ_ν , e quando radia, radia *toda* a sua energia; e 3) a radiação ou não de um oscilador quando a sua energia passa por um valor $n\epsilon_\nu$ é determinada estatisticamente, sendo a razão da probabilidade de não emissão para a probabilidade de emissão proporcional à intensidade da radiação que o excita. Então, tanto a suficiência do passo que consiste em discretizar a energia dos osciladores como a mencionada robustez do resultado de Planck são historicamente fictícias, embora fisicamente corretas; a aceitação do resultado (que foi originalmente escrito sob a forma (1.1) e relacionado com o resultado de Wien para altas frequências escrito como na Eq. (1.3) *sem explicitar a forma da constante α*) pode ser tida como motivada mais por seu sucesso quantitativo que por suas eventuais virtudes estruturais. O seu desenvolvimento posterior foi na realidade bastante árduo, pois teve também que proceder ao enxugamento e à radicalização das hipóteses efetivamente utilizadas, inicialmente por demais “conservadoras”. A posição do próprio Planck com relação às suas hipóteses pode ser elucidativa a esse respeito:

É verdade que não provaremos que essa(s) hipótese(s) representa(m) a expressão única, ou sequer mesmo a mais adequada, da lei dinâmica da vibração

de osciladores. Pelo contrário, tenho como muito provável que ela(s) possa(m) ser melhorada(s) tanto na sua forma como no seu conteúdo. No entanto, não existe outro método de testar a sua admissibilidade que não seja testar as suas consequências, e enquanto não forem descobertas contradições internas ou com a experiência, enquanto não for proposta para substituí-la(s) outra hipótese mais adequada, ela(s) podem com justiça reivindicar uma certa importância (M. Planck e M. Masius, *The Theory of Heat Radiation*, McGraw-Hill, N.Y. 1914, p. 154).

Por outro lado, é curioso notar o ingrediente de natureza probabilística na hipótese 3) da formulação de Planck, de certa forma “premonitório” do que viria a ser efetivamente introduzido na *interpretação* da teoria madura que finalmente resultaria dela, quase três décadas mais tarde, até hoje objeto de grande número de discussões e controvérsias.

1.2 Trabalhos de Einstein e Bohr.

O período que se seguiu ao tratamento da radiação de corpo negro por Planck pode ser pensado como um período de sondagens destinadas a avaliar as estranhas hipóteses de quantização, tanto no que diz respeito à sua natureza (sendo nesse caso importante ter em mente que o trabalho de Planck de certa forma procura evitar a discretização completa da energia do oscilador harmônico, deixando a “quantização” numa espécie de limbo, do qual nem mesmo algum papel eventualmente desempenhado pela radiação parece poder ser excluído) quanto no que diz respeito ao seu alcance, isto é, à sua relevância em situações diferentes daquela para a qual foram formuladas.

Um passo absolutamente decisivo para explicitar a vinculação entre energia e frequência utilizada por Planck, reforçando ao mesmo tempo o seu caráter fundamental, foi dado por Einstein, em 1905, em sua teoria do efeito fotoelétrico. A idéia apresentada nesse trabalho é tão simples quanto radical e poderosa, e consiste simplesmente em notar que o fato de que a *energia* dos elétrons emitidos no efeito fotoelétrico depende da *frequência* da radiação absorvida e não de sua intensidade pode ser explicado quantitativamente com a hipótese de que a radiação eletromagnética de frequência ν se comporta como se constituída de “pacotes” indivisíveis de energia $h\nu$. Essa hipótese, então, *discretiza completamente a radiação*. Outro passo, devido também a Einstein, que contribuiu de forma muito importante para expandir o alcance das hipóteses de Planck, foi o tratamento, dado em 1907, para o comportamento dos calores específicos de sólidos a temperaturas muito baixas. Supondo que a energia de um sólido à temperatura T seja representada pela energia média de uma coleção de $3N$ “osciladores de Planck” unidimensionais com frequência ν , para os quais vale

$$\langle \epsilon \rangle = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1},$$

a energia do sólido deve ser algo como

$$\langle E \rangle \simeq 3N \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}$$

donde resulta imediatamente que

$$c_v = \frac{\partial \langle E \rangle}{\partial T} \simeq 3Nk \left(\frac{h\nu}{kT} \right)^2 \frac{e^{\frac{h\nu}{kT}}}{\left(e^{\frac{h\nu}{kT}} - 1 \right)^2}.$$

Para $h\nu \ll kT$ essa expressão reproduz o resultado “clássico” $3Nk$ conhecido como lei de Dulong e Petit, mas diverge dele para temperaturas baixas (na escala $h\nu/k$), indo a zero no limite $T \rightarrow 0$. Esse comportamento reproduz razoavelmente os valores medidos de c_v para diversos sólidos, mediante um ajuste do valor de ν para cada caso.

O fato de que a adiabaticidade, tão espertamente utilizada por Wien em sua análise geral da radiação de corpo negro, pode ter tido um papel também importante na questão da discretização completa da energia dos osciladores de Planck, é indicado claramente em uma discussão entre Lorentz e Einstein registrada nos anais da Conferência Solvay de 1911. Nesse registro se lê que

(...) O Sr. LORENTZ recorda uma conversa tida há algum tempo com o Sr. EINSTEIN, na qual discutiu-se um pêndulo simples cujo comprimento era diminuído tomando o fio entre dois dedos e deslocando-os para baixo. Supondo que inicialmente o pêndulo tivesse exatamente um elemento de energia que corresponde à frequência de suas oscilações (*isto é, energia $h\nu_0$, sendo ν_0 a frequência inicial do pêndulo*), parece que no fim da experiência sua energia será menor que o elemento que corresponde à nova frequência (*isto é, $h\nu_1$, sendo a nova frequência $\nu_1 > \nu_0$*).

Sr. EINSTEIN - Se o comprimento do pêndulo for modificado de forma infinitamente lenta a energia de oscilação permanece igual a $h\nu$ se ela era inicialmente igual a $h\nu$; ela varia proporcionalmente à frequência. O mesmo acontece para um circuito elétrico oscilante, sem resistência, e também para a radiação livre.

Sr. LORENTZ - Esse resultado é muito curioso e faz desaparecer a dificuldade (...)

A referência à “radiação livre” sugere um dos possíveis ingredientes que pode ter estado envolvido na elaboração do trabalho sobre o efeito fotoelétrico. Além disso, o argumento de Einstein contém já essencialmente o que viria a ser conhecido, depois de 1913, como a “hipótese adiabática” de Ehrenfest (que de fato atribui essa denominação a Einstein). Essa hipótese seria

mais tarde elaborada por Sommerfeld para, admitindo de forma explícita a discretização completa da energia dos osciladores de Planck, estendê-la a outros sistemas, e se tornaria (em 1917) um dos elementos cruciais invocados por Bohr (sob a denominação de *princípio da transformabilidade mecânica*) para a descrição do espectro do átomo de hidrogênio. Essa hipótese foi o ingrediente crucial na montagem do esquema *dinâmico* usado para estender a quantização dos osciladores de Planck a sistemas mais gerais.

O tratamento de Bohr do átomo de hidrogênio permitiu obter corretamente as frequências das linhas espectrais observadas, embora tenha mostrado limitações de princípio na descrição das intensidades relativas dessas linhas.

1.2.1 A “velha Mecânica Quântica” do átomo de hidrogênio: frequências.

A “geometria” do átomo de hidrogênio adotada por Bohr foi a sugerida pelos trabalhos de Rutherford envolvendo o espalhamento de partículas α por átomos mais pesados: uma carga central, pesada, distribuída com um raio da ordem de 10^{-13} cm circundada por uma distribuição muito mais extensa e leve de carga de sinal oposto. No caso do hidrogênio essas duas distribuições correspondem a apenas um próton e um elétron, respectivamente. Para descrever o comportamento dinâmico de um tal sistema, Bohr propõe dois “princípios” fundamentais que são formulados do seguinte modo:

1. Um sistema atômico pode existir, e somente pode existir permanentemente, numa série de estados que correspondem a uma série descontínua de valores de sua energia; conseqüentemente, qualquer mudança da energia do sistema, incluindo emissão e absorção de radiação, deve ocorrer através de uma *transição completa* entre dois desses estados, que serão chamados *estados estacionários* do sistema.

2. A radiação emitida ou absorvida durante a transição completa entre dois estados estacionários é “unifrequêntica” e possui uma frequência ν dada pela relação

$$E' - E'' = h\nu$$

onde h é a constante de Planck e E' e E'' são os valores da energia nos estados considerados.

O primeiro princípio declara, definitivamente, a discretização completa da energia de “sistemas atômicos”, e o segundo é fortemente reminescente da quantização da radiação proposta em 1905 por Einstein, razão pela qual ficou conhecido como “relação de Einstein-Bohr”. Ambos contrariam radicalmente os resultados da física clássica. No entanto, é curioso observar o papel

mediador essencial desempenhado por resultados puramente clássicos, através da hipótese adiabática de Ehrenfest, na escolha dos estados estacionários mencionados no primeiro princípio.

Esses estados são determinados através do seguinte procedimento: para um oscilador linear de período τ a variável de ação I é

$$I = \int_0^\tau p \dot{q} dt = 2 \int_0^\tau \frac{p^2}{2m} dt = \tau E = \frac{E}{\nu}.$$

Entendendo as hipóteses de Planck como exigindo que a energia seja $E_n = nh\nu$, segue que os valores possíveis da variável de ação para o oscilador são $I_n = nh$. Essa tradução da condição de Planck em termos da variável de ação é “útil” porque I é um *invariante adiabático* do oscilador. Se o movimento de um sistema mecânico, clássico, suposto ainda unidimensional e espacialmente limitado, é regido por uma hamiltoniana dependente de um parâmetro externo λ , quando esse parâmetro é variado muito lentamente (ou “adiabaticamente”, no sentido que a sua variação $\delta\lambda$ em um período τ do movimento é muito menor que λ), existem determinadas variáveis dinâmicas, chamadas invariantes adiabáticos, cujo valor permanece *immune* à variação de λ . Essas quantidades fazem parte do ferramental padrão mais sofisticado da mecânica clássica⁴, e na realidade estão presentes também na discussão entre Lorentz e Einstein transcrita na página 12. Aqui, a sua utilidade se deve ao fato de que permitem identificar propriedades que devem ser *iguais* em sistemas *diferentes* (isto é, com hamiltonianas diferentes), sempre que as diferenças possam ser pensadas como introduzidas de forma suficientemente gradual e lenta. No caso do oscilador, ele pode ser visto como correspondendo a um caso particular de uma classe de sistemas descritos por uma hamiltoniana dotada de um parâmetro λ cuja variação altera o termo de energia potencial, por exemplo

$$H(p, q, \lambda) = \frac{p^2}{2m} + \lambda \frac{m\omega^2}{2} q^2 + (1 - \lambda)V(q)$$

onde $V(q)$ é algum potencial anarmônico que admite também trajetórias periódicas (para garantir isso, no caso simples de sistemas unidimensionais, é suficiente exigir que o movimento seja sempre espacialmente limitado). Essa hamiltoniana descreve um oscilador harmônico de frequência angular $\omega = 2\pi\nu$ quando $\lambda = 1$, uma partícula de massa m no potencial $V(q)$ para $\lambda = 0$ e interpola entre esses dois limites para valores intermediários de λ . Como a invariância adiabática de I significa que o seu valor permanece constante se, durante o movimento, o parâmetro λ sofre uma variação *muito lenta* no sentido que

$$\frac{\tau}{\lambda} \frac{d\lambda}{dt} \ll 1$$

⁴V. e.g. L. Landau e E. M. Lifshitz, *Mécanique*, Éditions en Langues Etrangères, Moscou, 1960, §49.

ela pode ser tomada como um argumento em favor da suposição de que a condição de quantização $I_n = nh$, uma vez admitida para o oscilador harmônico, pode ser também utilizada para movimentos periódicos em potenciais unidimensionais mais gerais (isto é o que Bohr chama princípio da transformabilidade mecânica). Os *estados estacionários* do potencial $V(q)$ são os que correspondem às órbitas periódicas clássicas para as quais essa condição de quantização esteja satisfeita, isto é

$$I = \int_0^\tau p\dot{q} dt = nh. \quad (1.4)$$

Sommerfeld adotou uma generalização desse procedimento que funciona, em particular, para partículas sujeitas a um potencial central em três dimensões. Nesse caso, usando coordenadas polares (r, φ) no plano do movimento, os momentos canônicos correspondentes são, no caso de uma partícula de massa m , $p_r = m\dot{r}$ e $p_\varphi = mr^2\dot{\varphi}$, sendo este último (o momento angular) uma constante do movimento, $p_\varphi = l$. Como para estados ligados de um potencial central atrativo o movimento de cada uma dessas variáveis é periódico, existem *duas* condições de quantização, uma para cada um dos graus de liberdade, que são

$$\int_0^{\tau_r} m\dot{r}^2 dt = n_r h \quad \text{e} \quad \int_0^{\tau_\varphi} l \dot{\varphi} dt = 2\pi l = n_\varphi h$$

onde τ_r e τ_φ são, respectivamente, os períodos do movimento radial e do movimento angular. Essas foram as condições de quantização usadas por Bohr para determinar os estados estacionários do átomo de hidrogênio, em que o potencial central é $V(r) = -e^2/r$. Os estados estacionários que tem $n_r = 0$ correspondem a uma família de órbitas circulares que podem ser rotuladas pelo valor correspondente do momento angular, dado pela segunda condição de quantização como $l = n_\varphi h/2\pi$. Um cálculo simples dá para a energia⁵ dessas órbitas circulares $E_l = -me^4/2l^2$ ou, em termos do “número quântico” n_φ ,

$$E_{n_\varphi} = -\frac{2\pi^2 me^4}{n_\varphi^2 h^2}.$$

Essa fórmula simples contém na realidade um dos grandes triunfos da teoria de Bohr. De fato, usando a relação entre as energias dos estados estacionários e as frequências correspondentes à transição entre eles, o que se obtém para estas últimas é

⁵Como pode se imediatamente observado na expressão seguinte, essa energia tende a $-\infty$ quando l (ou n_φ) $\rightarrow 0$, como resultado do fato de que o potencial é singular para $r = 0$ e uma órbita circular de momento angular zero tem raio também zero (isto é, se reduz a um ponto). Essas dificuldades estão relacionadas com as que levarão Heisenberg à substituição de (1.4) por algo como $\Delta I = h\Delta n$, o que permite que se tenha $I = I_0 + nh$, com I_0 independente de n , v. pag. 24.

$$\nu_{n \rightarrow n'} = \frac{2\pi^2 m e^4}{h^3} \left(\frac{1}{n'^2} - \frac{1}{n^2} \right), \quad n > n'.$$

No caso particular correspondente a $n' = 2$, essa expressão reproduz a fórmula empírica montada por Balmer em 1885 para representar os comprimentos de onda associados às linhas do espectro visível do hidrogênio

$$\lambda = \frac{c}{\nu} \simeq \frac{3645.6}{n^2 - 4} 10^{-8} \text{ cm}, \quad n = 3, 4, 5, \dots$$

que foi reescrita por Rydberg, em 1890, sob a forma

$$\frac{\nu}{c} = \frac{1}{\lambda} \simeq 109720 \left(\frac{1}{2^2} - \frac{1}{n^2} \right) \text{ cm}^{-1}.$$

Além disso, ela dá uma expressão para a constante empírica que aparece na fórmula de Rydberg

$$\sim 109720 \text{ cm}^{-1} \rightarrow \frac{2\pi^2 m e^4}{h^3 c}$$

que a reproduz *quantitativamente* para o valor empírico da constante de Planck h (obtido, por exemplo, do espectro da radiação de corpo negro).

Outra propriedade notável da formula de Bohr para as freqüências é a de que ela satisfaz ainda um outro “princípio”, este chamado por Bohr *princípio da correspondência*: ele consiste em que

pode-se esperar que qualquer teoria capaz de descrever este fenômeno de acordo com a experiência [*Bohr fala dos espectros atômicos*] deve ser algum tipo de generalização natural da teoria comum [*‘teoria comum’ significa aqui a mecânica e a eletrodinâmica clássicas aplicadas ao problema da radiação*].

Esse princípio pode na realidade ser visto como uma extensão ao problema dos espectros atômicos da virtude estrutural da fórmula de Planck para a densidade espectral da radiação de corpo negro, segundo a qual ela reproduz o resultado clássico (de Rayleigh-Jeans) no limite de baixas freqüências. O que se deve esperar aqui é que, no limite de baixas freqüências (que corresponde a transições $n + 1 \rightarrow n$ com $n \gg 1$, isto é, a números quânticos *grandes*) as freqüências de Bohr reproduzam as previstas pela ‘teoria comum’. Esta prevê, como é sabido, que a freqüência da radiação seja igual à freqüência de revolução dos elétrons. O que a fórmula de Bohr dá, para $n' = n + 1 \gg 1$, é

$$\nu_{n+1 \rightarrow n} = \frac{4\pi^2 m e^4}{h^3 n^3} + \mathcal{O}(n^{-4}) \simeq \frac{m e^4}{2\pi l^3}, \quad (1.5)$$

que de fato coincide com a freqüência clássica de revolução dos elétrons.

1.2.2 Intensidades.

Além de seus respectivos comprimentos de onda, ou frequências, as linhas dos espectros atômicos têm intensidades relativas que podem ser medidas e que devem também poder ser calculadas a partir da teoria, e aqui a teoria de Bohr encontra dificuldades graves. O procedimento seguido por Bohr nessa questão pode ser parafraseado da seguinte forma. Classicamente, a energia radiada por unidade de tempo por uma carga e acelerada (não relativística) é dada pela fórmula de Larmor

$$\delta\epsilon = \frac{2}{3} \frac{e^2}{c^3} \langle |\ddot{q}|^2 \rangle.$$

No caso de um oscilador harmônico de frequência angular $\omega = 2\pi \nu$ a coordenada q pode ser expressa em termos da variável de ação I e da variável angular canonicamente conjugada $\theta(t) = \omega t$ como $q(t) = q_0(I) \exp(i\theta)$. Para tratar o caso mais geral do movimento periódico de um oscilador anarmônico pode-se usar uma expansão em série de Fourier

$$q(t) = \sum_k q_k(I) \exp(ik\theta)$$

de modo que a energia radiada pode ser escrita como

$$\delta\epsilon = \sum_k \delta\epsilon_k = \frac{2}{3} \frac{e^2}{c^3} \sum_k |q_k(I)|^2 k^4 \omega^4. \quad (1.6)$$

Bohr procura então obter uma versão quântica dessa expressão, usando como guia o princípio de correspondência: no caso de radiação de baixa frequência, a frequência clássica $k\omega$ e a variável clássica de ação I podem ser reinterpretadas como $\sim 8\pi^3 m e^4 / h^3 n^3$ e nh respectivamente. Para frequências que não são baixas no sentido do princípio de correspondência, no entanto, aparece a dificuldade “de princípio” de que a expressão quântica para a frequência envolve *dois* estados estacionários com valores *diferentes* para a variável de ação, o que introduz uma ambigüidade insuperável na reinterpretação de $q_k(I)$. Bohr observa finalmente que, na teoria quântica, esses coeficientes de Fourier estão associados a “probabilidades de transição espontânea” entre pares de estados estacionários, e que “evidentemente”, sem uma teoria detalhada do mecanismo dessas transições, não se pode calcular essas probabilidades exatamente.

1.2.3 Einstein, os princípios de Bohr e a radiação de corpo negro.

Na mesma época (1917) em que o longo e elaborado trabalho de Bohr sobre os espectros atômicos estava sendo escrito, aparece um trabalho de Einstein em que a fórmula de Planck para a densidade espectral da radiação de corpo negro é reobtida a partir dos “princípios”

de Bohr e de resultados gerais de mecânica estatística. Um dos aspectos mais importantes desse trabalho é que nele intervêm explicitamente as “probabilidades de transição” referidas por Bohr. Combinando hipóteses muito gerais sobre elas com condições naturais sobre o limite de altas temperaturas, Einstein obtém, em particular, relações gerais entre essas probabilidades e as frequências de Bohr associadas à transição.

As hipóteses utilizadas por Einstein foram de fato as seguintes:

1) A cavidade que constitui o corpo negro é formada de objetos quânticos (“moléculas”), que podem existir, no sentido proposto por Bohr, em estados estacionários caracterizados por certos valores discretos E_n de energia.

2) A uma dada temperatura T , a probabilidade de que uma das moléculas esteja num estado estacionário de energia E_n é proporcional ao fator de Boltzmann

$$W_n = p_n \exp\left(\frac{-E_n}{kT}\right)$$

em que a constante p_n é um *peso estatístico* característico do estado estacionário.

3) Ocorrem transições *espontâneas* entre estados estacionários de energia E_m e E_n ($< E_m$) com a emissão de radiação eletromagnética. A probabilidade de ocorrência dessas transições depende, e depende *apenas*, dos dois estados estacionários envolvidos nela, sendo escrita como

$$dW_{\text{ex}} = A_n^m dt,$$

onde a notação foi escolhida de forma que o índice superior corresponde ao estado inicial e o índice inferior ao estado final da transição.

4) As moléculas podem absorver radiação presente no meio efetuando também transições entre estados estacionários. Sendo $\rho(\nu)$ a densidade espectral da radiação, a probabilidade de absorção associada a uma transição entre os estados E_n e E_m tem a forma

$$dW_{\text{abs}} = B_m^n \rho(\nu_{mn}) dt$$

onde os coeficientes B_m^n dependem apenas dos estados envolvidos na transição.

5) A presença, no meio, de radiação com densidade espectral ρ dá lugar também a *emissão induzida* de radiação entre níveis E_m e E_n com probabilidade

$$dW_{\text{ei}} = B_n^m \rho(\nu_{mn}) dt$$

onde os coeficientes B_n^m também dependem apenas dos estados envolvidos na transição.

Enquanto as hipóteses 4) e 5) têm análogos diretos em um cálculo puramente clássico, a *emissão espontânea* de radiação introduzida na hipótese 3) é um fenômeno de natureza essencialmente quântica.

Embora os vários ingredientes envolvidos nessas hipóteses dependam da natureza detalhada da dinâmica das moléculas e de sua interação com a radiação, a densidade espectral *de equilíbrio* deve ser independente disso. De fato, a condição de equilíbrio das moléculas com a radiação, isto é, a igualdade das taxas totais de emissão e de absorção de radiação de frequência ν_{mn} se exprime como

$$W_n dW_{\text{abs}} = W_m [dW_{\text{ei}} + dW_{\text{ex}}] \quad \text{ou} \quad W_n B_m^n \rho(\nu_{mn}) = W_m [B_n^m \rho(\nu_{mn}) + A_n^m].$$

Por outro lado, no limite de altas temperaturas, $T \rightarrow \infty$, $W_n \rightarrow p_n$ e também, pela lei de Wien, $\rho \rightarrow \infty$, donde se obtém a relação

$$p_n B_m^n = p_m B_n^m.$$

Usando essa relação e resolvendo a condição de equilíbrio para $\rho(\nu_{mn})$ resulta

$$\rho(\nu_{mn}) = \frac{A_n^m / B_n^m}{\exp \frac{E_m - E_n}{kT} - 1}$$

que reproduz a expressão de Planck na forma (1.1) desde que $E_m - E_n = h\nu_{mn}$ (o que constitui na realidade uma dedução da relação de Bohr entre energias e frequência) e $A_n^m / B_n^m = \alpha \nu_{mn}^3$. Este último resultado é um bonus extremamente importante da dedução, pois ele relaciona a probabilidade, essencialmente quântica, de emissão espontânea com a probabilidade de emissão induzida, que possui um análogo clássico direto. Incorporando a forma explícita da constante α obtida em (1.2) ele aparece como

$$\frac{A_n^m}{B_n^m} = \frac{8\pi h}{c^3} \nu_{mn}^3$$

onde o caráter “essencialmente quântico” da emissão espontânea é indicado explicitamente pela presença da constante de Planck h como *fator*.

Uma crítica tão interessante quanto importante a esse trabalho de Einstein, bem como a correspondente reformulação do ponto criticado, foi a ele encaminhada em 1924 por Satyendranath Bose, então na Universidade de Dacca (Índia), e publicada em *Zeitschrift für Physik* **26**, 178 (1924) com o título *Plancks Gesetz und Lichtquantenhypothese*, em tradução para o alemão feita pelo próprio Einstein, e na qual o autor é identificado simplesmente como Bose, sem qualquer prenome ou inicial. Nesse trabalho Bose critica a utilização da lei de Wien (um resultado clássico) na re-dedução feita por Einstein da fórmula de Planck, e substitui o seu uso por um tratamento dentro dos cânones da mecânica estatística dos “quanta de luz” introduzidos também por Einstein no seu tratamento do efeito fotoelétrico. Numa *Nota do Tradutor*, acrescentada ao trabalho de Bose,

Einstein enfatiza a importância desse desenvolvimento, que “pode ser estendido de forma a dar uma teoria quântica dos gases ideais”, como promete fazer (e faz, efetivamente) num trabalho futuro (*Sitzber. Kgl. Preuss. Akad. Wiss.* 1925, 3). É na extensão escrita por Einstein do trabalho de Bose que aparece pela primeira vez o resultado conhecido como “condensação de Bose-Einstein” de um gás ideal a temperatura muito baixa. Esse fenômeno, freqüentemente citado hoje como um exemplo notável de comportamento quântico à escala macroscópica, foi portanto identificado *antes* das primeiras formulações consistentes da mecânica quântica por Heisenberg e Schrödinger.

1.3 A nova mecânica de Heisenberg.

O tratamento dado por Einstein, no trabalho sobre a radiação de corpo negro, à questão da emissão e absorção de radiação, representa uma atitude bastante diferente da adotada, por exemplo, por Bohr, frente à mesma questão. De fato, enquanto Bohr parte de resultados clássicos, procurando reinterpretá-los em termos quânticos com ajuda do princípio de correspondência, Einstein começa por introduzir, como objetos hipotéticos básicos, “probabilidades de transição” entre estados estacionários. A existência e relevância de tais objetos é claramente identificada por Bohr como uma consequência de seus “princípios”, sem que no entanto isso chegue a ser usado de forma ativa no desenvolvimento teórico. Nesse desenvolvimento, o dicionário fornecido pelo princípio de correspondência se mostra insuficiente para calculá-los em geral, devido, em última análise, justamente à incompatibilidade entre a teoria clássica e os “princípios” quânticos. No tratamento de Einstein, por outro lado, as probabilidades de transição são tratadas simplesmente como “novos” objetos a determinar, dotados desde o início das propriedades específicas essenciais ao contexto quântico, particularmente dependência dos *dois* estados estacionários envolvidos na transição e a existência de uma probabilidade de emissão espontânea de radiação, de natureza puramente quântica, além das probabilidades de absorção e de emissão induzida, que tem análogos clássicos. O problema da densidade espectral da radiação de corpo negro de fato dispensa, devido à sua natureza estatística e de equilíbrio, uma especificação mais detalhada das diferentes probabilidades de transição, embora acabe por dar indicações quantitativas bastante específicas sobre relações existente entre elas bem como sobre a adequação geral das hipóteses feitas. Desse modo, embora sem colocar a questão da construção de uma teoria quântica consistente, o trabalho de Einstein indica claramente as vantagens que podem resultar de um distanciamento controlado de ingredientes clássicos que se mostrem em conflito com os princípios quânticos.

Os passos decisivos para a construção de uma teoria quântica completa foram dados por Heisenberg, em 1925, com base em dois pontos centrais. O primeiro é simplesmente o de que *a mecânica clássica não é válida na escala atômica*. O segundo tem na devida conta a inegável

adequação das teorias clássicas nos contextos em que ela se aplica, e consiste na observação de que *o princípio de correspondência de Bohr é uma condição fundamental para a nova teoria*. Foi essencial, no entanto, interpretar a não validade da mecânica clássica na escala atômica num sentido muito mais radical do que havia sido feito anteriormente por Bohr. Bohr, de fato, modificou o comportamento dinâmico dos sistemas atômicos através de seus princípios quânticos, mas *reteve o substrato cinemático* da teoria clássica. Heisenberg, no entanto, escreve na introdução de seu trabalho de 1925 (que tem como título “Reinterpretação quântica de relações cinemáticas e mecânicas”) que

a condição de frequência de Einstein-Bohr (que é válida em todos os casos) por si mesma já representa uma ruptura tão completa da mecânica clássica, ou antes (adotando o ponto de vista da teoria ondulatória) da cinemática que serve de base a essa mecânica, que mesmo para os problemas quânticos mais simples a validade da mecânica clássica simplesmente não pode ser mantida.

A formulação da nova teoria deve incluir portanto uma reformulação também dos ingredientes cinemáticos da teoria clássica.

O caminho seguido por Heisenberg consistiu, basicamente, em reter a equação de movimento clássica (o sistema tratado tendo um único grau de liberdade)

$$M \frac{d^2 q}{dt^2} - f(q) = 0 \quad (1.7)$$

o que significa reter de certo modo o esquema dinâmico formal da mecânica clássica, mas sem que isso implique necessariamente em interpretar $q(t)$ como uma função do tempo com o significado de posição. Pode-se dizer então que esse caminho vai consistir basicamente numa reformulação cinemática *apenas*, que vai no entanto afetar profundamente os procedimentos envolvidos no seu uso e a interpretação dos resultados. Na realidade, Heisenberg propõe que $q(t)$ deva ser visto em geral como um objeto associado a “quantidades atômicas observáveis”, como níveis de energia, frequências de Bohr, intensidades.

O processo seguido na especificação de tal objeto pode ser esquematizado da seguinte forma. No caso de um movimento clássico, periódico, de período $\tau = 2\pi/\omega$ e em geral anarmônico, é possível extrair “elementos” de frequência bem definida expandindo $q(t)$ em série de Fourier

$$q(t) = \sum_{k=-\infty}^{\infty} q_k e^{ik\omega t}$$

com a realidade de $q(t)$, $q(t) = q^*(t)$, garantida pela condição $q_k^* = q_{-k}$. Na “velha” teoria quântica de Bohr $q(t)$, e portanto também ω e os q_k , se referem a órbitas clássicas particulares

para as quais é satisfeita uma condição de quantização. Dessa forma a expansão de Fourier aparece ali como

$$q(n, t) = \sum_{k=-\infty}^{\infty} q_k(n) e^{ik\omega_n t} \quad (1.8)$$

onde o número quântico n provém da condição de quantização para a ação, $I = nh$. Nesse processo a interpretação de $q(n, t)$ como posição, função do tempo, é mantida, e faz com que as frequências que aparecem na expansão sejam os múltiplos $k\omega_n$ da frequência angular da órbita quantizada. Como as frequências observadas não são em geral essas, mas sim as frequências de Bohr associadas a *transições* entre dois estados estacionários, Heisenberg *substitui* o conjunto de termos da expansão de Fourier (visto como “conjunto de elementos de frequência bem definida” que caracterizam o movimento clássico) por um “conjunto” alternativo, quântico, cujos elementos não devem necessariamente ser vistos como termos de uma série, relacionado com o primeiro através das substituições $k\omega_n \rightarrow \omega(n, n - k)$ e $q_k(n) \rightarrow q(n, n - k)$, onde agora as frequências correspondem a frequências de Bohr para a transição entre dois estados estacionários n e $n - k = m$, isto é

$$\omega(n, m) = \frac{2\pi}{h}(E_n - E_m)$$

e as novas amplitudes $q(n, n - k)$ são entendidas como estando *também* associadas a *transições* entre os estados estacionários n e $n - k$. Dessa forma

$$q_k(n) e^{ik\omega_n t} \rightarrow q(n, n - k) e^{i\omega(n, n - k)t}.$$

Os reflexos dessas regras de substituição para o momento $p(t) = M\dot{q}(t)$ (e seus “elementos” de frequência bem definida) podem ser obtidos facilmente. Da expansão de Fourier para $q(n, t)$ obtém-se, de fato,

$$p(n, t) = \sum_{k=-\infty}^{\infty} p_k(n) e^{ik\omega_n t}$$

com $p_k(n) = iMk\omega_n q_k(n)$, o que leva à substituição

$$p_k(n) \rightarrow p(n, n - k) = iM\omega(n, n - k)q(n, n - k).$$

Uma classe importante de variáveis dinâmicas que deve ser considerada é a que envolve *produtos* de vários fatores $q(t)$ e $p(t)$, como $q^2(t)$, $p^2(t)$, $p(t)q(t)$ etc. Classicamente (e supondo já implementadas regras de quantização no sentido da escolha de órbitas clássicas particulares, como feito por Bohr) a expansão de Fourier

$$q_n^2(t) = \sum_{k=-\infty}^{\infty} \mathcal{Q}_k(n) e^{ik\omega_n t}$$

pode ser imediatamente relacionada com a expansão correspondente de $q(t)$ usando a “regra de composição das freqüências” $k\omega_n + k'\omega_n = (k + k')\omega_n$, que dá

$$\mathcal{Q}_k(n) = \sum_{l=-\infty}^{\infty} q_{k-l}(n) q_l(n). \quad (1.9)$$

Fazendo, por outro lado, as substituições $\mathcal{Q}_k(n) \rightarrow \mathcal{Q}(n, n - k)$ e $k\omega_n \rightarrow \omega(n, n - k)$, e tendo em conta a regra de composição das freqüências quânticas

$$\omega(n, n - k) = \omega(n, n - l) + \omega(n - l, n - k) \quad (1.10)$$

(que, na realidade, corresponde à chamada regra de combinação de Rydberg-Ritz, conhecida da análise fenomenológica de espectros atômicos), resulta que a tradução quântica “natural” da relação (1.9) é

$$\mathcal{Q}(n, n - k) = \sum_{l=-\infty}^{\infty} q(n, n - l) q(n - l, n - k). \quad (1.11)$$

Uma característica importante desta “regra de multiplicação quântica” é a de que o produto de variáveis dinâmicas diferentes, $x(t)$ e $y(t)$, deve ser visto em geral como *não comutativo*, $[xy](t) \neq [yx](t)$.

Resta agora determinar as freqüências $\omega(n, n - k)$ e as amplitudes $q(n, n - k)$ a partir da equação de movimento (1.7). Para isso deve-se lembrar que a dependência com n de $q(n, n - k)$ provém de uma reinterpretação da dependência com n de $q_k(n)$, que por sua vez resulta da condição de quantização de Bohr (1.4). É preciso portanto, antes de mais nada, reinterpretar essa condição em termos da cinemática quântica, o que na realidade é uma das passagens ao mesmo tempo mais “livres” e mais cruciais do trabalho de Heisenberg.

O primeiro ponto levantado consistiu em enfraquecer a condição de quantização de Bohr (1.4), (a qual, nessa forma, “não se encaixa bem no cálculo dinâmico”), mantendo-a no entanto consistente com o princípio de correspondência. Em particular, para que o período clássico das órbitas de Bohr coincida, para valores grandes da ação $I = nh$, $n \gg 1$, com as freqüências de Bohr $(E_n - E_{n-1})/h$, basta que valha $\Delta I = h$ para duas órbitas consecutivas. Isso leva a que, desse ponto de vista, a condição (1.4) possa ser substituída por algo como

$$h = \frac{d}{dn} (nh) = \frac{d}{dn} \int_0^\tau p_n(t) \dot{q}_n(t) dt$$

ou, usando as expansões de Fourier de $p_n(t)$ e $q_n(t)$,

$$h = 2\pi M \sum_{k=-\infty}^{\infty} k \frac{d}{dn} \left(k\omega_n |q_k(n)|^2 \right).$$

O sentido preciso da derivação com relação a n nessas expressões não chega a ter maior relevância dado que elas servem apenas para motivar, de certa forma, a reinterpretação quântica da condição de quantização como sendo

$$h = 4\pi M \sum_{k=0}^{\infty} \left[|q(n, n+k)|^2 \omega(n+k, n) - |q(n, n-k)|^2 \omega(n, n-k) \right] \quad (1.12)$$

com o entendimento de que, existindo um estado estacionário “fundamental” rotulado por $n = n_0$, cuja energia E_{n_0} é menor que a de qualquer outro estado estacionário, deve-se ter ainda

$$q(n_0, n_0 - j) \equiv 0, \quad \text{para} \quad j > 0.$$

É claro que a passagem de (1.4) a (1.12), passando pela derivação simbólica com relação a n , não deve ser vista de forma alguma como uma dedução, mas sim como uma proposta de “adaptação de postulado” no sentido da construção de uma teoria quântica consistente.

O funcionamento desse esquema pode ser ilustrado aplicando-o ao caso simples do oscilador harmônico. Nesse caso $f(q) = -M\omega_0^2 q$ na equação de movimento (1.7), onde agora q corresponde ao conjunto de objetos $\{q(n, n+k) \exp[i\omega(n, n+k)t]\}$. Como neste caso o movimento clássico é harmônico e conseqüentemente a expansão de Fourier (1.8) tem um único termo, é razoável supor que apenas as amplitudes do tipo $q(n, n \pm 1)$ são diferentes de zero. Substituindo esses objetos na equação de movimento resulta

$$\left[-\omega(n, n \pm 1)^2 + \omega_0^2 \right] q(n, n \pm 1) = 0$$

o que mostra que $\omega(n, n \pm 1) = \mp\omega_0$ para qualquer n . As amplitudes $q(n, n \pm 1)$ são agora determinadas pela condição de quantização (1.12). Usando $n = 0$ para o estado fundamental resulta que $q(0, -1)$ deve se anular (por estar associado a uma transição para um estado que estaria *abaixo* do estado fundamental). A condição de quantização calculada para $n = 0$ dá portanto

$$|q(0, 1)|^2 = \frac{h}{4\pi M\omega_0}.$$

A mesma condição de quantização calculada para $n = 1$ dá

$$|q(1, 2)|^2 - |q(1, 0)|^2 = \frac{h}{4\pi M\omega_0}.$$

O termo $|q(1, 0)|^2$ pode ser obtido notando que as amplitudes $q(1, 0)$ e $q(0, 1)$ estão associadas, respectivamente, às frequências $\omega(1, 0) = \omega_0$ e $\omega(0, 1) = -\omega_0$. Desse modo, uma “condição de realidade” análoga à que deve ser imposta sobre os coeficientes de Fourier de (1.8) aparece como sendo $q(1, 0) = q(0, 1)^*$. Portanto $|q(1, 0)|^2 = |q(0, 1)|^2$, donde segue que

$$|q(1, 2)|^2 = 2\frac{h}{4\pi M\omega_0}.$$

Em geral, a condição de quantização estabelece uma relação de recorrência entre $|q(n, n+1)|^2$ e $|q(n-1, n)|^2$ da qual se obtém

$$|q(n, n+1)|^2 = (n+1)\frac{h}{4\pi M\omega_0}. \quad (1.13)$$

“Um passo seguinte óbvio”, como escreveu Heisenberg em 1925, seria usar esse resultado em conjunção com a expressão clássica para a energia,

$$E = \frac{M}{2}\dot{q}^2 + \frac{M}{2}\omega_0^2 q^2,$$

sendo no entanto necessário lembrar que $\dot{q}^2(n, n-k)$ e $q^2(n, n-k)$ são agora objetos de dois índices que devem ser calculados usando a “regra de multiplicação quântica” (1.11). Desse modo resulta que os únicos produtos não nulos são aqueles com $k = 0$ e $k = \pm 2$ (isto é, do tipo $\dot{q}^2(n, n)$, $\dot{q}^2(n, n \pm 2)$, $q^2(n, n)$, $q^2(n, n \pm 2)$), os quais são dados por

$$\begin{aligned} \dot{q}^2(n, n) &= \dot{q}(n, n-1)\dot{q}(n-1, n) + \dot{q}(n, n+1)\dot{q}(n+1, n) = (2n+1)\frac{h\omega_0}{4\pi M} \\ q^2(n, n) &= q(n, n-1)q(n-1, n) + q(n, n+1)q(n+1, n) = (2n+1)\frac{h}{4\pi M\omega_0} \\ \dot{q}^2(n, n \pm 2) &= \dot{q}(n, n \pm 1)\dot{q}(n \pm 1, n \pm 2) = -\omega_0^2 q(n, n \pm 1)q(n \pm 1, n \pm 2) = -\omega_0^2 q^2(n, n \pm 2) \end{aligned}$$

de modo que apenas os termos com $k = 0$ contribuem para a energia total. Um ponto importante relativo a este cálculo “quântico” é o de que, de forma semelhante ao que ocorre no contexto clássico, ele dá um resultado *independente do tempo* para a energia *total* embora os termos correspondentes à energia cinética e à energia potencial dependam do tempo quando considerados

separadamente. O resultado do calculo quântico para a energia total, vista também como um objeto quântico de dois índices, pode portanto ser escrito

$$E(n, n - k) = \left(n + \frac{1}{2}\right) \frac{h\omega_0}{2\pi} \delta_{k,0}$$

onde o delta de Kronecker $\delta_{k,0}$ indica explicitamente que apenas os objetos do tipo $E(n, n)$ são diferentes de zero. Esses valores correspondem aos valores possíveis da energia do oscilador quântico, e em particular *diferem por uma constante aditiva* dos valores

$$E_n = n \frac{h\omega_0}{2\pi}$$

obtidos usando diretamente a regra de quantização de Bohr com a cinemática clássica. A cinemática quântica, por outro lado, resolve completamente a ambigüidade existente na expressão (1.6) para as intensidades das transições. Substituindo $|q_k(I)|^2$ pelos objetos quânticos $|q(n, n - k)|^2$ resulta que no oscilador harmônico haverá apenas transições entre estados de energia consecutivos, $E(n, n)$ e $E(n - 1, n - 1)$, e que suas intensidades são proporcionais a n , como mostrado pela Eq. (1.13).

1.3.1 A teoria de Heisenberg passada a limpo: as matrizes de Born e Jordan.

O trabalho de Heisenberg foi recebido para publicação em *Zeitschrift für Physik* em 29 de julho de 1925. Menos de dois meses depois, em 27 de setembro, o mesmo periódico recebeu um trabalho de Born e Jordan, intitulado “Sobre a Mecânica Quântica” em cuja introdução lê-se o seguinte:

O tratamento teórico de Heisenberg publicado recentemente neste periódico, que teve como objetivo estabelecer um novo formalismo cinemático e mecânico em conformidade com as exigências básicas da teoria quântica, nos parece ter um significado potencial considerável. Ele representa uma tentativa de fazer justiça aos novos fatos estabelecendo um sistema conceitual novo e realmente adequado em vez de adaptar as idéias costumeiras de forma mais ou menos artificial e forçada. A argumentação física que levou Heisenberg a esse desenvolvimento foi tão claramente exposta por ele que quaisquer observações suplementares parecem supérfluas. Mas, como ele mesmo observa, o seu tratamento está apenas em seu estágio inicial no que se refere aos aspectos formais, matemáticos. Suas hipóteses foram aplicadas apenas a exemplos simples sem ser estendidas completamente a uma teoria generalizada. Tendo estado

na vantajosa posição de nos familiarizarmos com suas idéias ao longo de seu estágio de elaboração, buscamos agora (pois suas investigações foram concluídas) esclarecer o conteúdo matematicamente formal de seu tratamento e apresentar aqui alguns de nossos resultados. Eles indicam que é de fato possível, a partir das premissas colocadas por Heisenberg, construir uma teoria matemática fechada da mecânica quântica que mostra analogias notavelmente próximas com a mecânica clássica, mas que ao mesmo tempo preserva os aspectos característicos dos fenômenos quânticos.

A formulação geral de Born e Jordan adota um ponto de vista hamiltoniano, no sentido de que utiliza coordenadas e momentos em vez de coordenadas e velocidades. Ela começa por reconhecer que, dentro das premissas de Heisenberg, um sistema dinâmico (quântico), com um único grau de liberdade e classicamente periódico, deve ser descrito por uma coordenada espacial \mathbf{q} e um momento \mathbf{p} que devem ser representados por *matrizes* dependentes do tempo, cujos elementos têm a forma

$$\mathbf{q} \equiv \{q_{nm}e^{i\omega_{nm}t}\}, \quad \mathbf{p} \equiv \{p_{nm}e^{i\omega_{nm}t}\} \quad (1.14)$$

onde ω_{nm} são as frequências associadas a transições entre estados descritos pelos números quânticos n e m . Essas matrizes devem ser *hermiteanas*, isto é, para qualquer t real, devem ser iguais às complexo-conjugadas de suas transpostas. Dessa forma

$$q_{nm}q_{mn} = |q_{nm}|^2 \quad \text{e} \quad \omega_{nm} = -\omega_{mn}.$$

A primeira dessas quantidades estará eventualmente associada a probabilidades de transições $n \rightleftharpoons m$. A relação entre ω_{nm} e ω_{mn} , por outro lado, deve ser entendida como ligada à existência de quantidades E_n tais que

$$h\frac{\omega_{nm}}{2\pi} = E_n - E_m \quad (1.15)$$

o que leva à regra de composição de frequências (1.10). A “regra de multiplicação quântica” (1.11) corresponde então simplesmente ao *produto de matrizes*:

$$\mathbf{q}^2 \equiv \left\{ \sum_k q_{nk}e^{i\omega_{nk}t} q_{km}e^{i\omega_{km}t} \right\} = \left\{ \left(\sum_k q_{nk}q_{km} \right) e^{i\omega_{nm}t} \right\}$$

pois $\omega_{nk} + \omega_{km} = \omega_{nm}$. Outras variáveis dinâmicas (funções de \mathbf{p} e \mathbf{q}) podem ser construídas de forma semelhante, com a propriedade de que o elemento n, m da matriz resultante terá sempre a dependência temporal dada pelo fator $e^{i\omega_{nm}t}$. Variáveis dinâmicas *independentes do tempo* devem portanto corresponder a *matrizes diagonais*, isto é, a matrizes cujos únicos

elementos não nulos sejam aqueles com $n = m$, pois $\omega_{nn} = 0$. Em particular, para que a energia total seja independente do tempo, é preciso que a hamiltoniana (que é uma função de \mathbf{p} e \mathbf{q}) seja representada por uma matriz diagonal $\mathbf{H} \equiv \{H_n \delta_{nm}\}$, sendo ainda natural identificar os elementos diagonais H_n com as energias E_n envolvidas nas frequências ω_{nm} .

A própria condição de quantização (1.12) pode ser reescrita de forma muito simples em termos de produtos de matrizes. De fato, usando o fato de que na condição usada por Heisenberg $Mq(n+k, n)\omega(n+k, n)$ corresponde a $-ip_{n+k, n}$, é fácil ver que ela pode ser reescrita como

$$\sum_k (p_{nk}q_{kn} - q_{nk}p_{kn}) = \frac{h}{2\pi i}.$$

O lado esquerdo dessa relação corresponde à diferença dos elementos diagonais n, n dos produtos matriciais \mathbf{pq} e \mathbf{qp} respectivamente. O fato de que essa diferença não é nula indica a *não comutação* das duas matrizes \mathbf{p} e \mathbf{q} . Na realidade, Horn e Jordan argumentam ainda que o comutador $[\mathbf{p}, \mathbf{q}] \equiv \mathbf{pq} - \mathbf{qp}$ deve ser *independente do tempo*, e portanto diagonal, o que os leva a substituir essa condição por outra “condição de quantização mais forte”, que corresponde à relação matricial

$$\mathbf{pq} - \mathbf{qp} = \frac{h}{2\pi i} \mathbf{1} \quad (1.16)$$

onde $\mathbf{1} \equiv \{\delta_{nm}\}$ representa a matriz unidade, matriz diagonal cujos elementos diagonais são todos iguais a 1.

Resta ainda considerar as equações de movimento que devem ser satisfeitas pelas variáveis dinâmicas. Essa questão a rigor já foi decidida com a especificação da dependência temporal adotada para as variáveis dinâmicas juntamente com a identificação das energias E_n com os elementos da matriz diagonal correspondente à hamiltoniana. Dessas suposições resulta, de fato, para uma variável dinâmica representada por uma matriz $\mathbf{g}(\mathbf{p}, \mathbf{q})$ com elementos da forma $g_{nm}e^{i\omega_{nm}t}$,

$$\frac{d}{dt}g_{nm}e^{i\omega_{nm}t} = i\omega_{nm} g_{nm}e^{i\omega_{nm}t} = \frac{2\pi i}{h}(E_n - E_m) g_{nm}e^{i\omega_{nm}t}$$

o que pode ser reescrito de forma matricial, usando o fato de que $\mathbf{H} \equiv \{E_n \delta_{nm}\}$,

$$\frac{d\mathbf{g}}{dt} = \frac{2\pi i}{h}(\mathbf{Hg} - \mathbf{gH}) \equiv \frac{2\pi i}{h}[\mathbf{H}, \mathbf{g}]. \quad (1.17)$$

Equações de movimento para \mathbf{p} e \mathbf{q} correspondem a casos particulares dessa equação de movimento geral. É fácil verificar a estreita relação dessas equações de movimento com as equações de movimento clássicas no caso de uma hamiltoniana da forma $\mathbf{H} = \mathbf{K}(\mathbf{p}) + \mathbf{V}(\mathbf{q})$, em que $\mathbf{K}(\mathbf{p})$ e $\mathbf{V}(\mathbf{q})$ sejam exprimíveis como séries de potências de \mathbf{p} e \mathbf{q} respectivamente, isto é

$$\mathbf{K}(\mathbf{p}) = \sum_j a_j \mathbf{p}^j \quad \text{e} \quad \mathbf{V}(\mathbf{q}) = \sum_j b_j \mathbf{q}^j;$$

para isso basta usar uma propriedade algébrica geral de comutadores de matrizes (que pode ser verificada, por exemplo, por indução, ou recursivamente)

$$[\mathbf{x}^j, \mathbf{y}] = \sum_{k=1}^j \mathbf{x}^{k-1} [\mathbf{x}, \mathbf{y}] \mathbf{x}^{j-k} \quad \text{que se reduz a} \quad j \mathbf{x}^{j-1} [\mathbf{x}, \mathbf{y}] \quad \text{se} \quad [\mathbf{x}, [\mathbf{x}, \mathbf{y}]] = 0$$

a qual dá, usando a “condição de quantização mais forte” (1.16), equações de movimento formalmente idênticas às equações clássicas de Hamilton:

$$\begin{aligned} \frac{d\mathbf{q}}{dt} &= \frac{2\pi i}{h} [\mathbf{K}, \mathbf{q}] = \frac{2\pi i}{h} \sum_j j a_j \mathbf{p}^{j-1} [\mathbf{p}, \mathbf{q}] = \frac{\partial \mathbf{H}}{\partial \mathbf{p}} \\ \frac{d\mathbf{p}}{dt} &= \frac{2\pi i}{h} [\mathbf{V}, \mathbf{p}] = \frac{2\pi i}{h} \sum_j j b_j \mathbf{q}^{j-1} [\mathbf{q}, \mathbf{p}] = -\frac{\partial \mathbf{H}}{\partial \mathbf{q}}. \end{aligned} \tag{1.18}$$

Em resumo, o esquema proposto por Born e Jordan como realização mais geral e fechada do tratamento quântico de Heisenberg para sistemas periódicos com um grau de liberdade, consiste em obter duas matrizes \mathbf{q} e \mathbf{p} , cujos elementos são rotulados por índices associados aos “estados estacionários” do sistema, satisfazendo à condição “mais forte” de quantização (1.16) e sendo ainda tais que a matriz correspondente à hamiltoniana, $\mathbf{H}(\mathbf{p}, \mathbf{q})$, seja *diagonal*. A dependência temporal das variáveis dinâmicas é dada pelas equações (1.14) e (1.15), onde as energias E_n são os elementos diagonais de \mathbf{H} , que correspondem portanto às energias dos estados estacionários. Essa dependência temporal corresponde, então, à que é prescrita pelas equações de movimento (1.18).

Um ponto que é importante notar com relação a esse esquema é o de que a substituição das variáveis canônicas *clássicas* q e p pelas matrizes *não comutantes* \mathbf{q} e \mathbf{p} na construção das matrizes que são correspondentes quânticas de variáveis dinâmicas clássicas $g(p, q)$ leva em geral a ambigüidades quanto à ordenação de fatores \mathbf{q} e \mathbf{p} . Por exemplo, uma quantidade perfeitamente bem definida classicamente como $p^2 q$, pode ser interpretada como $\mathbf{p}^2 \mathbf{q}$, $\mathbf{p} \mathbf{q} \mathbf{p}$, $\mathbf{q} \mathbf{p}^2$, $(\mathbf{p}^2 \mathbf{q} + \mathbf{q} \mathbf{p}^2)/2$, etc, que, devido à não comutação de \mathbf{q} e \mathbf{p} , não coincidem entre si. Uma condição que reduz essas ambigüidades (sem no entanto elimina-las completamente) é a condição de hermiticidade $g_{nm} = g_{mn}^*$. Como nos casos das matrizes \mathbf{q} e \mathbf{p} , essa condição pode ser vista como correspondendo, na teoria quântica, à condição de realidade das variáveis dinâmicas

clássicas. Ela garante, em particular, que os elementos diagonais das matrizes que representam variáveis dinâmicas sejam reais. No caso da hamiltoniana \mathbf{H} ela garante a realidade das energias E_n dos estados estacionários. É fácil verificar que a condição de hermiticidade exclui, no caso considerado acima, combinações não simétricas como $\mathbf{p}^2\mathbf{q}$ e $\mathbf{q}\mathbf{p}^2$, mas não permite optar entre as outras possibilidades simétricas, por exemplo. A ambigüidade remanescente é uma dificuldade real na “quantização” de um sistema a partir de sua descrição clássica. Ela pode ser vista como uma indicação da precedência da descrição quântica do sistema, no sentido de que esta não pode, em geral, ser univocamente inferida da descrição de seu análogo clássico.

1.3.2 O oscilador harmônico unidimensional segundo a mecânica das matrizes.

O tratamento de Heisenberg, na sua versão consolidada por Born e Jordan, pode mais uma vez ser ilustrado com o caso simples, mas relevante, do oscilador harmônico unidimensional. A hamiltoniana tem nesse caso a forma conhecida

$$\mathbf{H} = \frac{1}{2M}\mathbf{p}^2 + \frac{M\omega_0^2}{2}\mathbf{q}^2$$

que em particular não envolve problemas de ordenamento. Born e Jordan observam com relação a ele que

mesmo para este problema simples é necessário suplementar a análise de Heisenberg. Esta última invoca considerações de correspondência para chegar a consequências importantes relativas à forma da solução: a saber, como classicamente só está presente *uma única* componente harmônica, Heisenberg escolhe uma matriz que representa transições entre estados adjacentes apenas, e que portanto [é tal que apenas $q_{n,n\pm1}$ são diferentes de zero]. Nós aqui procuramos construir toda a teoria de forma auto-referente, sem invocar a assistência da teoria clássica com base no princípio de correspondência. Vamos então investigar se [essa forma da matriz \mathbf{q}] pode ser por sua vez deduzida das fórmulas básicas ou, caso isso se mostre impossível, que postulados adicionais se fazem necessários.

O procedimento a seguir difere nos seus detalhes do que foi seguido no trabalho de Born e Jordan, mas no entanto preserva todos os pontos essenciais do procedimento original, embora com uma economia de recursos importante. Ele é hoje amplamente usado para tratar o oscilador harmônico quântico, e pode ser encontrado já na primeira edição, que data de 1930, dos *The Principles of Quantum Mechanics* de Dirac.

As matrizes \mathbf{p} e \mathbf{q} devem satisfazer a relação de comutação (1.16), a matriz \mathbf{H} deve ser diagonal, contendo os valores possíveis da energia do oscilador. Para tratar o problema de determinar tais matrizes é útil fazer a substituição linear

$$\mathbf{a} = \frac{1}{\sqrt{2\hbar M\omega_0}}(M\omega_0\mathbf{q} + i\mathbf{p}), \quad \mathbf{a}^\dagger = \frac{1}{\sqrt{2\hbar M\omega_0}}(M\omega_0\mathbf{q} - i\mathbf{p})$$

onde, e de agora em diante, $\hbar \equiv h/2\pi$. A escolha de constantes de proporcionalidade nessa substituição torna as novas quantidades \mathbf{a} e \mathbf{a}^\dagger adimensionais. A nova matriz \mathbf{a} não é hermiteana, e \mathbf{a}^\dagger é na realidade sua hermiteana adjunta (isto é, a matriz complexo-conjugada da transposta de \mathbf{a}). Usando a relação de comutação (1.16) é imediato verificar que ela corresponde à relação de comutação para as novas matrizes

$$[\mathbf{a}, \mathbf{a}^\dagger] = \mathbf{1}. \quad (1.19)$$

A inversão dessa substituição permite escrever \mathbf{p} e \mathbf{q} em termos de \mathbf{a} e \mathbf{a}^\dagger como

$$\mathbf{p} = i\sqrt{\frac{\hbar M\omega_0}{2}}(\mathbf{a}^\dagger - \mathbf{a}), \quad \mathbf{q} = \sqrt{\frac{\hbar}{2M\omega_0}}(\mathbf{a}^\dagger + \mathbf{a}) \quad (1.20)$$

e exprimir a hamiltoniana também em termos de \mathbf{a} e \mathbf{a}^\dagger :

$$\mathbf{H} = \frac{\hbar\omega_0}{4} [-(\mathbf{a}^\dagger - \mathbf{a})^2 + (\mathbf{a}^\dagger + \mathbf{a})^2] = \hbar\omega_0(\mathbf{a}^\dagger\mathbf{a} + \frac{1}{2}\mathbf{1}). \quad (1.21)$$

Devido à não comutatividade de \mathbf{a} e \mathbf{a}^\dagger , é preciso preservar a ordem dos fatores no cálculo de quantidades como $(\mathbf{a}^\dagger \pm \mathbf{a})^2$. Explicitamente, por exemplo,

$$(\mathbf{a}^\dagger + \mathbf{a})^2 = (\mathbf{a}^\dagger + \mathbf{a})(\mathbf{a}^\dagger + \mathbf{a}) = \mathbf{a}^{\dagger 2} + \mathbf{a}^\dagger\mathbf{a} + \mathbf{a}\mathbf{a}^\dagger + \mathbf{a}^2 = \mathbf{a}^{\dagger 2} + 2\mathbf{a}^\dagger\mathbf{a} + \mathbf{1} + \mathbf{a}^2$$

onde, na última passagem, foi feito uso da relação de comutação (1.19). O problema fica assim reduzido ao de encontrar as matrizes hermiteano-adjuntas \mathbf{a} e \mathbf{a}^\dagger , satisfazendo essa relação de comutação, e de forma que \mathbf{H} seja diagonal.

Esse novo problema pode ser resolvido do seguinte modo: primeiramente, é fácil verificar, usando mais uma vez a (1.19), que

$$[\mathbf{H}, \mathbf{a}] = -\hbar\omega_0\mathbf{a} \quad \text{e} \quad [\mathbf{H}, \mathbf{a}^\dagger] = \hbar\omega_0\mathbf{a}^\dagger$$

ou, usando $\mathbf{H} \equiv \{E_n \delta_{nm}\}$ e exprimindo essas relações explicitamente em termos de elementos de matriz,

$$(E_n - E_m + \hbar\omega_0)a_{nm} = 0 \quad \text{e} \quad (E_m - E_n - \hbar\omega_0)a_{mn}^\dagger = 0.$$

A primeira dessas equações exige que $a_{nm} = 0$ a menos que $E_n = E_m - \hbar\omega_0$; a segunda, por outro lado exige que $a_{mn}^\dagger = 0$ a menos que $E_m = E_n + \hbar\omega_0$. Isso significa que os elementos de matriz a_{nm} são diferentes de zero *apenas* quando n e m correspondem a estados estacionários tais que $E_m - E_n = \hbar\omega_0$, uma afirmação análoga valendo para a_{mn}^\dagger . Sendo \mathbf{q} uma combinação linear de \mathbf{a} e \mathbf{a}^\dagger , segue que os elementos de matriz q_{nm} não nulos correspondem necessariamente a estados estacionários tais que $E_m - E_n = \pm\hbar\omega_0$. Isso mostra como a condição imposta por Heisenberg com base em argumentos de correspondência emerge na realidade da própria estrutura da teoria quântica. De fato, para chegar precisamente à condição imposta por Heisenberg, de que apenas os elementos de matriz $q_{nn\pm1}$ são diferentes de zero, é necessário apenas *rotular* os estados estacionários de forma conveniente. Isso se faz pondo

$$E_n = n\hbar\omega_0 + E_0, \quad n \text{ inteiro} \quad (1.22)$$

onde E_0 é uma constante ainda por determinar. Como observado por Born e Jordan, esta escolha particular de rótulos “serve unicamente para estabelecer a forma normal da solução. Nada de novo é introduzido *fisicamente*”.

Com esta particular “forma normal” de rotular os estados estacionários resulta então que

$$a_{nm} = a_{m-1n} \delta_{n,m-1} \quad \text{e} \quad a_{mn}^\dagger = a_{n+1m}^\dagger \delta_{m,n+1}$$

e portanto

$$E_m = \hbar\omega_0 \left(\sum_n a_{mn}^\dagger a_{nm} + \frac{1}{2} \right) = \hbar\omega_0 \left(a_{mm-1}^\dagger a_{m-1m} + \frac{1}{2} \right) = \hbar\omega_0 \left(|a_{mm-1}^\dagger|^2 + \frac{1}{2} \right)$$

onde na última passagem foi usado o fato de que $a_{m-1m} = a_{mm-1}^{\dagger*}$. Comparando com (1.22) resulta finalmente que

$$E_0 = \frac{1}{2}\hbar\omega_0 \quad \text{e} \quad |a_{nn-1}^\dagger|^2 = n, \quad n = 0, 1, 2, \dots \quad (1.23)$$

pois, neste último caso, $|a_{nn-1}^\dagger|^2 \geq 0$. A matriz diagonal (e portanto independente do tempo) que representa a hamiltoniana, em vista de (1.22) e (1.23), é dada explicitamente por

$$\mathbf{H} = \hbar\omega_0 \begin{pmatrix} 1/2 & 0 & 0 & 0 & \dots \\ 0 & 3/2 & 0 & 0 & \dots \\ 0 & 0 & 5/2 & 0 & \dots \\ 0 & 0 & 0 & 7/2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix} \quad (1.24)$$

e a matriz correspondente a \mathbf{a}^\dagger , com a dependência temporal prescrita por (1.14) e usando também a transformação (1.20),

$$\mathbf{a}^\dagger = e^{-i\omega_0 t} \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix}.$$

A matriz \mathbf{a} é a complexo-conjugada da transposta desta última. Todas essas matrizes têm evidentemente dimensão infinita, e é imediato verificar explicitamente que elas satisfazem a relação (1.21).

A forma explícita da matriz \mathbf{a}^\dagger escrita acima envolve a rigor uma particular escolha de fases para os elementos a_{nn-1}^\dagger . De fato, tudo o que é exigido pela segunda relação (1.23) é que

$$a_{nn-1}^\dagger = \sqrt{n} e^{i\phi_n}$$

onde as fases ϕ_n são constantes arbitrárias. A forma escolhida para \mathbf{a}^\dagger corresponde a tomar $\phi_n = 0$ para todo n . É imediato verificar que as relações de comutação (1.19) e (1.16) bem como a forma de \mathbf{H} e as quantidades observáveis associadas às intensidades das transições $|q_{nn\pm 1}|^2$ não dependem dessas constantes, embora as matrizes correspondentes a \mathbf{p} e \mathbf{q} dependam delas. Essa ambigüidade de fase é na realidade um caso particular bastante trivial de uma simetria da teoria quântica análoga à simetria que permite efetuar transformações canônicas na mecânica clássica hamiltoniana, e será tratada em detalhe oportunamente.

1.4 Opções determinantes e decorrências da teoria de Heisenberg, Born e Jordan.

A teoria quântica de Heisenberg, na versão consolidada por Born e Jordan, contém já explícita ou implicitamente todos os ingredientes formais essenciais da teoria quântica tal como ela é entendida hoje. Suas características mais estranhas, como a reformulação cinemática segundo a qual coordenada e momento devem ser representados por matrizes não comutantes satisfazendo equações de movimento formalmente idênticas às equações clássicas, resultaram na realidade da opção de exprimir a dinâmica dos processos de emissão e absorção de radiação à escala atômica estritamente em termos de "quantidades observáveis", isto é, frequências e intensidades das

diferentes transições possíveis. Sua aplicação a sistemas simples, como o oscilador harmônico, mostra que o novo esquema efetivamente gera, de forma completamente auto-referente, sem necessidade de recursos mais ou menos artificiosos a hipóteses externas, tanto os valores possíveis da energia (e conseqüentemente as as frequências da radiação associada às transições possíveis) como as quantidades que permitem calcular suas intensidades relativas.

Os aspectos intuitivamente mais insólitos da teoria quântica foram portanto gerados a partir de uma adesão bastante estrita (bem mais estrita, em particular, que no caso da teoria quântica de Bohr, por exemplo) a critérios de observabilidade para a escolha dos objetos básicos a serem considerados. Esses critérios são extremamente pragmáticos de fato ⁶, e fornecem efetivamente um conjunto mínimo de ingredientes interpretativos, indispensável para o desenvolvimento do aparato formal. Desse modo, um certo nível na elaboração formal da teoria pôde (ou teve que) ser atingido *antes* que mais atenção fosse dada ao desenvolvimento sistemático de esquemas interpretativos mais amplos. Em particular, é bastante claro que temas 'filosoficamente' favoritos como o princípio de incerteza não tiveram *nenhum* papel nos cruciais primeiros estágios da elaboração da teoria quântica, tendo sido na realidade desenvolvidos apenas em uma fase posterior. O trabalho de Heisenberg sobre o princípio de incerteza, que tem como título "Interpretação da cinemática e da dinâmica da teoria quântica", apareceu apenas em 1927 (*Z. Physik* **43**, 172). Ao contrário dos princípios de Bohr, cujo papel foi o de fornecer orientação necessária no âmbito da situação criada pelas repetidas violações da consistência interna das teorias clássicas, esse princípio deve ser entendido como a explicitação de determinadas *conseqüências da nova teoria*, complementada por uma análise de situações observacionais em que elas se manifestam de forma mais explícita. A questão da interpretação geral da teoria quântica, no sentido em que é ainda hoje tida como questão em aberto, passou a ser tratada de forma mais intensa a partir da Conferência Solvay de 1929. Antes desses dois marcos, ainda em 1925, foi publicado um trabalho conjunto de Born, Heisenberg e Jordan estendendo a teoria para sistemas com mais de um grau de liberdade; e, em 1926, Pauli pode obter da nova teoria o espectro de Balmer do átomo de Hidrogênio, utilizando a conservação do vetor de Runge-Lenz para resolver as equações relevantes no caso. O desenvolvimento dos aspectos operacionais da teoria teve portanto precedência de fato sobre a maior elaboração dos esquemas interpretativos.

Como já mencionado, tanto o interesse como também a controvérsia acerca da interpretação da teoria quântica continuam sendo bastante intensos um século depois do trabalho de Planck. No entanto, é também bastante claro que tal interesse e controvérsia desempenham na realidade um papel bastante limitado na aplicabilidade e no uso efetivo da teoria. Poder-se-ia dizer que o problema da interpretação da teoria quântica se manifesta quase que apenas em circunstâncias

⁶Em que se deve pesar no entanto a posição fortemente *anti-pragmática* adotada por Heisenberg no relato que faz de suas conversas com Barton e Dirac, v. W. Heisenberg, *Diálogos sobre Física Atômica*, Editorial Verbo, Lisboa-São Paulo, 1975, Cap. VIII.

nas quais o interesse predominante recai sobre a questão da interpretação da teoria quântica. Essa situação parece portanto não diferir essencialmente daquela que pode ter prevalecido imediatamente após o aparecimento do trabalho pioneiro de Heisenberg. Em última análise, os aspectos “problemáticos” da teoria quântica decorrem de um *hiato de princípio* criado entre a cinemática das teorias clássicas e a cinemática quântica, e constantemente exacerbado pelo enorme sucesso *de ambas* em seus respectivos domínios. Em alguma medida, eles serão abordados onde apropriado, em capítulos subsequentes.

Problemas.

1.1 - Com relação ao diálogo entre Lorentz e Einstein reproduzido na página 12,

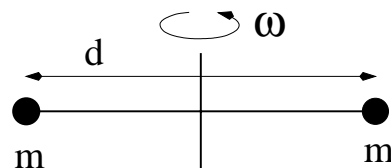
a) mostre explicitamente que a constância da integral de ação $I = \int_0^T p \dot{q} dt$ implica, para um oscilador harmônico, a proporcionalidade entre a energia e a frequência;

b) calcule a amplitude das oscilações, nas condições descritas por Lorentz, para um pêndulo de 10 cm de comprimento com massa de 1 g; estime a ordem de grandeza de efeitos anarmônicos nesse caso;

c) implemente a transposição do argumento para o caso de um circuito oscilante, indicando uma possível realização do circuito e uma variação de parâmetros que seja análoga à descrita para o caso do pêndulo.

1.2 - Use a condição de quantização (1.4) para determinar os valores permitidos para a energia de uma molécula diatômica, vista como um sistema rígido de dois átomos de massa m separados por uma distância d , capaz de girar em torno de um eixo perpendicular à linha que une os dois átomos passando pelo centro de massa da molécula (v. figura 1.1).

Figura 1.1: Molécula diatômica.



a) Escreva o momento $|\vec{p}|$ de cada átomo em termos da sua massa, da velocidade angular ω da molécula e dos seus parâmetros geométricos. O integrando de (1.4) é neste caso $2p^2/m$ e a integração é sobre um período do movimento, $0 \leq t \leq 2\pi/\omega$. A condição de quantização restringe os valores permitidos ω_n para a velocidade angular ω .

b) Use os valores permitidos para a velocidade angular para deduzir os valores permitidos E_n para a energia e I_n para o momento angular da molécula. Verifique que estes últimos são múltiplos de inteiros de $\hbar = h/2\pi$, e portanto independentes de m e d .

c) Mostre que os valores possíveis da energia podem ser obtidos diretamente da expressão clássica $E = I^2/2\mathcal{I}$, onde \mathcal{I} é o momento de inércia da molécula, usando a “quantização” do momento angular obtida no item anterior.

1.3 - Mostre que o lado direito da relação (1.5) corresponde de fato, para o modelo clássico mais simples do átomo de hidrogênio, que consiste de uma partícula de massa M sujeita ao potencial externo $-e^2/r$, à frequência de revolução da órbita circular associada ao valor l do momento angular.

1.4 - a) Mostre que, se uma matriz \mathbf{g} satisfaz

$$\frac{d\mathbf{g}}{dt} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{g}] = 0$$

com \mathbf{H} diagonal, isto é, $H_{nm} = E_n \delta_{nm}$, então segue que $g_{nm} = 0$ sempre que $E_n \neq E_m$;

b) usando as equações de movimento de Heisenberg

$$\frac{d\mathbf{q}}{dt} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{q}] \quad \text{e} \quad \frac{d\mathbf{p}}{dt} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{p}]$$

com

$$\mathbf{H} = \frac{\mathbf{p}^2}{2M} + \mathbf{V}(\mathbf{q}) \quad \text{e} \quad \mathbf{V}(\mathbf{q}) = \sum_j b_j \mathbf{q}^j$$

mostre que $\frac{d}{dt}[\mathbf{q}, \mathbf{p}] = 0$;

c) do resultado b), combinado com a), tomando $\mathbf{g} = [\mathbf{q}, \mathbf{p}]$, Jordan inferiu que os elementos não diagonais do comutador $[\mathbf{q}, \mathbf{p}]$ devem ser zero. Essa inferência é *sempre* correta?

1.5 - Obtenha explicitamente a matriz \mathbf{q} para o oscilador harmônico.

1.6 - Se \mathbf{g} é uma matriz com elementos g_{nm} , chama-se um *autovetor* de \mathbf{g} um conjunto de números c_n , dos quais pelo menos um é diferente de zero, tais que

$$\sum_m g_{nm} c_m = \gamma c_n$$

e nesse caso γ (que é um número não necessariamente real que pode, em particular, ser igual a zero) é chamado *autovalor* associado a esse autovetor.

a) Use essa definição de autovetor para mostrar que os autovetores da matriz $\mathbf{a}(t)$ são definidos pela relação de recorrência $c_n = \gamma e^{-i\omega_0 t} c_{n-1} / \sqrt{n}$, e que portanto $c_n = (\gamma e^{-i\omega_0 t})^n c_0 / \sqrt{n!}$, γ sendo qualquer número complexo;

b) mostre que a matriz \mathbf{a}^\dagger não tem autovetores.

1.7 - Mostre que a matriz que corresponde à hamiltoniana de um oscilador harmônico de frequência ω_0 pode ser escrita em termos das matrizes \mathbf{a}^\dagger e \mathbf{a} como $H = \hbar\omega_0(\mathbf{a}^\dagger\mathbf{a} + \frac{1}{2})$. Lembre-se que a forma dessas matrizes é

$$\mathbf{H} = \hbar\omega_0 \begin{pmatrix} 1/2 & 0 & 0 & 0 & \dots \\ 0 & 3/2 & 0 & 0 & \dots \\ 0 & 0 & 5/2 & 0 & \dots \\ 0 & 0 & 0 & 7/2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \mathbf{a}^\dagger = e^{-i\omega_0 t} \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

e \mathbf{a} é a transposta da complexo-conjugada de \mathbf{a}^\dagger (também chamada *adjunta* de \mathbf{a}^\dagger).

Um vetor $|\nu\rangle$ representando um estado estacionário com energia $E_\nu = \hbar\omega_0(\nu + 1/2)$ (na verdade um auto-vetor de \mathbf{H}) pode ser escrito como um vetor coluna.

a) Represente esse vetor coluna para os estados $|0\rangle$, $|1\rangle$ e $|2\rangle$. Obtenha $\mathbf{a}|0\rangle$, $\mathbf{a}|1\rangle$, $\mathbf{a}|2\rangle$ e $\mathbf{a}^\dagger|0\rangle$, $\mathbf{a}^\dagger|1\rangle$ e $\mathbf{a}^\dagger|2\rangle$. **Sugestão:** faça os cálculos envolvendo as matrizes \mathbf{a}^\dagger e \mathbf{a} tomando $t = 0$. Os fatores exponenciais dependentes do tempo podem ser acrescentados posteriormente sem dificuldade.

b) Dos resultados obtidos em a) escreva uma expressão geral para $\mathbf{a}|\nu\rangle$ e $\mathbf{a}^\dagger|\nu\rangle$ em função de ν . Calcule também $\mathbf{a}^\dagger(\mathbf{a}|\nu\rangle)$.

c) Obtenha a matriz $\mathbf{a}^\dagger\mathbf{a}$ e mostre que o vetor $(\mathbf{a}^\dagger\mathbf{a})|\nu\rangle$ reproduz o último resultado obtido em b), e portanto $\mathbf{a}^\dagger(\mathbf{a}|\nu\rangle) = (\mathbf{a}^\dagger\mathbf{a})|\nu\rangle$.

1.8 - Usando a equação de movimento de Heisenberg

$$\frac{d\mathbf{g}}{dt} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{g}]$$

obtenha a dependência temporal da energia cinética $\mathbf{K}(t)$ e da energia potencial $\mathbf{V}(t)$ para um oscilador harmônico. Use para isso

$$\mathbf{H}(t) = \mathbf{H}(t = 0) = \mathbf{K}(t = 0) + \mathbf{V}(t = 0) = \mathbf{K}(t) + \mathbf{V}(t)$$

com

$$\mathbf{K}(t) = \frac{\mathbf{p}^2(t)}{2M} \quad \text{e} \quad \mathbf{V}(t) = \frac{M\omega_0^2}{2}\mathbf{q}^2(t).$$

Verifique explicitamente, a partir do resultado, que $\mathbf{H}(t) = \mathbf{K}(t) + \mathbf{V}(t) = \mathbf{H}(t = 0)$, isto é, que $\mathbf{H}(t)$ resulta de fato independente do tempo. Mostre também que a diferença $\mathbf{K}(t) - \mathbf{V}(t)$ oscila harmonicamente com frequência $2\omega_0$, e compare esse fato com o resultado clássico correspondente. **Sugestões:** 1) Escreva equações para $\frac{d}{dt}\mathbf{K}$ e para $\frac{d}{dt}\mathbf{V}$ calculando explicitamente os comutadores, lembrando que $[\mathbf{q}(t), \mathbf{p}(t)] = i\hbar\mathbf{1}$ é independente do tempo; 2) derive essas duas equações com relação ao tempo e mostre que as derivadas segundas podem ser re-expressas em termos de $\mathbf{K}(t)$ e $\mathbf{V}(t)$; verifique que daí resultam equações diferenciais acopladas para \mathbf{K} e \mathbf{V} que se desacoplam considerando combinações simples dessas duas variáveis dinâmicas; ache essas combinações e resolva as equações.

Referências gerais.

Os trabalhos clássicos de Einstein, Ehrenfest, Bohr, Heisenberg, Born, Jordan e outros, ligados à origem da mecânica quântica foram reunidos (em traduções para o inglês, quando originalmente escritos em outra língua) no livro

★ **B. L. van der Waerden**, *Sources of Quantum Mechanics*, North Holland Publishing Company (Amsterdam) 1967.

Esse livro contém ainda uma discussão do conteúdo dos trabalhos bem como um estudo acerca das circunstâncias históricas relativas ao seu aparecimento, incluindo depoimentos posteriores obtidos de seus autores. Uma versão apreciavelmente reelaborada da história das origens da mecânica quântica ao longo dessa mesma linha é encontrada em

★ **S. Tomonaga**, *Quantum Mechanics* (Volume I), North Holland Publishing Company (Amsterdam) 1968.

Um livro de ensaios não técnicos de Heisenberg, muitos dos quais com caráter de depoimentos referentes aos primeiros tempos da mecânica quântica é, na sua tradução portuguesa,

★ **W. Heisenberg**, *Diálogos sobre Física Atômica*, Editorial Verbo (Lisboa-São Paulo) 1975.

Um relato sobre as idéias que levaram Schrödinger à formulação da mecânica quântica sob a forma da mecânica ondulatória pode ser lido, em português, em

★ **R. Köberle**, *Sobre a Gênese da Mecânica Ondulatória*, Revista Brasileira de Física **9**, 243-273 (1979).

Alguns textos mais antigos sobre física atômica contém informações didaticamente elaboradas muito mais ricas e abundantes que seus sucedâneos mais recentes, em particular acerca dos trabalhos de Wien, Rayleigh, Jeans, Planck, etc. Dois exemplos disso são

★ **G. P. Harnwell** e **J. J. Livingood**, *Experimental Atomic Physics*, McGraw-Hill Book Company (New York and London) 1933.

★ **Robert B. Leighton**, *Principles of Modern Physics*, McGraw-Hill Book Company (New York Toronto London) 1959.

Esses livros contem também referências bibliograficas completas aos trabalhos originais dos autores citados.

Capítulo 2

Estrutura algébrica da teoria de Heisenberg. Postulados interpretativos.

O propósito deste capítulo é explicitar a estrutura formal *geral* da teoria quântica, tomando como ponto de partida a teoria de Heisenberg na sua versão consolidada por Born e Jordan. Isso vai também permitir estabelecer contato com outras realizações da mesma teoria, especialmente a que resultou do processo paralelo de invenção que culminou com a teoria publicada em 1926 por Schrödinger. A equivalência dessa teoria com a de Heisenberg, não advertida inicialmente, foi provada nesse mesmo ano também por Schrödinger. A consolidação do arcabouço formal comum às duas teorias se deve sobretudo a Dirac, nas sucessivas edições dos *The Principles of Quantum Mechanics*, a primeira das quais data de 1930 e a quarta de 1958. Curiosamente, a mecânica quântica foi novamente reinventada por Feynman nos anos 40 (v. R. P. Feynman, *The Space-Time Approach to Non-relativistic Quantum Mechanics*, Revs. Mod. Phys. **20**, 367 (1948)), numa forma suficientemente diferente das versões anteriores para dar início a um terceiro tipo de realização que, embora seja tecnicamente bastante mais complicado que as realizações de Heisenberg e de Schrödinger, oferece ocasião para pontos de vista novos, interessantes e úteis.

A explicitação de um formalismo geral para a teoria quântica resulta porém incompleta, no sentido que deixa em aberto a relação entre vários de seus ingredientes e o domínio mais concreto dos resultados de medidas realizadas ou realizáveis. Na realidade, a própria teoria de Heisenberg teve necessariamente que incluir alguns elementos de natureza interpretativa, como a identificação dos elementos da matriz hamiltoniana em sua forma diagonal com os valores da energia dos estados estacionários possíveis do sistema descrito, e a identificação do quadrado dos módulos dos elementos da matriz \mathbf{q} correspondente, $|q_{nm}|^2$, como os momentos de dipolo associados a transições entre dois desses estados estacionários, associados à intensidade da radiação emitida nessas transições. A generalização do formalismo exige no entanto também

uma extensão do aparato interpretativo da teoria. A forma corrente dessa extensão também será tratada neste capítulo. Seus ingredientes básicos apareceram pela primeira vez em um trabalho de Born (M. Born, Z. Physik **38**, 803 (1926)) e foram em seguida elaborados por Heisenberg (Z. Physik **43**, 172 (1927)) e Bohr (Naturwiss. **16**, 245 (1928), *ibid.* **17**, 483 (1929) e *ibid.* **18**, 72 (1930)). Esses desenvolvimentos foram suficientes para dar à mecânica quântica uma base operacional adequadamente sólida para sua aplicação efetiva desde então. A elaboração do arcabouço interpretativo da mecânica quântica, e até mesmo o seu questionamento com proposta de alternativas, continua de fato até os dias de hoje, sem que no entanto os resultados de suas aplicações tenham tido que ser de qualquer forma revistos.

2.1 Estrutura algébrica da teoria de Heisenberg: discussão heurística.

Antes de passar a uma formulação mais completa da estrutura formal subjacente à teoria de Heisenberg, pode ser útil ter um panorama geral do que está envolvido e do que deve estar contido em tal formulação. Os ingredientes necessários para caracterizar esse panorama fazem parte do arsenal padrão da álgebra linear em espaços vetoriais sobre números complexos. De fato, matrizes complexas como as que aí aparecem representam operadores lineares nesses espaços. Nesta seção esse arsenal vai ser tratado de forma mais ou menos incidental e incompleta, na estrita medida do indispensável, para não eclipsar o foco principal da discussão que é a estrutura algébrica da teoria quântica. Um tratamento mais completo aparece na seção seguinte.

Como ponto de partida, vamos supor portanto que a teoria seja formulada em termos de um espaço vetorial (em geral de dimensão infinita) sobre os números complexos, no qual está definido um produto escalar hermiteano. Isso significa que, designando os vetores desse espaço como φ , ψ etc. o produto escalar de dois vetores (ψ, φ) é um número complexo que, em particular, tem a propriedade

$$(\psi, \varphi) = (\varphi, \psi)^*, \quad (2.1)$$

onde o asterisco indica conjugação complexa. Dessa propriedade resulta que o produto escalar de um vetor consigo mesmo é um número real, que pode ser tomado como o quadrado da *norma* desse vetor. O anulamento do produto escalar de dois vetores, nenhum dos quais tendo norma nula, corresponde à *ortogonalidade* dos dois vetores.

Qualquer vetor φ desse espaço pode ser expandido em uma *base* de (em geral infinitos) vetores $\{\phi_n\}$, $n = 1, 2, \dots$, mutuamente ortogonais e de norma igual a 1, isto é, tais que

$$(\phi_n, \phi_m) = \delta_{nm},$$

como

$$\varphi = \sum_n c_n \phi_n \quad (2.2)$$

onde os coeficientes c_n são números complexos que podem ser expressos em termos do produto escalar como (ϕ_n, φ) . Uma base desse tipo é chamada *base ortonormal*, o que significa que ela é constituída de vetores ortogonais entre si e de norma 1.

Uma classe de objetos de crucial importância é a de *operadores lineares* definidos sobre o espaço vetorial. Um tal operador age sobre vetores do espaço dando como resultado outros vetores do mesmo espaço, e satisfaz condições de linearidade. Isso significa que, se g é um tal operador, então um objeto como $g\varphi$ é um vetor, e que vale a propriedade

$$g(z_1\varphi_1 + z_2\varphi_2) = z_1 g\varphi_1 + z_2 g\varphi_2$$

onde z_1 e z_2 são números complexos quaisquer. Paralelamente ao fato de que a escolha de uma base ortonormal permite exprimir um vetor φ em termos de uma coleção de componentes $\{c_n\}$ como em (2.2), ela permite também exprimir um operador linear g em termos de uma matriz $\{g_{nm}\}$. De fato, a componente n do vetor $g\varphi$ é $d_n = (\phi_n, g\varphi)$, e escrevendo φ também em termos da base escolhida na forma (2.2) resulta, usando a linearidade de g , que

$$d_n = \sum_m (\phi_n, g\phi_m) c_m \equiv \sum_m g_{nm} c_m.$$

Um operador linear g é dito *hermiteano* se ele satisfaz

$$(\psi, g\varphi) = (g\psi, \varphi)$$

para quaisquer dois vetores do espaço. Ela deve valer, em particular, para quaisquer dois vetores de qualquer base ortonormal que seja escolhida. Usando a propriedade (2.1) do produto escalar, é fácil verificar que as matrizes que representam operadores hermiteanos em bases ortonormais são sempre tais que $g_{nm} = g_{mn}^*$.

As matrizes de Born e Jordan podem portanto ser vistas como representações de operadores lineares hermiteanos de *um certo espaço vetorial, numa certa base ortonormal*, com o que a revisão cinemática levada a cabo por Heisenberg passa a ser vista como a substituição das variáveis dinâmicas clássicas (em geral funções reais das coordenadas e momentos que definem o estado do sistema) por operadores lineares hermiteanos sobre um espaço vetorial complexo, representados numa determinada base. Esse ponto de vista deixa duas questões em aberto, que consistem respectivamente em especificar o espaço de que se trata, e a base a ser usada. Elas são porém de ordens diferentes, e podem ser tratadas independentemente uma da outra. Enquanto

especificar o espaço de que se trata deve significar entender de alguma forma os vetores de que ele consiste, especificar a base é uma questão que pode ser tratada independentemente de tal entendimento, e que de fato é praticamente respondida pela condição de que a matriz que representa a hamiltoniana seja *diagonal*. Sendo esta uma representação de um operador linear hermiteano H , essa condição significa que os vetores ϕ_n que constituem a base ortonormal escolhida são na realidade, independentemente do sentido que possam ter, *autovetores* desse operador, isto é

$$H\phi_n = E_n\phi_n. \quad (2.3)$$

De fato, a propriedade de hermiticidade de H garante eventualmente a existência de um conjunto de tais autovetores que seja uma base ortonormal no espaço em que H age, além de garantir a realidade dos autovalores E_n .

Esse ponto de vista permite então formular o esquema de Heisenberg, Born e Jordan da seguinte forma:

1. As variáveis canônicas clássicas devem ser substituídas por operadores lineares hermiteanos q e p , agindo sobre um espaço vetorial (de dimensão infinita), os quais devem satisfazer a relação de comutação $[q, p] = i\hbar\hat{1}$, onde $\hat{1}$ é o operador identidade do espaço vetorial (isto é, o operador cuja ação sobre qualquer vetor φ do espaço leva ao próprio vetor, $\hat{1}\varphi = \varphi$).

2. O hamiltoniano é um operador linear hermiteano obtido da hamiltoniana clássica através da substituição das variáveis canônicas clássicas pelos operadores q e p .

3. Os valores possíveis da energia dos estados estacionários são os autovalores E_n de H , obtidos da equação (2.3).

4. O “momento de dipolo” associado a transições entre os estados estacionários de energia E_m e E_n corresponde ao elemento de matriz $q_{nm} \equiv (\phi_n, q\phi_m)$, onde ϕ_m e ϕ_n são autovetores de H correspondentes a esses dois autovalores.

O esquema de Born e Jordan corresponde a essa formulação na situação particular em que os operadores são representados na base de autovetores de H . As matrizes que representam q e p são então $q_{nm} = (\phi_n, q\phi_m)$ e $p_{nm} = (\phi_n, p\phi_m)$ e a que representa o hamiltoniano é diagonal, $H_{nm} = E_n \delta_{nm}$. A formulação mais geral, em termos de operadores, tem no entanto a característica importante de não conter qualquer referencia essencial a uma base específica. Esta *independência de base* foi amplamente explorada na formulação geral da mecânica quântica desenvolvida por Dirac e utilizada constantemente até hoje.

É interessante estudar também a dinâmica (isto é, a dependência temporal das variáveis dinâmicas) dentro deste ponto de vista mais geral. A equação de movimento matricial (1.17) se torna uma equação de movimento para operadores

$$\frac{dg}{dt} = \frac{i}{\hbar}[H, g]$$

e é fácil verificar que essa equação admite a solução formal

$$g(t) = e^{\frac{i}{\hbar}Ht}g(0)e^{-\frac{i}{\hbar}Ht} \quad (2.4)$$

onde $g(0)$ é a condição inicial para o operador $g(t)$ em $t = 0$. Para dar sentido às exponenciais que contêm o operador H no expoente basta interpreta-las em termos da série de potências, e.g.

$$e^{\frac{i}{\hbar}Ht} = \hat{1} + \frac{i}{\hbar}Ht + \frac{1}{2!} \left(\frac{i}{\hbar}Ht \right)^2 + \dots$$

onde as potências do operador H são também operadores interpretados de forma natural como $H^2\varphi = H(H\varphi)$, etc.

Nos casos em que $g(0)$ não comuta com H , não é possível rearranjar os fatores da solução formal de modo a simplesmente cancelar as duas exponenciais. Quando $[H, g(0)] = 0$, no entanto, esse cancelamento é possível, e a variável dinâmica correspondente ao operador $g(0)$ é uma constante do movimento (independente do tempo). O próprio hamiltoniano H é claramente uma tal constante do movimento, o que faz com que os seus autovetores e autovalores, Eq. (2.3), sejam também independentes do tempo. A matriz de Born e Jordan correspondente ao operador $g(t)$, incluindo a dependência temporal, pode, no caso geral, ser facilmente recuperada dessas equações entre operadores usando a base constituída pelos autovetores do hamiltoniano. De fato,

$$g_{nm}(t) = (\phi_n, e^{\frac{i}{\hbar}Ht}g(0)e^{-\frac{i}{\hbar}Ht}\phi_m) = e^{\frac{i}{\hbar}(E_n - E_m)t}g_{nm}(0) \quad (2.5)$$

onde, para obter a última expressão, foi usada a expansão em série das exponenciais e a propriedade de hermiticidade de H , além da equação (2.3). Uma consequência importante que pode ser inferida dessa expressão geral é que, no caso particular em que $g(0)$ é uma constante do movimento (isto é, no caso em que esse operador comuta com H), os elementos de matriz $g_{nm}(0) = (\phi_n, g(0)\phi_m)$ devem ser *nulos* sempre que $E_n \neq E_m$, pois do contrário a matriz $\{g_{nm}(t)\}$ dependeria do tempo.

Finalmente, é interessante observar que, definindo a partir dos autovetores ϕ_n de H os vetores *dependentes do tempo*

$$\tilde{\phi}_n(t) \equiv e^{-\frac{i}{\hbar}Ht}\phi_n = e^{-\frac{i}{\hbar}E_nt}\phi_n, \quad (2.6)$$

é possível obter os mesmos elementos de matriz $g_{nm}(t)$ como

$$g_{nm}(t) = \left(\tilde{\phi}_n(t), g(0)\tilde{\phi}_m(t) \right). \quad (2.7)$$

Enquanto em (2.5) eles aparecem como elementos de matriz do operador dependente do tempo $g(t)$ calculados com os vetores independentes do tempo ϕ_n, ϕ_m , em (2.7) eles aparecem expressos como elementos de matriz do operador independente do tempo $g(0)$ calculados com os vetores dependentes do tempo $\tilde{\phi}_n(t), \tilde{\phi}_m(t)$. Portanto, na medida em que quantidades observáveis associadas a sistemas quânticos possam ser sempre expressas em termos de tais elementos de matriz, existe uma forma de implementar a dinâmica, alternativa à que foi efetivamente usada por Heisenberg, Born e Jordan, e que leva aos mesmos resultados. Essa forma alternativa consiste em associar as variáveis dinâmicas a operadores *independentes do tempo*, e associar a dependência temporal das quantidades observáveis à dependência temporal dos *vetores*. Esta deve ser tal que os autovetores de H evoluem no tempo como prescrito em (2.6). A evolução temporal de uma combinação linear qualquer dos vetores dessa base, Eq. (2.2), deve ser então

$$\tilde{\varphi}(t) = \sum_n c_n e^{-\frac{i}{\hbar} E_n t} \phi_n = e^{-\frac{i}{\hbar} H t} \tilde{\varphi}(0) \quad (2.8)$$

que corresponde à equação de movimento *para os vetores*

$$i\hbar \frac{d\tilde{\varphi}(t)}{dt} = H\tilde{\varphi}(t).$$

Esta forma de implementar a dinâmica dos sistemas quânticos foi efetivamente a que foi adotada por Schrödinger, evidentemente por caminhos muito diferentes do que foi usado aqui. Este tem a ver mais com a estrutura geral comum, subjacente às teorias de Heisenberg e Schrödinger, como exposta por Dirac.

A formulação da teoria quântica em termos de um espaço vetorial passa necessariamente, como foi visto, pela introdução explícita dos vetores como uma classe de objetos relevantes. Esses objetos não aparecem explicitamente na formulação de Heisenberg, Born e Jordan, e por isso permaneceram desprovidos de interpretação na discussão precedente. A questão de 'entender de alguma forma os vetores' de que consiste o espaço vetorial utilizado é na realidade uma questão de natureza interpretativa que não foi ainda respondida, e que será tratada como tal na seção seguinte.

2.2 Estrutura formal e interpretação da teoria quântica.

Esta seção retoma a discussão da estrutura geral da teoria quântica de forma mais completa e precisa que a seção anterior, e trata também das questões interpretativas que foram deixadas

ali em aberto. Suas três partes iniciais tratam dos aspectos cinemáticos da teoria (isto é, de questões como os operadores que representam variáveis dinâmicas e o espaço vetorial em que eles são definidos) bem como da interpretação dos vários ingredientes e de conseqüências gerais da estrutura algébrica adotada, entre as quais estão, por exemplo, as chamadas relações de incerteza. As duas partes seguintes tratam dos aspectos dinâmicos da teoria, em particular de diferentes formas de implementar a dinâmica, de algumas conseqüências gerais das equações de movimento e finalmente de alguns aspectos relativos a processos de medida no âmbito da teoria quântica.

2.2.1 Espaços de Hilbert.

Como na seção anterior, começamos por supor que a teoria quântica seja formulada em termos de um espaço vetorial complexo, em geral de dimensão infinita, munido de um produto escalar hermiteano. A partir deste ponto, para escrever os vetores desse espaço vamos utilizar uma notação muito conveniente, devida a Dirac, na qual um vetor genérico é indicado pelo símbolo $|\rangle$. Vetores específicos, designados por rótulos identificadores como φ , ψ_1 , ψ_2 , etc. aparecem então sob a forma $|\varphi\rangle$, $|\psi_1\rangle$, $|\psi_2\rangle$, etc.

A estrutura de espaço vetorial significa que estão definidas operações de soma de vetores e de produto de um vetor por um número complexo z . A definição dessas operações é tal que

a) A soma de dois vetores associa a dois vetores quaisquer $|\varphi\rangle$ e $|\psi\rangle$ dados um terceiro vetor $|\chi\rangle = |\varphi\rangle + |\psi\rangle$ e satisfaz as propriedades

- a1)** associativa: $|\varphi\rangle + (|\psi_1\rangle + |\psi_2\rangle) = (|\varphi\rangle + |\psi_1\rangle) + |\psi_2\rangle$;
- a2)** comutativa: $|\varphi\rangle + |\psi_1\rangle = |\psi_1\rangle + |\varphi\rangle$;
- a3)** existe um vetor nulo $|\emptyset\rangle$ tal que $|\varphi\rangle + |\emptyset\rangle = |\varphi\rangle$ para todo vetor $|\varphi\rangle$;
- a4)** para todo vetor $|\varphi\rangle$ existe um outro vetor $|- \varphi\rangle$ tal que $|\varphi\rangle + |- \varphi\rangle = |\emptyset\rangle$.

b) O produto de um vetor $|\varphi\rangle$ por um número complexo z tem como resultado um vetor $z|\varphi\rangle$ e satisfaz as propriedades

- b1)** “associativa”: $(z_1 z_2)|\varphi\rangle = z_1(z_2|\varphi\rangle)$;
- b2)** e **b3)** distributivas:

$$(z_1 + z_2)|\varphi\rangle = z_1|\varphi\rangle + z_2|\varphi\rangle \quad \text{e} \quad z(|\varphi\rangle + |\psi\rangle) = z|\varphi\rangle + z|\psi\rangle;$$

b4) o produto de qualquer vetor por $z = 1$ reproduz o mesmo vetor, $1|\varphi\rangle = |\varphi\rangle$ para qualquer $|\varphi\rangle$.

O espaço vetorial a ser usado é também munido de um produto escalar, que é uma operação que associa a todo par de vetores $|\varphi\rangle, |\phi\rangle$, um número complexo, o qual será indicado pelo símbolo (φ, ϕ) , satisfazendo as propriedades

- p1)** o produto escalar de um vetor $|\varphi\rangle$ com a soma de dois outros $|\chi\rangle = |\psi_1\rangle + |\psi_2\rangle$ é igual à soma dos produtos escalares desse vetor com cada um dos vetores envolvidos na soma: $(\varphi, \chi) = (\varphi, \psi_1) + (\varphi, \psi_2)$;
- p2)** $(\varphi, z\psi) = z(\varphi, \psi)$, sendo z um número complexo;
- p3)** $(\varphi, \psi) = (\psi, \varphi)^*$, onde o asterisco indica conjugação complexa;
- p4)** para todo vetor $|\varphi\rangle$, $0 \leq (\varphi, \varphi) < \infty$, sendo que a igualdade a zero se dá se e somente se $|\varphi\rangle = |\emptyset\rangle$.

É claro que as propriedades **p2)** e **p3)** implicam ainda no fato importante que $(z\varphi, \psi) = z^*(\varphi, \psi)$. Juntamente com a propriedade **p1)**, elas implicam a *linearidade* do produto escalar com relação ao segundo argumento e o que se chama a *antilinearidade* com relação ao primeiro, o que corresponde às relações

$$(\varphi, z_1\psi_1 + z_2\psi_2) = z_1(\varphi, \psi_1) + z_2(\varphi, \psi_2) \quad \text{e} \quad (z_1\psi_1 + z_2\psi_2, \varphi) = z_1^*(\psi_1, \varphi) + z_2^*(\psi_2, \varphi).$$

O conjunto dessas duas propriedades é usualmente designado como a *sesquilinearidade* do produto escalar. A existência do produto escalar permite definir uma *norma* para os vetores do espaço como

$$\text{norma de } |\varphi\rangle \equiv \|\varphi\| = (\varphi, \varphi)^{1/2}. \quad (2.9)$$

Uma norma para os vetores de um espaço vetorial é definida em geral como uma função $\|\varphi\|$ com valores reais, tendo como argumento um vetor $|\varphi\rangle$ do espaço e satisfazendo as propriedades

- n1)** $0 \leq \|\varphi\| < \infty$ para qualquer vetor $|\varphi\rangle$, sendo que a igualdade a zero equivale a $|\varphi\rangle = |\emptyset\rangle$;
- n2)** $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$ para quaisquer $|\varphi\rangle, |\psi\rangle$;
- n3)** $\|z\varphi\| = |z|\|\varphi\|$, onde z é um número complexo e $|z|$ é o seu módulo, para qualquer $|\varphi\rangle$.

É fácil verificar que todas essas propriedades decorrem das propriedades do produto escalar quando se usa a definição (2.9) para a norma. Isso significa que a existência de um produto escalar *sempre* disponibiliza uma norma. A recíproca evidentemente não vale, já mesmo porque o produto escalar está definido para *pares* de vetores diferentes. Assim, o produto escalar implementa, por exemplo, a idéia “geométrica” de *ortogonalidade* de dois vetores

(através do anulamento de seu produto escalar), o que é evidentemente inacessível a partir da definição de uma norma independentemente do produto escalar. Nesse sentido, um espaço vetorial no qual está definida apenas uma norma para os vetores é “menos estruturado” que um espaço vetorial munido de um produto escalar.

Uma classe importante de objetos que é possível definir em um espaço vetorial complexo é a classe de todas as funções lineares, com valores complexos, cujo argumento é um vetor do espaço. A notação genérica introduzida por Dirac para esses objetos é $\langle \mid$. Funções lineares específicas, designadas por rótulos identificadores como a , b , etc. aparecem sob a forma $\langle a|$, $\langle b|$, etc. Dessa forma, o resultado da ação da função linear $\langle a|$ sobre o vetor $|\varphi\rangle$ será um número complexo, representado por $\langle a|\varphi\rangle$. A linearidade das $\langle \mid$ significa que

$$\langle \mid (z_1|\varphi\rangle + z_2|\psi\rangle) = z_1\langle \mid \varphi\rangle + z_2\langle \mid \psi\rangle,$$

isto é, o número complexo que resulta da aplicação de uma função linear qualquer a uma combinação linear de vetores é a mesma combinação linear dos números complexos que resultam da aplicação dessa função a cada um dos dois vetores separadamente.

Sobre o conjunto das funções lineares pode-se definir, de forma natural, a soma (associativa e comutativa) de duas funções e o produto de uma função por um número complexo respectivamente através das relações

$$(\langle a| + \langle b|) |\varphi\rangle = \langle a|\varphi\rangle + \langle b|\varphi\rangle \quad \text{e} \quad (z\langle a|) |\varphi\rangle = z\langle a|\varphi\rangle;$$

é também possível definir uma função nula $\langle \emptyset|$ através da relação $\langle \emptyset|\varphi\rangle = 0$ válida para qualquer vetor $|\varphi\rangle$; e ainda associar a cada função $\langle a|$ uma função $\langle -a|$ através de $\langle -a|\varphi\rangle = -\langle a|\varphi\rangle$. Disso resulta então que $\langle a| + \langle -a| = \langle \emptyset|$. Com essas definições resulta então que o conjunto das funções lineares sobre o espaço vetorial considerado tem por sua vez uma estrutura de espaço vetorial. Este novo espaço vetorial é chamado *espaço dual* do espaço de partida.

Um resultado importante é a possibilidade de estabelecer uma correspondência biunívoca (isto é, de um para um) entre os vetores do espaço de partida e os vetores (funções lineares) do espaço dual, através do produto escalar definido no espaço de partida. De fato, para estabelecer essa correspondência basta associar a cada vetor $|\varphi\rangle$ a função linear, a ser designada como $\langle \varphi|$, tal que o seu valor em qualquer vetor $|\psi\rangle$ seja igual ao produto escalar de $|\varphi\rangle$ e $|\psi\rangle$. Mais formalmente, a correspondência $|\varphi\rangle \leftrightarrow \langle \varphi|$ significa que

$$(\varphi, \psi) = \langle \varphi|\psi\rangle$$

para qualquer vetor $|\psi\rangle$ (nesta última relação, no primeiro membro φ denota o vetor $|\varphi\rangle$, enquanto no segundo membro aparece a função linear $\langle \varphi|$ associada a esse vetor). O fato

de que a função $\langle\varphi|$ associada desse modo ao vetor $|\varphi\rangle$ é única pode ser verificado notando que, se uma função linear $\langle\varphi'|$ é também tal que $(\varphi, \psi) = \langle\varphi'|\psi\rangle$ para qualquer $|\psi\rangle$, então $\langle\varphi| - \langle\varphi'| = \langle\varphi| + \langle-\varphi'| = \langle\emptyset|$, que corresponde à igualdade de $\langle\varphi|$ e $\langle\varphi'|$. É fácil também verificar que a função linear que corresponde ao vetor nulo $|\emptyset\rangle$ é a função nula $\langle\emptyset|$, isto é, $|\emptyset\rangle \leftrightarrow \langle\emptyset| \equiv \langle\emptyset|$.

A correspondência biunívoca entre vetores e funções lineares obtida através do produto escalar de vetores permite o uso da notação $\langle\varphi|\psi\rangle$ para indicar o produto escalar dos vetores $|\varphi\rangle$ e $|\psi\rangle$. Dirac chamou “bras” os vetores do espaço dual e “kets” os vetores do espaço vetorial de partida, de modo que usando essa notação os produtos escalares aparecem representados por “brackets” (parênteses) do tipo $\langle\varphi|\psi\rangle$, que correspondem a números complexos. Essa correspondência é *antilinear*, no sentido que a correspondência entre combinações lineares de “kets” e “bras” aparece como

$$z_1|\varphi\rangle + z_2|\psi\rangle \leftrightarrow z_1^*\langle\varphi| + z_2^*\langle\psi|,$$

o que pode ser facilmente deduzido da sesquilinearidade do produto escalar.

Para ilustrar o uso da notação de Dirac vamos usa-la para enunciar e demonstrar uma consequência geral importante da definição de produto escalar, conhecida como *desigualdade de Schwartz*. Ela afirma que, para dois vetores quaisquer $|\varphi\rangle$ e $|\psi\rangle$, vale a relação

$$|\langle\varphi|\psi\rangle|^2 \leq \langle\varphi|\varphi\rangle\langle\psi|\psi\rangle$$

que corresponde a $|(\varphi, \psi)|^2 \leq (\varphi, \varphi)(\psi, \psi)$ na notação introduzida inicialmente para o produto escalar. Para demonstrar essa desigualdade basta notar inicialmente que, em vista da propriedade **p4**,

$$\langle\varphi + z\psi|\varphi + z\psi\rangle \geq 0,$$

o que se escreve também como

$$\langle\varphi|\varphi\rangle + z\langle\varphi|\psi\rangle + z^*\langle\psi|\varphi\rangle + |z|^2\langle\psi|\psi\rangle \geq 0$$

onde foi feito uso da linearidade dos bras e da antilinearidade da correspondência entre kets e bras. Se $|\psi\rangle$ é o vetor nulo, então a desigualdade de Schwartz é válida (como igualdade), pois os seus dois lados se anulam. Se no entanto $|\psi\rangle$ não é o vetor nulo, tomando $z = -\langle\psi|\varphi\rangle/\langle\psi|\psi\rangle$ e usando também $\langle\psi|\varphi\rangle^* = \langle\varphi|\psi\rangle$ resulta

$$\langle\varphi|\varphi\rangle - 2\frac{|\langle\varphi|\psi\rangle|^2}{\langle\psi|\psi\rangle} + \frac{|\langle\varphi|\psi\rangle|^2}{\langle\psi|\psi\rangle} \geq 0$$

donde segue também a desigualdade de Schwartz.

Um espaço vetorial complexo (de dimensão infinita), munido de um produto escalar que define, em particular, a norma dos vetores que o constituem, é chamado *espaço de Hilbert* quando dotado ainda da propriedade adicional de *completeza*. Essa propriedade consiste em que toda seqüência infinita de vetores $\{|\psi_i\rangle\}$, $i = 1, 2, \dots$, tal que $\| |\psi_i\rangle - |\psi_j\rangle \| < \epsilon$, com ϵ arbitrariamente pequeno, desde que i e j sejam maiores que algum $N(\epsilon)$ (tais seqüências são chamadas *seqüências de Cauchy*), tem um limite $|\psi\rangle$ que é um vetor pertencente ao espaço. No que segue vamos supor também que o espaço a ser considerado é tal que admite bases constituídas de um conjunto enumerável de vetores, $\{\phi_n\}$, $n = 1, 2, \dots$, (que é uma propriedade conhecida como *separabilidade*). Um espaço de Hilbert separável será então tomado como o substrato natural sobre o qual pode ser desenvolvida a cinemática e a dinâmica da teoria quântica.

Além da caracterização abstrata de espaços de Hilbert que acaba de ser descrita, é útil dispor de um exemplo, realizado em termos de objetos relativamente familiares. Um exemplo relevante no contexto da Mecânica Quântica é o espaço cujos vetores são funções complexas $f(x)$ de uma variável real x e de quadrado integrável, isto é

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx < \infty .$$

A soma de vetores e o produto de um vetor por um número complexo podem ser tomadas como as definições usuais de soma de funções e de produto de uma função por um número complexo, e um produto escalar pode ser definido como

$$(f_1, f_2) \equiv \int_{-\infty}^{+\infty} f_1^*(x) f_2(x) dx .$$

Esse produto escalar leva à definição da norma de uma função $f(x)$ como

$$\|f\|^2 = \int_{-\infty}^{+\infty} |f(x)|^2 dx$$

e a desigualdade de Schwartz assume a forma

$$\left| \int_{-\infty}^{+\infty} f_1^*(x) f_2(x) dx \right|^2 \leq \int_{-\infty}^{+\infty} |f_1(x)|^2 dx \int_{-\infty}^{+\infty} |f_2(x)|^2 dx .$$

2.2.2 Operadores lineares.

Em vista de que variáveis dinâmicas são realizadas, na mecânica quântica, em termos de operadores lineares particulares do espaço de Hilbert que serve de substrato à cinemática da teoria,

esta seção é dedicada ao exame das propriedades mais relevantes dessa classe de objetos. O tratamento desenvolvido a seguir é suficientemente geral apenas no caso de espaços vetoriais de dimensão *finita*, que já se revelou insuficiente para tratar até mesmo o problema simples do oscilador harmônico quântico (v. seção 1.3.2). Isso significa que situações especiais, exigindo um tratamento mais completo e mais preciso, *podem* ocorrer no caso de espaços de dimensão infinita. Para evitar a parafernália maior de definições e cautelas, freqüentemente dispensável, do tratamento mais completo, vamos nos limitar a indicar ao longo do texto alguns dos pontos críticos em que tal tratamento pode se tornar necessário, indicando entre as referências gerais citadas no fim do capítulo textos em que ele pode ser encontrado.

Um operador linear g num espaço de Hilbert transforma vetores desse espaço em outros vetores do mesmo espaço, $g|\varphi\rangle = |\psi\rangle$, e se comporta com relação às operações lineares envolvendo vetores como

$$g(z_1|\varphi_1\rangle + z_2|\varphi_2\rangle) = z_1 g|\varphi_1\rangle + z_2 g|\varphi_2\rangle.$$

Uma base ortonormal enumerável $\{|\phi_n\rangle\}$, $n = 1, 2, \dots$ do espaço permite representar vetores do espaço em termos da coleção de suas componentes nessa base e um operador linear em termos de uma matriz determinada através da ação do operador sobre cada um dos elementos da base. De fato, usando as expansões

$$|\varphi\rangle = \sum_n c_n |\phi_n\rangle \quad \text{e} \quad |\psi\rangle = \sum_n d_n |\phi_n\rangle,$$

que identificam as componentes dos vetores $|\varphi\rangle$ e $|\psi\rangle$, a ação de g sobre $|\varphi\rangle$ pode ser representada como

$$d_n = \sum_m g_{nm} c_m$$

onde os vetores $|\varphi\rangle$ e $|\psi\rangle$ são representados por suas respectivas componentes na base ortonormal escolhida, $\{c_n\}$ e $\{d_n\}$, $n = 1, 2, \dots$ e $g_{nm} = \langle\phi_n|(g|\phi_m\rangle)$. Devido à ortonormalidade dos vetores da base, a norma dos vetores pode ser expressa em termos de suas componentes como

$$\|\varphi\| = \langle\varphi|\varphi\rangle^{\frac{1}{2}} = \left(\sum_n |c_n|^2\right)^{\frac{1}{2}}$$

e uma expressão análoga para $\|\psi\|$. A notação $\langle\phi_n|(g|\phi_m\rangle)$ usada para os elementos da matriz que g_{nm} indica explicitamente que eles correspondem a tomar o produto escalar do vetor que resulta da aplicação de g a $|\phi_m\rangle$ com o vetor $|\phi_n\rangle$. Ela é usualmente aliviada notando que a ação de um bra (função linear) $\langle\phi_n|$ dado sobre o resultado da aplicação de um operador linear

g também dado a um ket genérico $|\rangle$, na realidade define uma nova função linear sobre os kets, pois resulta em um número complexo. Denotando essa nova função linear por $(\langle\phi_n|g)$ resulta que

$$(\langle\phi_n|g)|\rangle \equiv \langle\phi_n|(g|\rangle)$$

o que permite ignorar os parênteses. Dessa forma, os elementos da matriz que representa o operador g na base escolhida podem ser escritos simplesmente como $g_{nm} = \langle\phi_n|g|\phi_m\rangle$.

No caso de um espaço vetorial de dimensão finita, a ação de um operador linear sobre qualquer vetor do espaço estará bem definida sempre que a sua ação sobre cada um dos vetores de uma base esteja bem definida. Isso no entanto pode não ser verdade no caso de um espaço de dimensão infinita. Um exemplo do tipo de problema que pode ocorrer nesse caso pode ser dado usando a representação matricial do hamiltoniano do oscilador harmônico como escrita na Eq. (1.24), construída em termos da base ortonormal (infinita mas enumerável) formada pelos autovetores do hamiltoniano. Seja então o vetor representado nessa mesma base pelas componentes $c_m = \sqrt{6}/[(m+1)\pi]$, $m = 0, 1, 2, \dots$. Esse vetor tem norma 1, o que decorre de que

$$\sum_{n=0}^{\infty} \frac{1}{(n+1)^2} = \frac{\pi^2}{6}.$$

A ação do hamiltoniano sobre esse vetor dá

$$d_n = \sum_{m=0}^{\infty} H_{nm} c_m = \sum_{m=0}^{\infty} \hbar\omega_0 \left(n + \frac{1}{2}\right) \delta_{nm} \frac{\sqrt{6}}{(m+1)\pi} = \hbar\omega_0 \frac{\sqrt{6}}{\pi} \left(1 - \frac{1}{2(n+1)}\right)$$

o que não corresponde a um vetor do espaço, pois a quantidade correspondente à norma diverge, $\sum_{n=0}^{\infty} |d_n|^2 = \infty$. Operadores desse tipo, isto é, cuja ação sobre vetores de norma 1 podem levar a imagens com norma arbitrariamente grande, são chamados operadores *não limitados*. Uma característica importante desses operadores, que precisa ser levada em conta para que o seu tratamento seja tornado mais preciso, é a classe de vetores sobre os quais sua ação está bem definida, que é em geral um subconjunto de vetores que não coincide com o espaço todo, e que é chamado *domínio* do operador. O vetor acima, representado pelas componentes c_m , está fora dessa classe no caso do hamiltoniano do oscilador harmônico.

A soma e o produto de operadores lineares pode ser definida de forma natural através das relações

$$(g + h)|\varphi\rangle = g|\varphi\rangle + h|\varphi\rangle \quad \text{e} \quad (gh)|\varphi\rangle = g(h|\varphi\rangle)$$

e correspondem também a operadores lineares. O produto de operadores é não comutativo, pois em geral $g(h|\varphi\rangle) \neq h(g|\varphi\rangle)$, o que equivale a dizer que o comutador de dois operadores lineares (que é por sua vez um operador linear) é em geral diferente do operador nulo $\hat{0}$ (que leva qualquer vetor ao vetor nulo), $[g, h] \equiv gh - hg \neq \hat{0}$. O produto de um operador por um número complexo também é definido simplesmente como $(zg)|\varphi\rangle = z(g|\varphi\rangle)$.

Para alguns operadores g é possível obter operadores correspondentes \bar{g} tais que $g\bar{g} = \bar{g}g = \hat{1}$, onde $\hat{1}$ é o operador unidade, que associa qualquer vetor do espaço a si próprio. Nesses casos \bar{g} é chamado operador *inverso* de g , escrito como $\bar{g} \equiv g^{-1}$. É fácil verificar que se g e h são dois operadores que admitem inversos g^{-1} e h^{-1} , então o produto gh também admite um inverso $(gh)^{-1}$ dado por

$$(gh)^{-1} = h^{-1}g^{-1}.$$

Em termos da representação por matrizes, a soma, o produto e o inverso de operadores são dados em termos das operações *matriciais* correspondentes.

De forma até certo ponto paralela à usada para estabelecer uma correspondência entre kets e bras através do produto escalar, é possível associar a cada operador g outro operador g^\dagger , chamado *operador adjunto* de g , através da relação

$$\langle\varphi|g^\dagger|\psi\rangle = \langle\psi|g|\varphi\rangle^* \quad (2.10)$$

para quaisquer vetores $|\varphi\rangle$ e $|\psi\rangle$. Em termos de uma representação matricial, é claro que a matriz que representa o operador adjunto g^\dagger é a complexo-conjugada da transposta da matriz que representa g , isto é, $(g^\dagger)_{nm} = \langle\phi_n|g^\dagger|\phi_m\rangle = \langle\phi_m|g|\phi_n\rangle^* = g_{mn}^*$. Com essa definição, o bra $\langle\varphi'|$ associado ao ket $|\varphi'\rangle = g|\varphi\rangle$, e que portanto satisfaz com qualquer $|\psi\rangle$ a relação

$$\langle\varphi'|\psi\rangle = \langle\psi|\varphi'\rangle^* = \langle\psi|g|\varphi\rangle^* = \langle\varphi|g^\dagger|\psi\rangle,$$

pode ser escrito como $\langle\varphi'| = \langle\varphi|g^\dagger$. A definição de operador adjunto permite ainda verificar facilmente as propriedades

- c1)** $(g^\dagger)^\dagger = g$; de fato, $\langle\varphi|(g^\dagger)^\dagger|\psi\rangle = \langle\psi|g^\dagger|\varphi\rangle^* = \langle\varphi|g|\psi\rangle$ para quaisquer vetores $|\varphi\rangle$ e $|\psi\rangle$;
- c2)** $(g + h)^\dagger = g^\dagger + h^\dagger$;
- c3)** $(gh)^\dagger = h^\dagger g^\dagger$; de fato, $\langle\varphi|(gh)^\dagger|\psi\rangle = \langle\psi|gh|\varphi\rangle^* = \langle\varphi|h^\dagger g^\dagger|\psi\rangle$ para quaisquer vetores $|\varphi\rangle$ e $|\psi\rangle$;
- c4)** $(zg)^\dagger = z^* g^\dagger$; o operador zg é definido de forma natural como o produto do operador “multiplicação pelo número complexo z ” com g .

Um operador g é dito *hermiteano* se $g^\dagger = g$, e *anti-hermiteano* se $g^\dagger = -g$. Operadores hermiteanos e anti-hermiteanos são em muitos aspectos semelhantes a quantidades reais e imaginárias puras, respectivamente, o que significa também que a definição de operador adjunto é em muitos aspectos semelhante a uma conjugação complexa para operadores. Assim, por exemplo, $\langle \varphi | g^\dagger | \varphi \rangle = \langle \varphi | g | \varphi \rangle^* = \pm \langle \varphi | g | \varphi \rangle$, onde o duplo sinal corresponde a $g^\dagger = \pm g$.

A definição (2.10) de operador adjunto apresenta problemas no caso de operadores não limitados, quando a ação do operador g não está bem definida para todos os vetores do espaço. De fato, nesse caso o lado direito de (2.10) não estará bem definido para qualquer vetor $|\varphi\rangle$ do espaço, mas apenas para vetores pertencentes ao domínio de g . Essa dificuldade pode ser contornada considerando o operador adjunto g^\dagger definido apenas para vetores $|\psi\rangle$ tais que a equação

$$\langle \psi | g | \varphi \rangle^* = \langle \chi | \varphi \rangle^*$$

admita uma única solução $|\chi\rangle$ para qualquer $|\varphi\rangle$ pertencente ao domínio de g , sendo a definição de g^\dagger completada então com

$$g^\dagger |\psi\rangle = |\chi\rangle.$$

Dessa forma, o operador g^\dagger estará por sua vez definido apenas em um certo domínio de vetores $|\psi\rangle$ condicionado pela existência da solução $|\chi\rangle$. Esses fatos devem ser levados em conta também na definição de operadores hermiteanos, $g^\dagger = g$, pois embora coincidindo na sua ação sobre vetores para os quais essa ação esteja definida, esses dois operadores poderão ainda diferir quanto aos respectivos domínios. É comum reservar para o caso em que tanto as ações quanto os domínios de g^\dagger e g coincidam a denominação de operador *auto-adjunto*.

A definição do adjunto de um operador permite caracterizar duas classes extremamente importantes de operadores lineares. A primeira delas é a dos chamados *operadores unitários*, identificados pela propriedade $u^\dagger = u^{-1}$ ou equivalentemente

$$uu^\dagger = u^\dagger u = \hat{1}.$$

A ação de operadores unitários preserva o produto escalar (e portanto também a norma) de vetores. De fato,

$$(\langle \psi | u^\dagger)(u | \varphi \rangle) = \langle \psi | u^\dagger u | \varphi \rangle = \langle \psi | \varphi \rangle.$$

Disso resulta, em particular, que um operador unitário transforma uma base ortonormal em outra base ortonormal do espaço vetorial. É possível também verificar facilmente que o produto de operadores unitários é ainda um operador unitário.

A segunda classe de operadores é a dos chamados *operadores de projeção*, identificados pelas propriedades

$$P^\dagger = P \quad \text{e} \quad P^2 \equiv PP = P.$$

A segunda dessas propriedades é conhecida como propriedade de *idempotência*. É imediato verificar que, se P é um operador de projeção, então $Q \equiv \hat{1} - P$ é também um operador de projeção. Este segundo operador de projeção é *complementar* a P , no sentido que $P + Q = \hat{1}$ e $PQ = QP = \hat{0}$. Usando esses dois operadores, um vetor qualquer $|\varphi\rangle$ do espaço pode ser decomposto na soma de dois vetores *ortogonais* como $|\varphi\rangle = P|\varphi\rangle + Q|\varphi\rangle$.

Um exemplo muito simples de operador linear, mas de ocorrência freqüente, é um objeto do tipo

$$g_{\chi_1\chi_2} \equiv |\chi_1\rangle\langle\chi_2|$$

onde nenhum dos dois vetores $|\chi_1\rangle$ e $|\chi_2\rangle$ é o vetor nulo. A ação deste objeto sobre um vetor qualquer $|\varphi\rangle$ é definida como

$$g_{\chi_1\chi_2}|\varphi\rangle = |\chi_1\rangle\langle\chi_2|\varphi\rangle.$$

A linearidade de tal operador decorre diretamente da linearidade da ação dos bras sobre vetores. O seu adjunto pode ser facilmente obtido a partir de

$$\langle\psi|g_{\chi_1\chi_2}^\dagger|\varphi\rangle = \langle\varphi|g_{\chi_1\chi_2}|\psi\rangle^* = (\langle\varphi|\chi_1\rangle\langle\chi_2|\psi\rangle)^* = \langle\psi|\chi_2\rangle\langle\chi_1|\varphi\rangle,$$

que vale para quaisquer vetores $|\varphi\rangle$ e $|\psi\rangle$, e que portanto o identifica como sendo

$$g_{\chi_1\chi_2}^\dagger = |\chi_2\rangle\langle\chi_1|.$$

Isso mostra que operadores do tipo $g_{\chi\chi} = |\chi\rangle\langle\chi|$ são hermiteanos. Notando ainda que

$$g_{\chi\chi}^2 = |\chi\rangle\langle\chi|\chi\rangle\langle\chi| = \langle\chi|\chi\rangle g_{\chi\chi}$$

resulta que

$$P_\chi \equiv \frac{g_{\chi\chi}}{\langle\chi|\chi\rangle} = \frac{|\chi\rangle\langle\chi|}{\langle\chi|\chi\rangle}$$

é um operador de projeção. Se, em particular, $|\chi\rangle$ é tomado como sendo um dos vetores $|\phi_n\rangle$ de norma 1 pertencente a uma base ortonormal, então

$$P_{\phi_n}|\varphi\rangle = |\phi_n\rangle\langle\phi_n|\varphi\rangle \quad (2.11)$$

isto é, P_{ϕ_n} associa a cada vetor $|\varphi\rangle$ a sua componente ao longo do vetor $|\phi_n\rangle$ da base. É imediato verificar que a soma de vários operadores P_{ϕ_n} , isto é

$$P_s \equiv \sum_{n_i \in \{s\}} P_{\phi_{n_i}},$$

onde $\{s\}$ indica um subconjunto qualquer de vetores da base ortonormal que tenham sido selecionados, é ainda um operador de projeção, cujo efeito sobre um vetor qualquer $|\varphi\rangle$ é

$$P_s|\varphi\rangle = \sum_{n_i \in \{s\}} |\phi_{n_i}\rangle\langle\phi_{n_i}|\varphi\rangle$$

isto é, ele associa a qualquer vetor $|\varphi\rangle$ a sua componente pertencente ao *sub-espaço* vetorial gerado pelos vetores $|\phi_{n_i}\rangle$ selecionados. Em particular, a expansão de um vetor qualquer $|\varphi\rangle$ na base completa pode ser escrita como

$$|\varphi\rangle = \sum_n P_{\phi_n}|\varphi\rangle = \sum_n |\phi_n\rangle\langle\phi_n|\varphi\rangle$$

o que mostra que o operador unidade pode ser visto como um “operador de projeção sobre o espaço todo” que pode ser escrito sob a forma

$$\hat{1} = \sum_n P_{\phi_n} = \sum_n |\phi_n\rangle\langle\phi_n| \quad (2.12)$$

que é usualmente chamada *resolução da unidade* em termos da base ortonormal $\{|\phi_n\rangle\}$.

2.2.3 Espectro de operadores lineares e a interpretação da Mecânica Quântica.

Dado um operador linear g , se existir um vetor *não nulo* $|\varphi_\gamma\rangle$ tal que valha a relação

$$g|\varphi_\gamma\rangle = \gamma|\varphi_\gamma\rangle, \quad (2.13)$$

chamada usualmente *equação de autovalores*, onde γ é um número (em geral complexo), então γ é um *autovalor* de g , e $|\varphi_\gamma\rangle$ é um *autovetor* associado a esse autovalor. É importante notar

que o autovalor γ pode eventualmente ser zero (e nesse caso o lado direito da relação se anula), embora seja essencial que o autovetor *não* seja o vetor nulo. O conteúdo dessa relação pode ser expresso em palavras dizendo que o resultado da ação do operador sobre um autovetor do operador é um vetor proporcional ao próprio autovetor, e que o autovalor associado ao autovetor é a constante de proporcionalidade entre a imagem do autovetor e ele próprio. Em vista do caráter homogêneo da equação linear (2.13), é claro que ela caracteriza o autovetor apenas a menos de uma constante multiplicativa complexa (que deve no entanto ser diferente de zero). Parte dessa ambigüidade pode ser removida impondo sobre o autovetor uma condição de normalização, tipicamente $\langle \varphi_\gamma | \varphi_\gamma \rangle = 1$. Nesse caso a ambigüidade remanescente se reduz a um fator complexo de módulo 1 (fator de fase, $e^{i\theta}$), isto é, condição de normalização define o autovetor a menos de um fator de fase.

O *espectro* de um operador linear consiste no conjunto de todos os seus autovalores. Tomando como exemplo o caso de um operador de projeção $P = P^\dagger = P^2$, não é difícil ver que seu espectro consiste apenas de dois números, 1 e 0. Isso resulta diretamente da propriedade de idempotência desse operador, pois essa propriedade, juntamente com a equação de autovalores

$$P|\varphi_\gamma\rangle = \gamma|\varphi_\gamma\rangle$$

dá a condição $\gamma^2 = \gamma$, que tem como soluções apenas $\gamma = 1$ e $\gamma = 0$. Esse argumento vale evidentemente para *qualquer* operador de projeção e não serve para identificar os autovetores correspondentes a cada um dos autovalores num caso específico que esteja sendo considerado. Para identificar também os autovetores no caso específico do operador P_{ϕ_n} escrito em (2.11), por exemplo, basta notar que, usando o operador de projeção complementar $Q_{\phi_n} = \hat{1} - P_{\phi_n}$, qualquer vetor $|\varphi\rangle$ pode ser decomposto como

$$|\varphi\rangle = P_{\phi_n}|\varphi\rangle + Q_{\phi_n}|\varphi\rangle$$

Desse modo a equação de autovalores fica

$$P_{\phi_n}|\varphi_\gamma\rangle = \gamma(P_{\phi_n}|\varphi_\gamma\rangle + Q_{\phi_n}|\varphi_\gamma\rangle)$$

cujas soluções são $Q_{\phi_n}|\varphi_\gamma\rangle = 0$ com $\gamma = 1$ ou $P_{\phi_n}|\varphi_\gamma\rangle = 0$ com $\gamma = 0$. No primeiro desses casos $|\varphi_{\gamma=1}\rangle$ deve ser proporcional a $|\phi_n\rangle$, e no segundo $|\varphi_{\gamma=0}\rangle$ deve ser ortogonal a $|\phi_n\rangle$. Enquanto no primeiro caso o autovetor é “único” (a menos das ambigüidades de normalização e fase), no segundo caso existe uma infinidade de autovetores linearmente independentes associados ao autovalor 0, correspondente à infinidade de vetores da base ortonormal $|\phi_m\rangle$, $m \neq n$, ortogonais a $|\phi_n\rangle$. Um autovalor ao qual correspondem vários autovetores linearmente independentes é dito *degenerado*, com grau de degenerescência igual ao número N de autovetores linearmente independentes associados a ele. Como qualquer combinação linear desses autovetores independentes

ainda é um autovetor associado ao mesmo autovalor degenerado, é sempre possível escolher um conjunto de N autovetores ortogonais para um autovalor com degenerescência N (aliás isso pode ser feito de infinitas maneiras). No caso do operador de projeção P_{ϕ_n} , o autovalor 0 é *infinitamente degenerado*.

A obtenção do espectro de um operador linear de um espaço de dimensão infinita é um problema que pode ser bastante complicado e que de qualquer forma tem que ser tratado caso a caso. Como exemplos extremos da variedade de resultados que se pode obter, basta considerar os operadores representados pelas matrizes \mathbf{a} e \mathbf{a}^\dagger introduzidas para o tratamento do oscilador harmônico através das relações (1.20). O espectro do operador \mathbf{a} , de fato, é todo o plano complexo, ao passo que o operador \mathbf{a}^\dagger não possui nenhum autovetor, e portanto tão pouco possui qualquer autovalor (v. Problema 1.5). Existem, no entanto, propriedades genéricas relativas seja ao espectro como um todo, seja a alguma característica que ele possua, que podem ser demonstradas para classes particulares de operadores. Um primeiro exemplo disso é o fato, já visto, que o espectro de operadores de projeção consiste *sempre* apenas dos dois autovalores reais 1 e 0. Os autovetores correspondentes a cada um desses autovalores (e portanto o seu respectivo grau de degenerescência) depende no entanto do particular operador que esteja sendo considerado, e tem que ser obtidos também caso a caso. Uma segunda propriedade genérica que pode ser demonstrada é a de que os autovalores de operadores unitários são números complexos de módulo 1 (fases). De fato, se u é um operador unitário e se $|\varphi_v\rangle$ é um seu autovetor com autovalor v , isto é,

$$u|\varphi_v\rangle = v|\varphi_v\rangle$$

segue que

$$\langle\varphi_v|u^\dagger = \langle\varphi_v|v^* ;$$

aplicando os bras da segunda equação sobre os kets da primeira

$$\langle\varphi_v|u^\dagger u|\varphi_v\rangle = |v|^2 \langle\varphi_v|\varphi_v\rangle$$

de onde resulta que $|v|^2 = 1$ em vista de que $u^\dagger u = \hat{1}$ devido à unitariedade de U . Uma classe importante de operadores unitários é a dos operadores do tipo e^{ig} onde g é um operador hermiteano, $g = g^\dagger$. A unitariedade de tais operadores resulta de que

$$(e^{ig})^\dagger = e^{-ig^\dagger} = e^{-ig} = (e^{ig})^{-1}.$$

Se então $|\varphi_\gamma\rangle$ é um autovetor de g com autovalor γ

$$e^{ig}|\varphi_\gamma\rangle = e^{i\gamma}|\varphi_\gamma\rangle$$

isto é, esse mesmo autovetor é também autovetor do operador unitário e^{ig} com autovalor $e^{i\gamma}$. Como este último autovalor deve ser uma fase, os autovalores de operadores hermiteanos são *reais*. De uma forma inteiramente análoga é possível verificar que a exponencial e^g de um operador anti-hermiteano $g = -g^\dagger$ é também um operador unitário, e que os autovalores de um operador anti-hermiteano são números imaginários puros. A natureza dos autovalores de operadores hermiteanos e anti-hermiteanos é portanto mais uma instância da analogia mencionada anteriormente com quantidades reais e imaginárias.

Existe ainda uma propriedade genérica importante que se aplica aos *autovetores* de operadores hermiteanos, anti-hermiteanos e unitários, que é a de que autovetores que correspondem a autovalores diferentes são ortogonais. No caso de um operador hermiteano $g = g^\dagger$, supondo que se tenha

$$g|\varphi_\gamma\rangle = \gamma|\varphi_\gamma\rangle \quad \text{e} \quad g|\varphi_\delta\rangle = \delta|\varphi_\delta\rangle,$$

tomando o produto escalar da primeira equação com o autovetor $|\varphi_\delta\rangle$ resulta

$$\langle\varphi_\delta|g|\varphi_\gamma\rangle = \gamma\langle\varphi_\delta|\varphi_\gamma\rangle;$$

usando a hermiticidade de g , o lado esquerdo dessa equação pode ser reescrito como

$$\langle\varphi_\delta|g|\varphi_\gamma\rangle = \langle\varphi_\gamma|g^\dagger|\varphi_\delta\rangle^* = \langle\varphi_\gamma|g|\varphi_\delta\rangle^* = \delta\langle\varphi_\delta|\varphi_\gamma\rangle$$

onde foi usado também o fato de que o autovalor δ é real. Desse modo resulta que

$$\delta\langle\varphi_\delta|\varphi_\gamma\rangle = \gamma\langle\varphi_\delta|\varphi_\gamma\rangle \quad \text{ou} \quad (\delta - \gamma)\langle\varphi_\delta|\varphi_\gamma\rangle = 0;$$

portanto, se $\delta \neq \gamma$ segue que $\langle\varphi_\delta|\varphi_\gamma\rangle = 0$. Procedimentos análogos estendem esse resultado também a operadores anti-hermiteanos e unitários, tendo em conta a natureza dos autovalores (números imaginários puros e fases, respectivamente) em cada caso. Para operadores unitários é preciso usar o fato de que, se v e $|\varphi_v\rangle$ resolvem a equação de autovalores

$$u|\varphi_v\rangle = v|\varphi_v\rangle$$

para o operador unitário u , então

$$\langle\varphi_v|u = v\langle\varphi_v|,$$

que resulta imediatamente da aplicação de u^\dagger sobre a equação de autovalores. De fato, isso dá, em vista da unitariedade de u ,

$$|\varphi_v\rangle = vu^\dagger|\varphi_v\rangle,$$

que mostra que a ação de u^\dagger sobre o autovetor $|\varphi_v\rangle$ de u equivale a multiplicação pelo complexo conjugado da fase v .

Existe uma forma de definir o espectro de um operador linear “sem falar em autovetores”. Para chegar ao espectro do operador g , ela se baseia na consideração do operador $g - z\hat{1}$, onde z é um número complexo, e o espectro é definido em termos dos valores de z nos quais a inversão desse operador é problemática. É claro que quando z é igual a um autovalor γ correspondente a um autovetor $|\varphi_\gamma\rangle$ esse operador não tem um inverso, pois $|\varphi_\gamma\rangle$ é uma solução diferente do vetor nulo de $(g - \gamma\hat{1})|\varphi_\gamma\rangle = 0$. Este tipo de problema na inversão de $g - z\hat{1}$ de fato esgota o repertório de problemas no caso de espaços de dimensão finita, de modo que a equivalência com o ponto de vista usual, “falando de autovetores”, é clara. No caso de espaços de dimensão infinita, pode ocorrer outro tipo de problema, em que o inverso não existe como operador limitado, isto é tal que sua ação sobre qualquer vetor de norma 1 tenha norma finita. Este tipo de problema está associado à existência de um espectro contínuo de autovalores, para quais a rigor não existem autovetores, no sentido de vetores de norma finita. A forma mais simples (e por isso bastante usual, mas menos precisa) de tratar esse tipo de espectro consiste em introduzir “autovetores impróprios”, de norma infinita.

A importância dessas várias definições e propriedades para a mecânica quântica está ligada diretamente ao fato de que as variáveis dinâmicas da teoria são implementadas como operadores hermiteanos particulares, num espaço vetorial, em geral de dimensão infinita. Em particular, no âmbito específico da mecânica das matrizes de Born e Jordan os valores possíveis da energia dos estados estacionários do sistema quântico são identificados com o espectro do operador hermiteano associado à energia total do sistema (o operador hamiltoniano para o sistema quântico, H). Essa identificação pode também ser pensada em termos de um *postulado interpretativo*, eventualmente passível de generalização, segundo o qual *o espectro de H deve ser interpretado como o conjunto de valores possíveis que a energia pode assumir em estados estacionários*. Essa interpretação não faz menção aos autovetores $|\phi_n\rangle$ associados aos autovalores E_n de H , mas em vista do fato que os elementos de matriz $q_{nm} = \langle\phi_n|q|\phi_m\rangle$ do operador q associado à posição correspondem a valores de um “momento de dipolo” associado a transições entre os estados estacionários associados aos autovalores E_n e E_m , os quais aparecem aí representados apenas por seus respectivos autovetores. Isso então sugere a interpretação dos autovetores de H como *objetos quânticos que caracterizam os estados estacionários do sistema*. Um óbvio passo seguinte é interpretar vetores mais gerais obtidos utilizando a estrutura de espaço vetorial, ou seja,

vetores escritos como combinações lineares de vetores que representam estados estacionários, do tipo

$$|\varphi\rangle = \sum_n c_n |\phi_n\rangle,$$

também como caracterizadores quânticos de estados (em geral não estacionários) do sistema quântico. A questão que se coloca aqui é a de que energia atribuir a esses estados mais gerais. A resposta a essa questão *não* está contida na estrutura formal montada para a teoria, mas depende de um postulado interpretativo adicional, formulado pela primeira vez por Born, em 1926. Para motiva-lo, vale observar inicialmente que a associação de $|\phi_n\rangle$ ao autovalor correspondente E_n de H pode ser feita através da relação

$$E_n = \langle \phi_n | H | \phi_n \rangle$$

com a normalização $\langle \phi_n | \phi_n \rangle = 1$. Escrevendo a expressão correspondente para o estado mais geral $|\varphi\rangle$, também normalizado de forma que $\langle \varphi | \varphi \rangle = 1$ (que corresponde a $\sum_n |c_n|^2 = 1$ quando todos os autovetores são normalizados da mesma forma), resulta

$$\langle E \rangle_\varphi \equiv \langle \varphi | H | \varphi \rangle = \sum_{n,m} c_m^* c_n \langle \phi_m | H | \phi_n \rangle = \sum_n |c_n|^2 E_n \quad (2.14)$$

onde foi usada também a ortogonalidade dos diferentes autovetores $|\phi_n\rangle$. A notação $\langle E \rangle_\varphi$ foi introduzida para sugerir a ideia de uma *energia média* associada ao vetor $|\varphi\rangle$, que aparece finalmente como uma média ponderada das energias dos estados estacionários incluídos na combinação linear que caracteriza esse vetor, com pesos dados pelo módulo quadrado dos coeficientes c_n . Segundo a interpretação proposta por Born, **a)** o resultado de uma medição da energia do sistema quântico é *sempre* um dos autovalores de H e **b)** a *probabilidade* de obter um particular resultado E_n é $|c_n|^2$ quando o *estado* do sistema é descrito pelo vetor $|\varphi\rangle$. Dessa forma o valor médio calculado corresponde de fato à média dos resultados de um número muito grande de medições da energia do sistema, quando ele se encontra nesse estado. É claro, em particular, que apenas quando uma só das probabilidades $|c_n|^2$ é diferente de zero (e ela deve então ser igual a 1 pela condição de normalização) o resultado da medição da energia pode ser previsto “com probabilidade 1”. Nesse caso o sistema se encontrará no correspondente estado estacionário.

O cálculo dessas probabilidades quânticas pode ser manejado de forma muito conveniente através do uso de operadores de projeção. De fato, é fácil ver que

$$|c_n|^2 = \langle \varphi | \phi_n \rangle \langle \phi_n | \varphi \rangle = \langle \varphi | P_{\phi_n} | \varphi \rangle \quad \text{com} \quad P_{\phi_n} = |\phi_n\rangle \langle \phi_n|$$

com todos os estados devidamente normalizados, sendo então P_{ϕ_n} um operador de projeção. Essa quantidade corresponde à probabilidade de obter o autovalor E_n numa medição da energia, *supondo que esse autovalor não seja degenerado*. No caso em que ele tiver degenerescência N haverá N autovetores $|\phi_{n,k}\rangle$, $k = 1, \dots, N$ linearmente independentes (que podem ser tomados como sendo ortogonais e devidamente normalizados) associados a E_n , e nesse caso a probabilidade de obtê-lo será dada por

$$\sum_{k=1}^N |c_{n,k}|^2 = \langle \varphi | P_n | \varphi \rangle$$

onde agora

$$P_n = \sum_{k=1}^N |\phi_{n,k}\rangle \langle \phi_{n,k}|, \quad (2.15)$$

ainda um operador de projeção. Embora o conjunto de N vetores ortonormais $|\phi_{n,k}\rangle$, $k = 1, \dots, N$ possa ser escolhido de infinitas maneiras diferentes, isso não afeta a unicidade *do operador de projeção* P_n , e portanto tão pouco a do quadrado da norma da projeção $P_n|\varphi\rangle$, que corresponde à probabilidade associada ao autovalor E_n nesse caso.

A chamada *interpretação probabilística* da teoria quântica, devida a Max Born, consiste numa generalização mais ou menos direta dessas idéias a *todas* as variáveis dinâmicas. Nela, o espaço de Hilbert que serve de substrato à teoria desempenha um papel análogo ao espaço de fases da teoria clássica, no sentido que diferentes vetores desse espaço correspondem a diferentes *estados* do sistema quântico. Eles são por isso freqüentemente chamados *vetores de estado*. Os valores que é possível obter numa medição de *qualquer* variável dinâmica g são aqueles que pertencem ao espectro do operador associado à variável dinâmica considerada. O caráter hermiteano desse operador garante a realidade dos autovalores e portanto do resultado da medição. A probabilidade de obter um determinado autovalor γ de g numa medição, quando o sistema se encontra em um estado descrito por um vetor de estado $|\psi\rangle$ dado, pode ser calculada como $\langle \psi | P_\gamma | \psi \rangle$, onde P_γ é o operador de projeção sobre o autovetor (ou autovetores independentes, no caso de um autovalor degenerado) associado(s) a γ . É importante ter sempre em vista o fato de que, no contexto da interpretação de Born, os vetores de estado são entendidos como devidamente normalizados, e que a modificação de um vetor apenas através de um fator de fase não afeta as quantidades fisicamente relevantes, que se exprimem em termos de módulos quadrados. Assim, “vetores diferentes”, no sentido de vetores de estado diferentes, significa a rigor *vetores normalizados que diferem mais que por um fator de fase apenas*. Isso em particular torna a indeterminação dos autovetores, devida ao caráter linear e homogêneo das equações de autovalores, irrelevante para fins da interpretação.

Uma outra forma de exprimir passagem à interpretação probabilística geral, tendo em vista a associação de vetores a estados do sistema, é interpretar a quantidade $|c_n|^2 = |\langle \phi_n | \varphi \rangle|^2$ na equação (2.14) como a *probabilidade de que o sistema esteja no estado $|\phi_n\rangle$* (no qual a energia tem o valor bem definido E_n), *dado que ele se encontra no estado $|\varphi\rangle$* (ou vice versa). Essa forma da interpretação pode então ser estendida a dois estados normalizados quaisquer $|\varphi\rangle$ e $|\psi\rangle$, um dos quais pode em particular ser um dos autovetores de alguma variável dinâmica, interpretado por sua vez como estado em que essa variável tem um valor bem definido. Essa formulação expõe claramente a consequência peculiar de que, além do fato de as variáveis dinâmicas não terem em geral valores bem definidos em estados quânticos completamente especificados pelos respectivos vetores de estado, as propriedades “estar no estado $|\varphi\rangle$ ” e “estar no estado $|\psi\rangle$ ” *não se excluem mutuamente* para estados diferentes, a menos do caso particular em que eles sejam *ortogonais*.

Para que essa interpretação probabilística da mecânica quântica seja uma interpretação completa, no sentido de poder ser aplicada a vetores de estado e variáveis dinâmicas quaisquer, é claro que é preciso que qualquer vetor de estado possa ser analisado em termos dos autovetores do operador correspondente a qualquer variável dinâmica. Essa condição não apresenta qualquer problema no caso especial de espaços vetoriais de dimensão *finita* D , pois nesse caso se pode mostrar que o espectro de qualquer operador hermiteano consiste sempre de um conjunto discreto de D números reais (não necessariamente diferentes devido à possibilidade de ocorrência de autovalores degenerados) aos quais se pode sempre associar uma base ortonormal de D autovetores. No caso de haver autovalores degenerados essa base apenas não é única (além da ambigüidade irrelevante de fases), sem que no entanto isso afete o cálculo de probabilidades de obtenção de cada um dos autovalores, as quais dependem na realidade de operadores de projeção como (2.15). No caso de espaços de dimensão infinita, a hermiticidade do operador (ou mais precisamente, seu caráter auto-adjunto) não é suficiente para garantir a existência de uma base de autovetores. Isso levou Dirac a distinguir, como uma classe especial de operadores hermiteanos de espaços de dimensão infinita, aqueles que permitem a análise de vetores de estado quaisquer em moldes essencialmente iguais aos que se aplicam a operadores hermiteanos em espaços de dimensão finita, dando-lhes o rótulo especial de *observáveis*.¹ Essa distinção por si só não resolve porém completamente o problema pois, no caso de espaços de dimensão infinita, o espectro dos observáveis pode ser parcial ou até totalmente *contínuo* (o que corresponde a variáveis dinâmicas para as quais os possíveis resultados de medições não se restringem a um conjunto discreto de valores, mas pelo contrário cobrem intervalos reais contínuos), e nesse caso há problemas especiais de normalização que afetam os ‘autovetores’ correspondentes.

¹A referência explícita a *observáveis* com este sentido aparece pela primeira vez na segunda edição dos *The Principles of Quantum Mechanics*, datada de 1934, e é desde então amplamente adotada.

Para identificar esses problemas de normalização, e também para mostrar um modo efetivo de contorná-los (também devido a Dirac), é conveniente considerar inicialmente o caso possível de um observável $g = g^\dagger$ com espectro puramente contínuo $\gamma_{\min} < \gamma < \gamma_{\max}$. Chamando $|\phi_\gamma\rangle$ os ‘autovetores’ correspondentes aos autovalores γ , da mesma forma que no caso do espectro discreto pode-se ver que $\langle\phi_{\gamma'}|\phi_\gamma\rangle = 0$ sempre que $\gamma' \neq \gamma$. A completeza dos ‘autovetores’ apropriadamente normalizados do espectro contínuo exprime-se formalmente através de uma resolução da unidade correspondente à (2.12) que pode ser escrita como

$$\hat{1} = \int_{\gamma_{\min}}^{\gamma_{\max}} |\phi_\gamma\rangle d\gamma \langle\phi_\gamma|, \quad (2.16)$$

sendo o sentido dessa integral explicitado como significando que o produto escalar de dois vetores $|\varphi\rangle$ e $|\psi\rangle$ quaisquer do espaço vetorial pode ser escrito como

$$\langle\psi|\varphi\rangle = \int_{\gamma_{\min}}^{\gamma_{\max}} \langle\psi|\phi_\gamma\rangle \langle\phi_\gamma|\varphi\rangle d\gamma$$

o que mostra que as funções da variável contínua γ que aparecem no integrando, $c(\gamma) \equiv \langle\phi_\gamma|\varphi\rangle$ e $d^*(\gamma) \equiv \langle\psi|\phi_\gamma\rangle$ efetivamente representam os vetores $|\varphi\rangle$ e $|\psi\rangle$ em termos dos ‘autovetores’ do espectro contínuo. Em particular, a norma de $|\varphi\rangle$ é dada por

$$\|\varphi\|^2 = \int_{\gamma_{\min}}^{\gamma_{\max}} |c(\gamma)|^2 d\gamma,$$

o que mostra que o caráter finito da norma implica que as funções representativas sejam funções de quadrado integrável definidas sobre o espectro contínuo. Finalmente, é preciso explicitar a condição de normalização envolvida na validade de (2.16). Para identifica-la, basta utilizar essa mesma resolução da unidade em

$$\langle\phi_\gamma|\varphi\rangle = \langle\phi_\gamma|\hat{1}|\varphi\rangle = \int_{\gamma_{\min}}^{\gamma_{\max}} \langle\phi_\gamma|\phi_{\gamma'}\rangle \langle\phi_{\gamma'}|\varphi\rangle d\gamma'$$

onde a variável de integração aparece escrita como γ' . Dessa relação é claro que a normalização dos autovetores do espectro contínuo deve ser tomada como sendo

$$\langle\phi_\gamma|\phi_{\gamma'}\rangle \rightarrow \delta(\gamma - \gamma') \quad (2.17)$$

onde a última função é a função delta de Dirac. Os ‘autovetores’ $|\phi_\gamma\rangle$ não têm, portanto, norma finita, e portanto não são, a rigor, vetores do espaço de Hilbert, embora possam, não obstante seu caráter “impróprio”, ser utilizados essencialmente em pé de igualdade com os objetos de

melhor linhagem num grande número de situações, algumas das quais são exemplificadas a seguir²

É fácil verificar que, usando a normalização (2.17), o operador

$$P_{\Delta} \equiv \int_{\gamma \in \Delta} |\phi_{\gamma}\rangle d\gamma \langle \phi_{\gamma}|,$$

onde a integração é limitada a um sub-intervalo (ou conjunto de sub-intervalos) Δ do espectro de g , é um operador de projeção. O sentido dessa integral é inteiramente análogo ao da que aparece em (2.16). Supondo que $|\varphi\rangle$ seja um vetor de estado normalizado, isto é, $\|\varphi\| = 1$, a quantidade

$$\langle \varphi | P_{\Delta} | \varphi \rangle = \int_{\gamma \in \Delta} \langle \varphi | \phi_{\gamma} \rangle \langle \phi_{\gamma} | \varphi \rangle d\gamma = \int_{\gamma \in \Delta} |c(\gamma)|^2 d\gamma$$

deve ser interpretada como a probabilidade de obter valores de γ pertencentes a Δ numa medição de g , dado que o sistema se encontre no estado descrito por $|\varphi\rangle$. Isso mostra que, no caso do espectro contínuo, $|\langle \phi_{\gamma} | \varphi \rangle|^2 = |c(\gamma)|^2$ tem o significado de uma *densidade de probabilidade*

²O ponto de vista de Dirac foi desde o início criticado por von Neumann, a quem se deve uma apresentação alternativa mas igualmente clássica da “teoria das transformações”, dentro de padrões matematicamente “menos objetáveis” (**John von Neumann**, *Mathematical Foundations of Quantum Mechanics*, Trad. Robert T. Beyer, Princeton University Press 1955). No prefácio a esse livro von Neumann diz

Dirac, in several papers, as well as in his recently published book, has given a representation of quantum mechanics which is scarcely to be surpassed in brevity and elegance, and which is at the same time of invariant character. It is therefore perhaps fitting to advance a few arguments on behalf of our method, which deviates considerably from that of Dirac.

The method of Dirac...in no way satisfies the requirements of mathematical rigor – not even if these are reduced in a natural and proper fashion to the extent common elsewhere in theoretical physics. ... There would be no objection here if these [“improper”] concepts...were intrinsically necessary for the physical theory. ... But this is by no means the case. ... It should be emphasized that the correct structure need not consist in a mathematical refinement and explanation of the Dirac method, but rather that it requires a procedure differing from the very beginning, namely, the reliance on the Hilbert theory of operators.

A edição original, em alemão, do livro de von Neumann foi publicada em 1932. O prefácio do autor (não datado) que aparece na tradução de 1955, de onde foram extraídos os trechos acima, faz referência, no entanto, à edição de 1936 do *The Principles...* de Dirac, e deve ter sido portanto escrito posteriormente a essa primeira edição de 1932.

Um testemunho interessante das idas e vindas conceituais a que está sujeito, inevitavelmente, o processo, sempre aberto, de elaboração de uma teoria é encontrado em uma carta de von Neumann ao matemático Garret Birkhoff, datada de uma quarta-feira, 13 de Novembro (presumivelmente de 1935), citada pelo seu destinatário em artigo publicado nos *Proc. Symp. Pure Math.*, American Mathematical Society (1961), vol. II, pag. 155. Nela von Neumann escreve, com relação à mecânica quântica:

definida sobre o espectro contínuo. Ao contrário do que acontece em geral no caso de espectros discretos, a probabilidade de obter um valor γ *dado* é zero.

Em geral, o espectro de um observável g num espaço de dimensão infinita é em parte discreto e em parte contínuo. A completeza dos seus autovetores, próprios e impróprios, que o caracteriza como observável, pode ser expressa em termos de uma resolução da unidade do tipo

$$\hat{1} = \sum_n |\phi_n\rangle\langle\phi_n| + \int |\phi_\gamma\rangle d\gamma \langle\phi_\gamma|$$

com as condições $\langle\phi_m|\phi_n\rangle = \delta_{mn}$ e $\langle\phi_\gamma|\phi_{\gamma'}\rangle = \delta(\gamma - \gamma')$ respectivamente para os autovetores próprios do espectro discreto e para os impróprios do espectro contínuo. Na parte referente ao espectro contínuo o domínio de integração é a parte contínua do espectro do observável considerado. Usando essa resolução da unidade, um vetor de estado qualquer $|\varphi\rangle$ pode ser expresso de forma simples e útil em termos de uma 'função', em parte discreta e em parte contínua, definida sobre o espectro do observável g considerado. A correspondência entre estados e 'funções' é dada por

$$|\varphi\rangle \rightarrow \begin{cases} \langle\phi_n|\varphi\rangle = c_n & (\text{espectro discreto}) \\ \langle\phi_\gamma|\varphi\rangle = c(\gamma) & (\text{espectro contínuo}) . \end{cases}$$

I would like to make a confession which may seem immoral: I do not believe absolutely in Hilbert space any more. After all Hilbert space (as far as quantum-mechanical things are concerned) was obtained by generalizing Euclidean space, footing on the principle of 'conserving the validity of all formal rules'....Now we begin to believe, that it is not the *vectors* which matter, but the lattice of all linear (closed) subspaces. Because (1) The vectors ought to represent the physical *states*, but they do it redundantly, up to a complex factor, only, (2) and besides, the states are merely a derived notion, the primitive (phenomenologically given) notion being the *qualities*, which correspond to the *linear closed subspaces*. But if we wish to generalize the lattice of all linear closed subspaces from a Euclidean space to infinitely many dimensions, then one does not obtain Hilbert space, but that configuration, which Murray and I called 'case II₁'. (The lattice of all linear closed subspaces of Hilbert space is our 'case I_∞'.)

Birkhoff chega a sugerir em 1961, a propósito dessa citação, alguma aproximação de seu conteúdo com o ponto de vista de Dirac, dizendo que

... many physicists do not make any explicit technical use of the infinite-dimensional character of Hilbert space. For example, though mathematicians might like to interpret the "bras" and "kets" of Dirac as vectors in Hilbert space and its dual, and his "observables" as self-adjoint operators, there is no evidence that Dirac himself thinks them in this way. Thus, he makes no effort to fit his delta-function consistently into a Hilbert space framework - nor is it obvious how he could.

Em termos dessas funções, o resultado da ação do operador g sobre o estado $|\varphi\rangle$ corresponde a uma nova função que pode ser facilmente identificada calculando $\langle\phi_n|g|\varphi\rangle$ e $\langle\phi_\gamma|g|\varphi\rangle$. Tendo em conta a hermiticidade de g , isso leva ao resultado

$$g|\varphi\rangle \rightarrow \begin{cases} \langle\phi_n|g|\varphi\rangle = \gamma_n\langle\phi_n|\varphi\rangle = \gamma_n c_n & (\text{espectro discreto}) \\ \langle\phi_\gamma|g|\varphi\rangle = \gamma\langle\phi_\gamma|\varphi\rangle = \gamma c(\gamma) & (\text{espectro contínuo}) \end{cases}$$

que é usualmente expresso dizendo que o observável que deu origem à resolução da unidade age *como um operador de multiplicação* sobre as funções que representam os vetores de estado através dessa mesma resolução da unidade.

2.2.4 Compatibilidade de observáveis e relações de incerteza.

No contexto da interpretação probabilística exposta na seção 2.2.3, um observável só tem um valor bem definido para um dado estado de um sistema quando o vetor que caracteriza esse estado é um autovetor do observável em questão. Isso significa que, para que *dois* observáveis possam ter simultaneamente valores bem definidos, é preciso que o estado do sistema seja *simultaneamente* um autovetor desses dois variáveis observáveis. É claro que a existência de estados *particulares* com essa propriedade relativamente a dois (ou mais) observáveis dados depende dos observáveis que estejam sendo considerados e deve ser estudada caso a caso. No entanto, é possível obter condições simples sob as quais dois (ou mais) observáveis admitam um conjunto *completo* de autovetores *comuns*. Observáveis com essa propriedade são chamados, por razões óbvias, *observáveis compatíveis*.

Sejam, de fato dois observáveis g e h (que, por simplicidade, serão supostos como tendo espectro puramente discreto) dotados de um conjunto completo de autovetores comuns $|\phi_n\rangle$, isto é

$$g|\phi_n\rangle = \gamma_n|\phi_n\rangle \quad \text{e} \quad h|\phi_n\rangle = \eta_n|\phi_n\rangle.$$

Nesse caso

$$gh|\phi_n\rangle = g\eta_n|\phi_n\rangle = \gamma_n\eta_n|\phi_n\rangle = h\gamma_n|\phi_n\rangle = hg|\phi_n\rangle$$

de modo que $(gh - hg)|\phi_n\rangle = 0$ para todos os autovetores $|\phi_n\rangle$. Como esses autovetores formam um conjunto completo, segue que g e h comutam, isto é, $[g, h] = 0$.

Reciprocamente, se g e h comutam, e usando o conjunto completo de autovetores de h com autovalores η_n ,

$$hg|\phi_n\rangle = gh|\phi_n\rangle = \eta_ng|\phi_n\rangle,$$

isto é, $g|\phi_n\rangle$ é também um autovetor de h com o mesmo autovalor η_n . Para analisar completamente as implicações disso é preciso considerar três casos:

i) $g|\phi_n\rangle$ é o vetor nulo $|\emptyset\rangle$. Dado que $|\phi_n\rangle$ não é nulo, pois é autovetor de h , segue que esse vetor é também autovetor de g com autovalor zero.

ii) $g|\phi_n\rangle$ não é nulo, e o autovalor η_n de h não é degenerado. Isso significa que o autovetor $|\phi_n\rangle$ associado a esse autovalor é único, e portanto $g|\phi_n\rangle$ deve ser proporcional a esse autovetor, sendo então também um autovetor de g .

iii) $g|\phi_n\rangle$ não é nulo, e o autovalor η_n de h tem degenerescência μ_n . Nesse caso existem μ_n autovetores ortogonais $|\phi_{nk}\rangle$, $k = 1, \dots, \mu_n$ com esse mesmo autovalor, e os vetores $g|\phi_{nj}\rangle$ podem ser escritos, para cada um dos valores de j , como combinações lineares dos $|\phi_{nk}\rangle$, isto é

$$g|\phi_{nj}\rangle = \sum_{k=1}^{\mu_n} c_{jk}|\phi_{nk}\rangle.$$

Os coeficientes c_{jk} definem então uma matriz hermiteana de dimensão μ_n de cuja diagonalização resultam μ_n autovetores simultâneos de h e g . Os autovalores correspondentes de h são evidentemente todos iguais a η_n , enquanto os autovalores de g são os autovalores da matriz $\{c_{jk}\}$. Isso mostra portanto que dois observáveis g e h têm um conjunto completo de autovetores comuns (isto é, são compatíveis) se e somente se eles comutam, $[g, h] = 0$.

Com relação ao caso iii), em que o autovalor η_n de h é degenerado, deve ser notado que os autovalores γ_{nk} da matriz $\{c_{jk}\}$, que correspondem aos autovalores de g para os autovetores comuns dos dois operadores poderão ter graus de degenerescência menores que μ_n , ou poderão até mesmo ser todos distintos. Neste último caso, a especificação dos autovalores η_n de h e γ_{nk} de g identificam completamente um *único* vetor de estado, pois a especificação do autovalor de g remove completamente a ambiguidade devida à degenerescência de η_n .

Esse resultado pode evidentemente ser estendido para mais de dois operadores. Uma família de observáveis g, h, l, \dots é dita compatível se os observáveis que a constituem comutam entre si quando tomados dois a dois em todas as possíveis combinações. Nesse caso existe um conjunto completo de autovetores comuns a todos os observáveis da família. Se a especificação dos autovalores correspondente a cada um dos observáveis identifica completamente um único vetor de estado (isto é, a especificação dos autovalores dos diferentes observáveis elimina completamente as eventuais ambiguidades associadas a degenerescências presentes em cada um dos espectros), então a família é designada como *conjunto completo de observáveis compatíveis* ou ainda como *observação maximal*. Um único observável com espectro não degenerado constitui por si só uma observação maximal: o vetor de estado do sistema está univocamente determinado se o observável tem um valor bem definido.

Quando um dado observável g não tem um valor bem definido para um determinado vetor de estado $|\varphi\rangle$, isto é, esse vetor se exprime como uma combinação linear de vários autovetores

$|\phi_n\rangle$ de g associados a autovalores γ_n diferentes, é conveniente dispor de uma quantidade que expresse quantitativamente a indefinição do valor de g em $|\varphi\rangle$. De acordo com a interpretação estatística, a indefinição provém da associação de probabilidades não nulas $p_n \equiv |\langle\phi_n|\varphi\rangle|^2$ a autovalores γ_n diferentes. Desse modo, além do *valor médio* do observável que, supondo o vetor de estado normalizado, $\|\varphi\| = 1$, se exprime como

$$\langle g \rangle_\varphi \equiv \sum_n p_n \gamma_n = \langle \varphi | g | \varphi \rangle,$$

é possível calcular também o desvio quadrático médio

$$\sigma_g^{(\varphi)^2} \equiv \sum_n p_n \gamma_n^2 - \left(\sum_n p_n \gamma_n \right)^2 = \langle \varphi | g^2 | \varphi \rangle - \langle \varphi | g | \varphi \rangle^2 \quad (2.18)$$

como medida da indefinição de g no estado $|\varphi\rangle$. Essa é uma quantidade não negativa, como pode ser verificado a partir da desigualdade de Schwartz, que permite escrever, usando a hermiticidade de g ,

$$\langle \varphi | g | \varphi \rangle^2 = |\langle \varphi | g | \varphi \rangle|^2 \leq \langle \varphi | \varphi \rangle \langle \varphi | g^2 | \varphi \rangle.$$

O desvio quadrático médio se anula quando essa desigualdade vale como *igualdade*. Para ver o que isso significa, introduzindo os operadores de projeção $P_\varphi \equiv |\varphi\rangle\langle\varphi|$ e $Q_\varphi \equiv \hat{1} - P_\varphi$ é possível escrever em geral

$$\langle \varphi | g^2 | \varphi \rangle = \langle \varphi | g P_\varphi g | \varphi \rangle + \langle \varphi | g Q_\varphi g | \varphi \rangle$$

o que mostra que para que a desigualdade valha como igualdade deve-se ter

$$\langle \varphi | g Q_\varphi g | \varphi \rangle = \langle \varphi | g Q_\varphi^2 g | \varphi \rangle = 0$$

ou seja, $Q_\varphi g | \varphi \rangle$, que é a componente de $g | \varphi \rangle$ ortogonal a $|\varphi\rangle$, tem norma nula, e portanto $g | \varphi \rangle = |\varphi\rangle\langle\varphi| g | \varphi \rangle$, isto é, $|\varphi\rangle$ é autovetor de g .

Uma forma útil de exprimir o desvio quadrático médio $\sigma_g^{(\varphi)}$ consiste em introduzir o operador 'deslocado'

$$g_{(\varphi)} \equiv g - \langle \varphi | g | \varphi \rangle \hat{1}$$

pois isso permite escreve-lo simplesmente como

$$\sigma_g^{(\varphi)^2} = \langle \varphi | g_{(\varphi)}^2 | \varphi \rangle.$$

Considerando agora os desvios quadráticos médios no estado $|\varphi\rangle$, $\sigma_g^{2(\varphi)}$ e $\sigma_h^{2(\varphi)}$, de *dois* observáveis g e h , não necessariamente compatíveis, resulta, usando mais uma vez a desigualdade de Schwartz, que o produto $\sigma_g^{2(\varphi)}\sigma_h^{2(\varphi)}$ satisfaz à desigualdade escrita em termos dos operadores deslocados

$$\sigma_g^{(\varphi)2}\sigma_h^{(\varphi)2} = \langle\varphi|g_{(\varphi)}^2|\varphi\rangle\langle\varphi|h_{(\varphi)}^2|\varphi\rangle \geq |\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle|^2.$$

O valor esperado do produto $g_{(\varphi)}h_{(\varphi)}$ é em geral complexo se esses operadores não comutam, donde segue que

$$|\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle|^2 = \left(\operatorname{Re}\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle\right)^2 + \left(\operatorname{Im}\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle\right)^2 \geq \left(\operatorname{Im}\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle\right)^2. \quad (2.19)$$

Usando a hermiticidade de g e h , esta última quantidade pode ser expressa em termos do comutador desses dois operadores. De fato

$$\begin{aligned} \operatorname{Im}\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle &= \frac{1}{2i} \left(\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle - \langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle^* \right) \\ &= \frac{1}{2i} \left(\langle\varphi|g_{(\varphi)}h_{(\varphi)}|\varphi\rangle - \langle\varphi|h_{(\varphi)}g_{(\varphi)}|\varphi\rangle \right) \\ &= \frac{\langle\varphi|[g_{(\varphi)}, h_{(\varphi)}]|\varphi\rangle}{2i} = \frac{\langle\varphi|[g, h]|\varphi\rangle}{2i} \end{aligned}$$

donde se obtém a desigualdade

$$\sigma_g^{(\varphi)}\sigma_h^{(\varphi)} \geq \left| \frac{\langle\varphi|[g, h]|\varphi\rangle}{2i} \right|. \quad (2.20)$$

Essa relação dá um *limite inferior* para o produto dos desvios quadráticos médios associados a dois observáveis num estado quântico $|\varphi\rangle$ em termos do valor médio nesse estado do comutador dos operadores associados a esses observáveis, e pode portanto ser vista como uma relação geral de incerteza. É importante notar que esse limite inferior *em geral depende do estado considerado*. Ele pode, em particular, ser zero mesmo no caso em que os observáveis não sejam compatíveis, como ocorre, por exemplo, no caso em que $|\varphi\rangle$ é autovetor de um dos observáveis. Existe porém uma situação em que essa relação geral de incerteza implica numa limitação *independente do estado considerado* para o produto dos desvios quadráticos médios. Ela corresponde ao caso em que o comutador $[g, h]$ é um múltiplo da unidade. Isso ocorre notadamente no caso dos operadores associados à coordenada e ao seu momento conjugado, que devem satisfazer à relação de comutação de Heisenberg, Born e Jordan

$$[p, q] = \frac{\hbar}{i} \hat{1}$$

que leva à relação de incerteza de Heisenberg

$$\sigma_p \sigma_q \geq \frac{\hbar}{2}$$

que significa que *para qualquer vetor de estado* o produto $\sigma_p \sigma_q$ é sempre maior que $\hbar/2$. Os estados para os quais essa última desigualdade vale como igualdade são por isso chamados *estados de incerteza mínima*.

O leitor atento poderá ter sentido um certo desconforto com a arbitrariedade envolvida na eliminação do quadrado de uma *parte real* na quantidade dada pela desigualdade de Schwartz como limite inferior para o produto das variâncias $\sigma_g^{(\varphi)^2}$ e $\sigma_h^{(\varphi)^2}$ na última passagem de (2.19), e notado que, a rigor, esse passo de fato restringe o conjunto de estados de incerteza mínima, pois para que a relação de incerteza de Heisenberg valha como igualdade é preciso não apenas que a desigualdade de Schwartz valha como igualdade mas também que o quadrado de parte real desprezado se anule. Precisamente esse mesmo passo foi discutido de forma mais completa por Schrödinger em uma comunicação pouco conhecida à Academia de Ciências da Prússia, datada de 1930³. Nessa comunicação Schrödinger conserva a contribuição da parte real, sendo portanto levado à desigualdade

$$\sigma_g^{(\varphi)^2} \sigma_h^{(\varphi)^2} \geq \left(\frac{\langle \varphi | (g_{(\varphi)} h_{(\varphi)} + h_{(\varphi)} g_{(\varphi)}) | \varphi \rangle}{2} \right)^2 + \left| \frac{\langle \varphi | [g, h] | \varphi \rangle}{2} \right|^2$$

que interpreta como estabelecendo uma relação entre *três* quantidades, a saber 1. o produto dos quadrados dos desvios quadráticos médios, 2. o módulo quadrado da metade do valor médio do comutador e 3. “uma quantidade que pode ser definida como o quadrado do desvio médio do produto, com a condição de que a não comutatividade seja levada em conta, isto é, o desvio médio do produto deve ser definido como a média aritmética de

$$\langle \varphi | g_{(\varphi)} h_{(\varphi)} | \varphi \rangle \quad \text{e} \quad \langle \varphi | h_{(\varphi)} g_{(\varphi)} | \varphi \rangle$$

que são as expressões ‘mistas’ completamente análogas a $\sigma_g^{(\varphi)^2}$ e $\sigma_h^{(\varphi)^2}$ ”. Esta terceira quantidade recebe hoje o nome de *covariança* de g e h em $|\varphi\rangle$. Designando-a como $\sigma_{gh}^{(\varphi)^2}$ e definindo a *matriz de covariança*

³E. Schrödinger, *Proceedings of the Prussian Academy of Sciences* **XIX**, pp. 296-303 (1930), reproduzida em *Erwin Schrödinger, Gesammelte Abhandlungen*, Band 3, Verlag der Österreichischen Akademie der Wissenschaften, Wien, pp. 348-356 (1984). Uma tradução para o inglês, anotada, dessa contribuição foi preparada por A. Angelow e M.-C. Batoni, “em honra ao septuagésimo aniversário da Relação de Incerteza de Schrödinger”.

$$\Sigma_{g,h} \equiv \begin{pmatrix} \sigma_g^{(\varphi)2} & \sigma_{gh}^{(\varphi)} \\ \sigma_{hg}^{(\varphi)} & \sigma_h^{(\varphi)2} \end{pmatrix}$$

é fácil verificar que a relação geral de incerteza de Schrödinger pode ser escrita em termos do determinante de $\Sigma_{g,h}$ como

$$\det(\Sigma_{g,h}) \geq \left| \frac{\langle \varphi | [g, h] | \varphi \rangle}{2} \right|^2, \quad (2.21)$$

e a forma geral dessa relação, no caso das variáveis p e q , satisfazendo as relações de comutação de Heisenberg, Born e Jordan, é dada por

$$\det(\Sigma_{g,h}) \geq \left(\frac{\hbar}{2} \right)^2.$$

Essa relação difere da relação de incerteza de Heisenberg por termos que envolvem a covariância $\sigma_{pq}^{(\varphi)}$. Schrödinger observa ainda que na teoria clássica das flutuações estatísticas o anulamento da covariância é uma condição *necessária* (mas não suficiente) para a independência estatística das flutuações de cada uma das variáveis, e que mesmo no caso de variáveis dinâmicas quânticas canonicamente conjugadas (isto é, satisfazendo relações de comutação do tipo da relação de Heisenberg, Born e Jordan para p e q), existem vetores de estado $|\varphi\rangle$ para os quais a covariância *não é nula*, indicando a presença de correlações entre as respectivas flutuações quânticas. Um exemplo concreto disso será encontrado e discutido mais adiante, v. seção 3.3.3.

Finalmente, é importante não perder de vista o fato de que o grau de indefinição do valor de um observável num dado estado $|\varphi\rangle$, medido pelo desvio quadrático médio $\sigma^{(\varphi)}$, exprime uma característica *da distribuição de probabilidades* sobre o espectro do observável, calculada a partir desse estado e dos autovetores $|\phi_n\rangle$ do observável como $\{p_n \equiv |\langle \phi_n | \varphi \rangle|^2\}$. Segundo os postulados interpretativos, o valor obtido em cada medição deve ser sempre um dos autovalores do observável. No caso de um estado com desvio quadrático médio suficientemente pequeno e de um observável com espectro discreto ele poderá diferir do valor médio (que, por outro lado, em geral diferirá dos autovalores) por uma quantidade necessariamente maior que o desvio calculado. O fato de que o desvio é, nesse sentido, pequeno, reflete apenas o fato de que a *probabilidade* predita para a obtenção de autovalores muito discrepantes da média é também pequena.

2.2.5 Muitos graus de liberdade, espaços produto e sistemas compostos.

A situação considerada nas seções precedentes corresponde, a rigor, à de um sistema quântico cujo análogo clássico tem um único grau de liberdade, ao qual estão associadas as variáveis canônicas de posição q e momento p . O espaço de fases quântico correspondente é um espaço de Hilbert no qual essas variáveis canônicas são implementadas como operadores hermiteanos satisfazendo as relações de comutação de Heisenberg, Born e Jordan $[q, p] = i\hbar$. Os estados $|\varphi\rangle$ desse sistema quântico são representados por vetores de norma 1 desse espaço, podendo ser expressos em termos de um conjunto ortonormal e completo de vetores de base $\{|\phi_n\rangle\}_{n=1,2,\dots}$ como

$$|\varphi\rangle = \sum_n c_n |\phi_n\rangle, \quad \sum_n |c_n|^2 = 1.$$

Para tratar sistemas cujo análogo clássico envolve mais que um grau de liberdade é preciso, portanto, estender de forma apropriada o tratamento considerado até aqui. Um ponto de vista alternativo, que em muitas situações pode ser não apenas conveniente mas também natural, consiste em considerar esses sistemas como constituídos de vários *sub-sistemas*. Isso é análogo à possibilidade de considerar, por exemplo, o problema (clássico) de uma partícula em duas dimensões como um problema de duas partículas em uma única dimensão, tomando cada uma delas como um “sub-sistema”; alternativamente, o problema de duas partículas em três dimensões pode também ser visto como o problema de uma única partícula em seis dimensões. A eventual artificialidade da separação em subsistemas no primeiro caso aparece, no segundo caso, associada ao tratamento dos dois subsistemas de uma partícula como um sistema único. Além disso, a própria identificação dos subsistemas envolve arbitrariedades, como pode ser visto descrevendo o sistema de duas partículas em termos de variáveis de centro de massa e relativas, e notando que estas últimas variáveis podem também ser utilizadas para definir os subsistemas. O tratamento que segue será, por simplicidade, limitado a dois subsistemas, ou graus de liberdade, apenas. As idéias gerais envolvidas nesse caso se estendem imediatamente, no entanto, para o caso de um número finito qualquer de graus de liberdade.

Para um sistema quântico com dois graus de liberdade (ou constituído de dois sub-sistemas de um grau de liberdade) existem duas variáveis canônicas de posição, q_1 e q_2 , às quais estão associados respectivamente momentos p_1 e p_2 . Essas variáveis devem ser implementadas em termos de operadores hermiteanos em um espaço de Hilbert \mathcal{H} que desempenha o papel de espaço de fases do sistema. Uma questão a ser considerada, portanto, é a relação existente entre esse espaço de fases e os espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 que correspondem a cada um dos sub-sistemas quando considerado isoladamente. Outra questão diz respeito às relações de comutação que devem ser impostas entre os operadores q_j e p_j , correspondentes às que se

aplicam ao caso de um único grau de liberdade, mas que agora envolvem os dois pares de variáveis canônicas.

Quanto à primeira questão, é razoável supor que o espaço de Hilbert correspondente ao conjunto de dois subsistemas inclua entre os seus vetores alguns que correspondem a estados quânticos bem definidos de cada um deles, descritos respectivamente por particulares vetores $|\varphi\rangle$ de \mathcal{H}_1 e $|\psi\rangle$ de \mathcal{H}_2 . Esses vetores particulares serão escritos como $|\varphi\rangle \otimes |\psi\rangle$, e designados como *produto tensorial* dos dois vetores $|\varphi\rangle$ e $|\psi\rangle$. É claro, no entanto, que o conjunto de todos os objetos desse tipo não constitui um espaço linear, pois combinações lineares de produtos tensoriais de dois vetores em geral *não* são, por sua vez, produtos tensoriais de dois vetores. De fato, considerando, por exemplo, a combinação linear

$$|\Psi_{1,2}\rangle \equiv z_1|\varphi_1\rangle \otimes |\psi_1\rangle + z_2|\varphi_2\rangle \otimes |\psi_2\rangle, \quad \text{com} \quad |\varphi_1\rangle \neq |\varphi_2\rangle, \quad |\psi_1\rangle \neq |\psi_2\rangle, \quad z_1, z_2 \neq 0$$

não corresponde a nenhum objeto do tipo $|\bar{\varphi}\rangle \otimes |\bar{\psi}\rangle$ com $|\bar{\varphi}\rangle$ em \mathcal{H}_1 e $|\bar{\psi}\rangle$ em \mathcal{H}_2 . Na realidade, a forma mais simples de construir um espaço de Hilbert que contenha, em particular, os produtos tensoriais de vetores de \mathcal{H}_1 e \mathcal{H}_2 como vetores é tomar alguma base (por conveniência ortonormal) $\{\phi_n\}_{n=1,2,\dots}$ de \mathcal{H}_1 e também alguma base (ortonormal) $\{\chi_k\}_{k=1,2,\dots}$ de \mathcal{H}_2 , e tomar para o espaço de Hilbert \mathcal{H} do sistema conjunto o espaço gerado pela base formada pelos produtos tensoriais

$$\{|\Phi_{n,k}\rangle\}_{n,k=1,2,\dots} \equiv \{|\phi_n\rangle \otimes |\chi_k\rangle\}_{n,k=1,2,\dots}.$$

O espaço assim construído é chamado *produto tensorial* dos espaços \mathcal{H}_1 e \mathcal{H}_2 , o que é designado pela notação $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, e a base $\{|\Phi_{n,k}\rangle\}_{n,k=1,2,\dots}$ é chamada a *base produto* das duas bases $\{\phi_n\}_{n=1,2,\dots}$ e $\{\chi_k\}_{k=1,2,\dots}$. Os vetores do espaço produto \mathcal{H} são expressos em termos da base produto como

$$|\psi\rangle = \sum_{n,k} c_{n,k} |\Phi_{n,k}\rangle = \sum_{n,k} c_{n,k} |\phi_n\rangle \otimes |\chi_k\rangle, \quad \sum_{n,k} |c_{n,k}|^2 = 1 \quad (2.22)$$

e um produto escalar em \mathcal{H} pode ser definido a partir dos produtos escalares de \mathcal{H}_1 e \mathcal{H}_2 através da relação

$$\langle \Phi_{n',k'} | \Phi_{n,k} \rangle \equiv \langle \phi_{n'} | \phi_n \rangle \langle \chi_{k'} | \chi_k \rangle = \delta_{n'n} \delta_{k'k}.$$

Isso dá, para o produto escalar de dois vetores $|\Psi_1\rangle$ e $|\Psi_2\rangle$ (cujos coeficientes de expansão na base produto são respectivamente $c_{n,k}^{(1)}$ e $c_{n,k}^{(2)}$)

$$\langle \Psi_1 | \Psi_2 \rangle \equiv \sum_{n,k} c_{n,k}^{(1)*} c_{n,k}^{(2)}.$$

É fácil verificar, de fato, que essa definição satisfaz todas as condições requeridas para um produto escalar hermiteano no espaço produto \mathcal{H} .

A segunda questão a ser tratada diz respeito às relações de comutação a serem impostas sobre as variáveis canônicas q_j e p_j ($j = 1, 2$), consideradas como operadores hermiteanos *no espaço produto* \mathcal{H} . Tomando como ponto de partida as variáveis de cada um dos dois subsistemas, \bar{q}_1, \bar{p}_1 e \bar{q}_2, \bar{p}_2 , definidas respectivamente nos espaços \mathcal{H}_1 e \mathcal{H}_2 , extensões naturais para o espaço produto podem ser obtidas através das definições

$$q_1|\Psi\rangle \equiv \sum_{n,k} c_{n,k}(\bar{q}_1|\phi_n\rangle) \otimes |\chi_k\rangle, \quad q_2|\Psi\rangle \equiv \sum_{n,k} c_{n,k}|\phi_n\rangle \otimes (\bar{q}_2|\chi_k\rangle)$$

além de duas relações análogas para p_1 e p_2 . Desse modo q_1 e p_1 agem da mesma forma que \bar{q}_1 e \bar{p}_1 sobre os fatores da base produto que correspondem ao espaço \mathcal{H}_1 enquanto agem como o operador identidade sobre os fatores que correspondem ao espaço \mathcal{H}_2 . Isso é usualmente expresso através da notação

$$q_1 \equiv \bar{q}_1 \otimes \hat{1}_2, \quad q_2 \equiv \hat{1}_1 \otimes \bar{q}_2$$

com duas relações análogas para p_1 e p_2 . Os operadores $\hat{1}_1$ e $\hat{1}_2$ são os operadores identidade dos espaços \mathcal{H}_1 e \mathcal{H}_2 respectivamente. É fácil verificar que os operadores q_j e p_j assim definidos satisfazem, como operadores em \mathcal{H} , as relações de comutação

$$[q_j, p_l] = i\hbar\delta_{jl}.$$

Essas relações de comutação implicam, em particular, na *compatibilidade* de variáveis de posição e de momento associadas a graus de liberdade *diferentes* do sistema composto.

Uma característica fundamental dessa extensão para sistemas com mais de um grau de liberdade (ou, equivalentemente, para sistemas que podem ser vistos como sendo constituídos de vários sub-sistemas) é a de que o espaço de fases do sistema composto resulta consideravelmente *maior*, devido à presença da estrutura linear (isto é, combinações lineares de estados são também estados possíveis do sistema), que o conjunto de estados obtidos especificando o estado de cada um dos subsistemas envolvidos. Este último conjunto contém apenas os estados que são produtos tensoriais de estados bem definidos de cada um dos subsistemas, que não é fechado com relação à estrutura linear, no sentido de que combinações lineares de estados deste tipo não são, em geral, também estados deste mesmo tipo. Em um trabalho de 1935, Schrödinger⁴ afirmou que esse fato “não é *um*, mas *o* traço característico da mecânica quântica, que obriga o seu afastamento completo das linhas de pensamento clássicas”. Para ter uma idéia do seu

⁴E. Schrödinger, *Discussion of probability relations between separated systems*, Proc. Camb. Phil. Soc. **31**, 555 (1935).

significado, pode-se notar, primeiramente, que a quantidade $P_{n,k} = |c_{n,k}|^2$ (v. equação (2.22)) deve ser interpretada como a probabilidade de que no estado $|\Psi\rangle$ do sistema composto o primeiro sub-sistema seja observado no estado $|\phi_n\rangle$ e o segundo subsistema seja observado no estado $|\chi_k\rangle$. Por outro lado, a probabilidade de que nesse estado o primeiro sub-sistema seja observado no estado $|\phi_n\rangle$ *independentemente do estado do segundo sub-sistema* será $P_{n,x} = \sum_k |c_{n,k}|^2$; enquanto a probabilidade de que o segundo sub-sistema seja observado no estado $|\chi_k\rangle$ *independentemente do estado do primeiro sub-sistema* será $P_{x,k} = \sum_n |c_{n,k}|^2$. É fácil ver então que, em geral, $P_{n,k} \neq P_{n,x}P_{x,k}$, o que significa que as probabilidades $P_{n,x}$ e $P_{x,k}$ *não são independentes*, indicando a presença de correlações entre os dois sistemas. Schrödinger exprimiu esse fato dizendo que, nessas condições, os dois sub-sistemas estão *emaranhados* (“entangled”). Isso pode ser visto facilmente tomando o exemplo

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\phi_1\rangle \otimes |\chi_1\rangle + |\phi_2\rangle \otimes |\chi_2\rangle),$$

pois nesse caso $P_{1,1} = 1/2$ e $P_{1,x} = P_{x,1} = 1/2$, de modo que $P_{1,x}P_{x,1} = 1/4$, e indica a existência de correlações de natureza quântica entre os dois subsistemas. Por outro lado, se $|\Psi\rangle$ é um estado que é o produto tensorial $|\varphi\rangle \otimes |\psi\rangle$ de estados bem definidos de cada um dos subsistemas, segue que

$$|\varphi\rangle \otimes |\psi\rangle = \sum_{n,k} c_n^{(1)} |\phi_n\rangle \otimes c_k^{(2)} |\chi_k\rangle, \quad \sum_n |c_n^{(1)}|^2 = \sum_k |c_k^{(2)}|^2 = 1$$

e portanto $c_{n,k} = c_n^{(1)}c_k^{(2)}$ e conseqüentemente $P_n = |c_n^{(1)}|^2$ e $P_k = |c_k^{(2)}|^2$, de modo que $P_nP_k = P_{n,k}$, isto é, as probabilidades P_n e P_k são independentes e os dois subsistemas não estão emaranhados.

2.3 Avatares da dinâmica.

A descrição da evolução no tempo das propriedades observáveis de um sistema quântico aparece encarnada⁵ em várias formas diferentes, duas das quais foram adotadas respectivamente por Heisenberg e por Schrödinger já por ocasião da dupla invenção da teoria há três quartos de século. O propósito desta seção é apresentar essas duas descrições e mais uma terceira, de

⁵**Avatar.** [Do sânscr. *avatāra*, ‘descida’ (do Céu à Terra), pelo fr. *avatar*.] *S. m.* **1.** *Filos.* Reencarnação de um Deus, e, especialmente, no Hinduísmo, reencarnação do deus Vixnu. **2.** Transformação, transfiguração, metamorfose: “Aqui era a laranjeira-cravo junto da qual o vira, como em um a v a t a r , como em uma transfiguração, risonho, franco, comunicativo, sob o aspecto que em um momento a cativara.” (Júlio Ribeiro, *A Carne*, p. 91.) (Dicionário Aurélio).

uso também bastante corrente, enfatizando ao mesmo tempo o seu núcleo comum, que é um operador unitário dependente do tempo que rege, em última análise, a evolução temporal do sistema quântico. A sub-seção final trata ainda de uma relação importante existente entre as escalas de tempo envolvidas na evolução de um determinado estado do sistema e o grau de indeterminação da energia nesse estado, sendo este grau de indeterminação entendido no sentido utilizado na discussão das relações de incerteza.

2.3.1 O operador de evolução e as descrições de Heisenberg, de Schrödinger e “de interação”.

Como indicado na seção 2.1, com base na discussão da equação de movimento postulada no contexto da mecânica das matrizes, na *descrição de Heisenberg* a evolução temporal dos operadores hermiteanos que representam as variáveis dinâmicas (observáveis) de um sistema quântico é definida através de um operador também hermiteano H (o operador hamiltoniano do sistema físico considerado) por meio da equação de movimento

$$i\hbar \frac{dg(t)}{dt} = [g(t), H] \quad (2.23)$$

que, tendo em conta a relação de comutação postulada para coordenadas e momentos canonicamente conjugados, leva a equações de movimento para essas variáveis que mantêm uma analogia formal completa com as equações canônicas de Hamilton da mecânica clássica.

Supondo que H não dependa explicitamente do tempo, a equação de movimento (2.23), suplementada pela condição inicial de que $g(t = t_0) = g$, admite a solução formal

$$g(t) = e^{\frac{i}{\hbar}H(t-t_0)} g e^{-\frac{i}{\hbar}H(t-t_0)} \equiv U^\dagger(t, t_0) g U(t, t_0) \quad (2.24)$$

que revela explicitamente o fato de que a evolução temporal é efetivamente regida pelo *operador de evolução*

$$U(t, t_0) \equiv e^{-\frac{i}{\hbar}H(t-t_0)}. \quad (2.25)$$

Dada a hermiticidade de H , $U(t, t_0)$ é um operador unitário dependente do tempo. Esse operador pode ser definido, alternativamente, como a solução da equação diferencial

$$i\hbar \frac{dU(t, t_0)}{dt} = H U(t, t_0) \quad (2.26)$$

com a condição inicial $U(t_0, t_0) = \hat{1}$. A sua forma explícita, além de satisfazer essa condição inicial, mostra que ele satisfaz também às relações de unitariedade

$$U^\dagger(t, t_0) = U^{-1}(t, t_0) = U(t_0, t)$$

além da “lei de composição”

$$U(t, t')U(t', t_0) = U(t, t_0),$$

o que caracteriza o conjunto de transformações definidas pelos operadores $U(t, t')$ como um grupo contínuo (dado que o tempo é uma variável contínua) de transformações unitárias.

Nesta descrição da dinâmica, os observáveis são os objetos encarregados de dar conta da evolução temporal do sistema, sendo para isso associados a operadores hermiteanos que dependem do tempo da forma indicada pela equação (2.24). Como visto na seção anterior, quantidades essenciais para a interpretação estatística da teoria são objetos do tipo $|\langle \phi_n | \varphi \rangle|^2$, onde os $|\phi_n\rangle$ são autovetores do observável considerado e $|\varphi\rangle$ é o vetor de estado do sistema. Esses objetos são interpretados como probabilidades de obtenção de autovalores γ_n numa medição de g . Para calcular as probabilidades para um observável g no tempo t é preciso então, dado o vetor de estado do sistema, determinar autovetores e autovalores de $g(t)$.

A forma particular (2.24) de $g(t)$ permite relacionar facilmente esses autovetores e autovalores com os que correspondem ao tempo inicial de referência t_0 , através do operador de evolução. De fato, supondo que autovetores $|\phi_n\rangle$ e autovalores γ_n tenham sido obtidos para $g(t = t_0)$, isto é

$$g|\phi_n\rangle = \gamma_n|\phi_n\rangle$$

então a unitariedade de $U(t, t_0)$ permite escrever imediatamente

$$U^\dagger(t, t_0)gU(t, t_0)U^\dagger(t, t_0)|\phi_n\rangle = \gamma_n U^\dagger(t, t_0)|\phi_n\rangle$$

ou seja

$$g(t) U^\dagger(t, t_0)|\phi_n\rangle = \gamma_n U^\dagger(t, t_0)|\phi_n\rangle.$$

Isso mostra que $g = g(t = t_0)$ e $g(t)$ tem o mesmo espectro, e que os autovetores de $g(t)$ podem ser obtidos dos autovetores de g através do operador unitário $U^\dagger(t, t_0)$. Desse modo, a probabilidade calculada no tempo t_0 como

$$p_n(t_0) = |\langle \phi_n | \varphi \rangle|^2$$

evolui dinamicamente para a probabilidade, no tempo t ,

$$p_n(t) = |\langle \phi_n | U(t, t_0) | \varphi \rangle|^2. \quad (2.27)$$

É importante notar que o vetor de estado é o mesmo nos dois casos, dado que, na descrição de Heisenberg, os agentes dinâmicos são os observáveis.

A expressão obtida para $p_n(t)$ sugere imediatamente uma realização alternativa da dinâmica que leva a resultados idênticos para as probabilidades. Ela consiste em tratar os observáveis como operadores *independentes do tempo* e encarregar o vetor de estado da atribuição de dar conta da evolução temporal do sistema. Nesse caso a equação de movimento para o vetor de estado deve ser

$$i\hbar \frac{d|\varphi(t)\rangle}{dt} = H|\varphi(t)\rangle \quad (2.28)$$

cuja solução formal satisfazendo à condição inicial $|\varphi(t = t_0)\rangle = |\varphi\rangle$, no caso de um hamiltoniano H que não depende explicitamente do tempo, é

$$|\varphi(t)\rangle = e^{-\frac{i}{\hbar}H(t-t_0)}|\varphi\rangle = U(t, t_0)|\varphi\rangle.$$

É então evidente que se obtém o mesmo resultado para $p(t)$ seja analisando o vetor de estado fixo $|\varphi\rangle$ em termos dos autovetores dependentes do tempo $U^\dagger(t, t_0)|\phi_n\rangle$, seja analisando o vetor de estado dependente do tempo $|\varphi(t)\rangle$ em termos dos autovetores fixos $|\phi_n\rangle$. Esta última descrição da dinâmica quântica é conhecida como *descrição de Schrödinger*.

Existe uma forma, com utilidade prática considerável, de reescrever quantidades como $p_n(t)$ uma vez que o problema de autovalores para H tenha sido completamente resolvido. Ela tem de fato um peso considerável na importância freqüentemente atribuída à resolução desse problema. Supondo por simplicidade que o espectro seja puramente discreto, isso significa que se conheça os autovalores de H juntamente com o conjunto completo de autovetores correspondentes aos autovalores, isto é, que se tenha obtido todas as soluções do problema de autovalores

$$H|E_k\rangle = E_k|E_k\rangle.$$

O conhecimento dos autovetores permite escrever a resolução da unidade

$$\hat{1} = \sum_k |E_k\rangle\langle E_k|$$

em termos da qual é possível escrever o operador de evolução como

$$U(t, t_0) = \sum_k e^{-\frac{i}{\hbar}H(t-t_0)}|E_k\rangle\langle E_k| = \sum_k |E_k\rangle e^{-\frac{i}{\hbar}E_k(t-t_0)}\langle E_k|.$$

Usando esta última expressão em (2.27) o que se obtém é

$$p_n(t) = \left| \sum_k \langle \phi_n | E_k \rangle e^{-\frac{i}{\hbar} E_k (t-t_0)} \langle E_k | \varphi \rangle \right|^2$$

que exprime $p_n(t)$ em termos dos coeficientes de expansão dos vetores $|\phi_n\rangle$ e $|\varphi\rangle$ na base $\{|E_k\rangle\}$ e de fases dependentes do tempo e dos autovalores E_k . De uma forma completamente análoga se vê que a solução da equação de movimento para o vetor de estado na descrição de Schrödinger pode ser expressa como

$$|\varphi(t)\rangle = \sum_k |E_k\rangle e^{-\frac{i}{\hbar} E_k (t-t_0)} \langle E_k | \varphi \rangle$$

e também que a solução da equação de movimento para o observável $g(t)$ na descrição de Heisenberg fica

$$g(t) = \sum_{kk'} |E_k\rangle e^{\frac{i}{\hbar} (E_k - E_{k'}) (t-t_0)} \langle E_k | g | E_{k'} \rangle \langle E_{k'} |.$$

As descrições de Heisenberg e de Schrödinger não esgotam as alternativas de formulação para a dinâmica quântica. Também de uso freqüente são formulações “intermediárias”, em que o papel de agentes dinâmicos é de certa forma repartido entra as variáveis dinâmicas e os vetores de estado. A situação típica em que essas descrições são utilizadas é aquela em que o operador hamiltoniano, não dependente explicitamente do tempo, é escrito como a soma de duas partes que também não dependem explicitamente do tempo

$$H = H_0 + H_1.$$

Isso permite introduzir, além do operador de evolução definido como em (2.25), um operador de evolução também unitário mas associado somente a H_0

$$U_0(t, t_0) \equiv e^{-\frac{i}{\hbar} H_0 (t-t_0)}.$$

Com isso o objeto relevante para o cálculo de $p_n(t)$ pode ser reescrito sob a forma

$$\langle \phi_n | U(t, t_0) | \varphi \rangle = \langle \phi_n | U_0(t, t_0) U_0^\dagger(t, t_0) U(t, t_0) | \varphi \rangle$$

que pode ser interpretada em termos do produto escalar com o vetor de estado dependente do tempo $|\varphi^{(I)}(t)\rangle \equiv U_0^\dagger(t, t_0) U(t, t_0) | \varphi \rangle$ do autovetor também dependente do tempo $U_0^\dagger(t, t_0) | \phi_n \rangle$. Este corresponde porém ao autovetor da variável dinâmica dependente do tempo $g^{(I)}(t) \equiv U_0^\dagger(t, t_0) g U_0(t, t_0)$, que satisfaz a equação de movimento

$$i\hbar \frac{dg^{(I)}(t)}{dt} = [g^{(I)}(t), H_0] \quad (2.29)$$

enquanto o vetor de estado satisfaz

$$i\hbar \frac{d|\varphi^{(I)}(t)\rangle}{dt} = U_0^\dagger(t, t_0) H_1 U_0(t, t_0) |\varphi_I(t)\rangle \equiv H_1^{(I)}(t) |\varphi_I(t)\rangle. \quad (2.30)$$

A dependência temporal do hamiltoniano $H_1^{(I)}(t)$ provém dos operadores U_0^\dagger e U_0 e é portanto também a prescrita por uma equação como a (2.29). Desse modo, nesta descrição todos os operadores evoluem no tempo como numa descrição de Heisenberg, mas com relação apenas a H_0 . Por outro lado, o vetor de estado evolui como em uma descrição de Schrödinger, mas *apenas com a parte H_1 do hamiltoniano*, modificada de forma a incluir a dependência temporal característica da descrição de Heisenberg relativa a H_0 . Este tipo de descrição é chamado *descrição de interação*, pelo fato de que concentra *todos* os efeitos da parte “de interação” H_1 do hamiltoniano na evolução temporal do vetor de estado.

2.3.2 Extensão para hamiltonianos que dependem explicitamente do tempo.

Com alguns cuidados de natureza essencialmente técnica, as diferentes formas de implementação da dinâmica descritas acima podem ser estendidas para casos em que o hamiltoniano é um operador hermiteano que depende explicitamente do tempo. Isso permite descrever sistemas quânticos sujeitos a “forças externas” cuja dependência com o tempo é prescrita *a priori*, independentemente do comportamento do sistema quântico sujeito à sua ação. Além disso, em vista de que o hamiltoniano $H_1^{(I)}(t)$ que aparece na equação de evolução do vetor de estado na descrição de interação (2.30) depende do tempo, os aspectos técnicos envolvidos nesta extensão são relevantes também no tratamento dessa equação.

No caso de um hamiltoniano $H(t)$ dependente do tempo, um operador de evolução $U(t, t_0)$ pode ser definido generalizando trivialmente a equação diferencial (2.26) para

$$i\hbar \frac{dU(t, t_0)}{dt} = H(t)U(t, t_0), \quad (2.31)$$

mantendo inalterada a condição inicial $U(t_0, t_0) = \hat{1}$. A equação (2.31) pode também ser escrita sob a forma da equação integral, que incorpora a condição inicial,

$$U(t, t_0) = \hat{1} - \frac{i}{\hbar} \int_{t_0}^t dt' H(t') U(t', t_0). \quad (2.32)$$

A equivalência dessa equação integral com a equação diferencial (2.31) pode ser verificada imediatamente por simples diferenciação com relação a t . Neste caso, no entanto, não é possível escrever uma solução formal explícita sob forma exponencial como no caso independente do tempo. A dificuldade para este passo provém de que, em geral, $[H(t), H(t')] \neq 0$ se $t \neq t'$, isto é, em geral o hamiltoniano $H(t)$ não comuta consigo mesmo em tempos diferentes. Para perceber a melhor basta considerar, como exemplo, um caso em que o hamiltoniano é independente do tempo e igual a H_1 no intervalo $t_0 < t < t_0 + \Delta t$, e igualmente independente do tempo e igual a H_2 no intervalo subsequente $t_0 + \Delta t < t < t_0 + 2\Delta t$, mas com $H_1 \neq H_2$ e $[H_1, H_2] \neq 0$. Os operadores de evolução para cada um desses intervalos de tempo podem ser escritos como

$$U_1(t_0 + \Delta t, t_0) = e^{-\frac{i}{\hbar} H_1 \Delta t} \quad \text{e} \quad U_2(t_0 + 2\Delta t, t_0 + \Delta t) = e^{-\frac{i}{\hbar} H_2 \Delta t}.$$

Eles descrevem a evolução temporal do sistema de t_0 a $t_0 + \Delta t$ e de $t_0 + \Delta t$ a $t_0 + 2\Delta t$ respectivamente, de modo que a evolução no intervalo todo de t_0 a $t_0 + 2\Delta t$ é descrita por

$$\begin{aligned} U(t_0 + 2\Delta t, t_0) &= U_2(t_0 + 2\Delta t, t_0 + \Delta t) U_1(t_0 + \Delta t, t_0) \\ &= e^{-\frac{i}{\hbar} H_2 \Delta t} e^{-\frac{i}{\hbar} H_1 \Delta t} \neq e^{-\frac{i}{\hbar} (H_2 \Delta t + H_1 \Delta t)} \end{aligned}$$

onde a inequivalência das duas últimas formas exponenciais provém da não comutatividade de H_1 e H_2 , e pode ser facilmente verificada através de expansões em série de potências, que levam a ordenamentos diferentes desses operadores nos dois casos. Isso mostra que uma solução formal ingênua da equação (2.31), escrita possivelmente como

$$U(t, t_0) \rightarrow e^{-\frac{i}{\hbar} \int_{t_0}^t H(t) dt} \quad (?),$$

o que seria de fato a solução correta no caso em que $H(t)$ fosse uma função numérica, é na realidade excessivamente ingênua neste caso. A solução que tem devidamente em conta a não comutatividade do hamiltoniano em tempos diferentes pode porém ser pensada, como no caso do exemplo acima, em termos de uma sucessão de transformações unitárias infinitesimais, que avançam o tempo em intervalos dt durante os quais a variação de $H(t)$ é desprezível:

$$U(t + dt, t) = \hat{1} - \frac{i}{\hbar} H(t) dt + \mathcal{O}(dt^2).$$

A unitariedade (até termos de ordem dt inclusive) dessa transformação infinitesimal decorre diretamente da hermiticidade de $H(t)$, e a unitariedade da transformação finita, vista como uma sucessão de tais transformações infinitesimais, decorre de que o produto de operadores unitários é por sua vez um operador unitário.

Portanto, no caso de um hamiltoniano explicitamente dependente do tempo a equação (2.31), juntamente com a respectiva condição inicial, define também um operador de evolução unitário. É imediato também que esse operador satisfaz propriedades e obedece a uma “lei de composição” idênticas às que valem para o caso em que H é independente do tempo. Desse modo, todos os resultados obtidos nesse caso que dependem apenas da existência de um operador de evolução unitário são válidos também para o caso em que o hamiltoniano depende do tempo. Mais explicitamente, a evolução temporal dos observáveis na descrição de Heisenberg é dada por $g(t) \equiv U^\dagger(t, t_0)gU(t, t_0)$, como na última forma da equação (2.24). A equação de movimento correspondente pode ser obtida usando a equação (2.31) e sua adjunta

$$i\hbar \frac{dU^\dagger(t, t_0)}{dt} = -U^\dagger(t, t_0)H(t),$$

o que leva a

$$i\hbar \frac{dg(t)}{dt} = U^\dagger(t, t_0)g(H(t)U(t, t_0)) - (U^\dagger(t, t_0)H(t))gU(t, t_0) = [g(t), H_H(t)] \quad (2.33)$$

onde $H_H(t) \equiv U^\dagger(t, t_0)H(t)U(t, t_0)$ é a “versão Heisenberg” do hamiltoniano dependente do tempo $H(t)$. Ela difere da “versão Schrödinger” por conter os operadores de Heisenberg $p(t)$, $q(t)$ onde esta última contém os operadores independentes do tempo p e q . É importante notar que, devido à unitariedade da transformação que relaciona esses dois conjuntos de operadores, eles satisfazem as mesmas relações de comutação (com tempos iguais, no caso dos operadores de Heisenberg), isto é, $[q, p] = [q(t), p(t)] = i\hbar\hat{1}$.

É possível também formular uma descrição de Schrödinger na qual o vetor de estado satisfaz à equação diferencial de evolução

$$i\hbar \frac{d|\varphi(t)\rangle}{dt} = H(t)|\varphi(t)\rangle$$

bem como uma representação de interação em que os observáveis satisfazem à equação de movimento

$$i\hbar \frac{dg^{(I)}(t)}{dt} = [g^{(I)}(t), H_0(t)]$$

enquanto os vetores de estado evoluem de acordo com a equação

$$i\hbar \frac{d|\varphi^{(I)}(t)\rangle}{dt} = H_1^{(I)}(t)|\varphi^{(I)}(t)\rangle$$

com $H_1^{(I)}(t) = U_0^\dagger(t, t_0)H_1(t)U_0(t, t_0)$. Desde que H_0 e H_1 sejam sempre hermiteanos, é possível que eles sejam tomados como dependentes do tempo.

2.3.3 Relações de incerteza energia-tempo.

Se um observável g é compatível com o hamiltoniano H (independente do tempo) que determina a evolução temporal do sistema quântico, no sentido definido anteriormente de que $[g, H] = 0$, é fácil ver que o desvio quadrático médio de g não depende do tempo. A forma mais simples de obter esse resultado é usar a descrição de Heisenberg, que dá nesse caso $g(t) = U^\dagger(t, t_0)gU(t, t_0) = g$, pois nesse caso g e U comutam. Portanto

$$\sigma_g^{(\varphi)^2}(t) = \langle \varphi | g(t)^2 | \varphi \rangle - \langle \varphi | g(t) | \varphi \rangle^2 = \langle \varphi | g^2 | \varphi \rangle - \langle \varphi | g | \varphi \rangle^2.$$

Os observáveis desse tipo são chamados *constantes do movimento*. No caso em que g não é uma constante do movimento, a relação geral de incerteza (2.20) impõe a limitação

$$\sigma_g^{(\varphi)}(t)\sigma_H^{(\varphi)} \geq \left| \frac{\langle \varphi | [g(t), H] | \varphi \rangle}{2i} \right|$$

que, usando a equação de movimento de Heisenberg, pode ser escrita como

$$\sigma_g^{(\varphi)}(t)\sigma_H^{(\varphi)} \geq \frac{\hbar}{2} \left| \langle \varphi | \frac{dg(t)}{dt} | \varphi \rangle \right|.$$

Essa relação contém, por um lado, o desvio quadrático médio associado ao observável $g(t)$ e, por outro lado, uma taxa média de variação do valor médio de $g(t)$. Essas duas quantidades podem ser combinadas para definir um tempo característico para a evolução do observável o intervalo de tempo $\tau_g^{(\varphi)}$ durante o qual o seu valor médio sofre uma variação da ordem do seu desvio quadrático médio, isto é

$$\tau_g^{(\varphi)} \equiv \frac{\sigma_g^{(\varphi)}(t)}{\left| \langle \varphi | \frac{dg(t)}{dt} | \varphi \rangle \right|}.$$

Essa definição leva então à *relação de incerteza energia-tempo*

$$\sigma_H^{(\varphi)}\tau_g^{(\varphi)} \geq \frac{\hbar}{2} \quad (2.34)$$

que exprime uma limitação para $\tau_g^{(\varphi)}$ dependente da incerteza na energia $\sigma_H^{(\varphi)}$. Essa limitação consiste em que a escala de tempo característica para a evolução do valor médio de um observável que não é uma constante do movimento deve ser necessariamente *maior* que $\hbar/2\sigma_H^{(\varphi)}$, o que significa, em particular, que quanto mais próximo for o estado $|\varphi\rangle$ de um autovetor de H , no sentido de corresponder a incertezas $\sigma_H^{(\varphi)}$ menores, mais lentamente evoluirão os valores médios dos observáveis que não são constantes de movimento. No caso em que $|\varphi\rangle$ é um autovetor de H , esses valores médios se tornam independentes do tempo.

O uso da relação (2.20) pode ser substituído, no tratamento acima, pelo uso da relação mais completa (2.21), em que o termo correspondente à covariância de g e H é preservado. O que se obtém desse modo é, usando a equação de movimento de Heisenberg,

$$\det(\Sigma_{gH}) = \sigma_g^{(\varphi)^2} \sigma_H^{(\varphi)^2} - \sigma_{gH}^{(\varphi)^2} \geq \frac{\hbar^2}{4} \left| \langle \varphi | \frac{dg}{dt} | \varphi \rangle \right|^2$$

de onde se obtém, introduzindo nessa última relação o tempo característico $\tau_g^{(\varphi)}$,

$$\sqrt{\sigma_H^{(\varphi)^2} - \frac{\sigma_{gH}^{(\varphi)^2}}{\sigma_g^{(\varphi)^2}}} \tau_g^{(\varphi)} \geq \frac{\hbar}{2}. \quad (2.35)$$

Essa relação mostra que o efeito de uma covariância não nula de g e H é *aumentar* o limite inferior para a escala de tempo $\tau_g^{(\varphi)}$, correspondente a um valor dado da dispersão quadrática em energia, $\sigma_H^{(\varphi)^2}$.

É importante observar que, ao contrário de relações de incerteza como a de Heisenberg ou a de Schrödinger, que provém diretamente da estrutura cinemática da teoria quântica, as relações de incerteza energia-tempo tem um conteúdo dinâmico. Em termos qualitativos, elas mostram que processos de evolução rápidos requerem necessariamente o envolvimento de autovetores de H com energias muito diferentes e, reciprocamente, que processos que ocorrem em sistemas cuja energia é restrita a intervalos estreitos são necessariamente demorados. Quantitativamente, um processo (como, por exemplo, uma transição radiativa) que se desenvolve num dado tempo τ (tempo médio de vida do estado que sofre a transição) envolve uma incerteza na energia que é limitada inferiormente por $\Gamma \simeq \hbar/2\tau$ (essa incerteza aparece sob forma da “largura natural de linha” no caso da transição radiativa); um processo muito bem definido em energia (como, por exemplo, uma transição radiativa de “linha fina”, de largura Γ) necessariamente envolve estados cujo tempo médio de vida é maior que $\tau \simeq \hbar/2\Gamma$.

2.4 A “redução do pacote de ondas”.

Os postulados interpretativos de Born, tratados na seção **2.2.3**, permitem a previsão (em termos estatísticos) de resultados de medições feitas num dado sistema quântico, com base nos diversos ingredientes da teoria (vetor de estado, propriedades espectrais dos observáveis de interesse, evolução dinâmica do sistema antes da medição). Eles deixam, no entanto, completamente aberta a questão de como descrever o estado do sistema quântico *depois* de uma medição ter sido efetivamente completada, produzindo um resultado determinado.

Essa questão tem um conteúdo inequivocamente *dinâmico*, pois se refere aos efeitos da interação entre o sistema quântico observado e o dispositivo de medida, que é por certo indispensável para que o dispositivo de medida possa efetivamente funcionar como tal, e que pode alterar o estado do sistema submetido ao processo de medição. Um tratamento nesses termos envolve, no entanto, problemas sérios de diversas ordens. Por um lado, a dinâmica do dispositivo de medida e de sua interação com o sistema observado devem ser tratadas, por razões de consistência, também em termos da teoria quântica; então, não só essa descrição tende a ser proibitivamente complicada, devido à natureza macroscópica do dispositivo de medida, mas se defronta também, por outro lado, com questões graves e não inteiramente resolvidas, ligadas em última análise à peculiaridades da cinemática quântica. Para perceber a natureza dessas questões basta observar que os estados finais do sistema de medida são comumente entendidos como de certa forma “imunes ao princípio de superposição”, no sentido de que uma combinação linear de estados finais, tipo “posições do ponteiro”, do dispositivo de medida não são por sua vez estados possíveis, ou pelo menos não possíveis no mesmo sentido em que o são cada uma das “posições do ponteiro” tomada individualmente. Isso é freqüentemente expresso dizendo que o dispositivo de medida deve ter “características clássicas”. Razões de consistência exigem que a emergência de tais características clássicas seja dedutível da teoria quântica, o que novamente remete a um problema que não pode ser resolvido em geral.

Um tratamento padrão dessas questões consiste na introdução de um postulado adicional que, numa de suas versões correntes, efetivamente “prescreve” o estado do sistema quântico após a conclusão do processo de medida. Se o sistema que se encontra num estado descrito pelo vetor de estado $|\varphi\rangle$ é submetido a uma medição do observável g , cujos autovetores ortonormais são $\{|\phi_{nk}\rangle\}$ (supondo, por simplicidade, que o espectro seja puramente discreto), associados a autovalores γ_n e sendo k um índice para dar conta da eventual degenerescência dos autovalores γ_n , o efeito do processo de medição sobre o estado do sistema medido, após a obtenção de um particular autovalor γ_n , será

$$|\varphi\rangle = \sum_{nk} c_{nk} |\phi_{nk}\rangle \xrightarrow{\gamma_n} \frac{1}{\sqrt{\sum_k |c_{nk}|^2}} \sum_k c_{nk} |\phi_{nk}\rangle = \frac{1}{\sqrt{\langle\varphi|P_n|\varphi\rangle}} P_n |\varphi\rangle$$

com

$$P_n \equiv \sum_k |\phi_{nk}\rangle \langle \phi_{nk}|$$

isto é, $|\varphi\rangle$ é *reduzido* à sua projeção normalizada sobre o sub-espço dos autovetores degenerados de g associados ao autovalor obtido. No caso em que o autovalor γ_n não é degenerado, o estado do sistema após a medida será então dado simplesmente por $|\phi_n\rangle$. Este “postulado”, que a rigor deve encontrar apoio na dinâmica da interação do sistema medido com o aparato de medida, é freqüentemente designado como “redução do pacote de ondas”.

A prescrição contida nesse postulado tem como alegada virtude garantir a reprodutibilidade dos resultados de medições quânticas, no sentido de que a repetição imediata de uma medição que produziu um determinado resultado com certeza produzirá novamente o mesmo resultado. No entanto, é preciso ter em conta o fato de que medições feitas sobre sistemas quânticos *em geral não são reprodutíveis nesse sentido* (v. Landau e Lifshitz, *op. cit.* §7), de modo que a prescrição só pode ser tomada como válida em situações particulares que é importante identificar mais precisamente.

O conjunto formado pelo sistema submetido à medição mais o dispositivo de medida pode ser pensado como um sistema maior, composto de dois subsistemas, cada um dos quais tem seu próprio espaço de fases. Se $\{|\phi_n\rangle\}$ é a base do espaço de fases do sistema a ser medido formada pelos autovetores do observável que se pretende medir, g (ignorando agora, por simplicidade, questões ligadas à degenerescência), e se $\{|\Phi_n\rangle\}$ é uma base conveniente no espaço de fases do dispositivo de medida, uma base do *espaço produto* que desempenha o papel de espaço de fases para o sistema composto é formada pelos produtos distintos envolvendo um vetor de cada base, $\{|\phi_r\rangle \otimes |\Phi_s\rangle\} \equiv \{|\phi_r\Phi_s\rangle\}$. No estágio inicial da medida, esse sistema composto se encontra num estado que pode ser escrito como $|\varphi\rangle \otimes |\Phi_0\rangle \equiv \sum_n c_n |\phi_n\Phi_0\rangle$, o que significa que o sistema a ser medido e o dispositivo de medida devidamente “inicializado” tem como vetores de estado, respectivamente, $|\varphi\rangle \equiv \sum_n c_n |\phi_n\rangle$ e $|\Phi_0\rangle$. Para que o dispositivo de medida funcione como tal, é *suficiente* que a dinâmica da interação entre os dois subsistemas de que é formado o sistema composto faça com que ele evolua para um estado final do tipo $\sum_n c_n |\chi_n\rangle \otimes |\Phi_n\rangle$, em que os coeficientes c_n são *os mesmos* que aparecem na expansão do estado inicial a ser medido e os estados $|\chi_n\rangle$ são estados normalizados *quaisquer* do sistema que está sendo medido, cuja natureza pode depender do funcionamento do dispositivo de medida utilizado⁶. De fato, de acordo com os postulados interpretativos de Born, a probabilidade de que o *dispositivo de medida* se encontre então no estado $|\Phi_n\rangle$ será dada por $|c_n|^2$, de modo que a sua observação corresponde efetivamente à medida do autovalor de g associado a $|\phi_n\rangle$ no sistema submetido à medição. O estado final correspondente deste sistema será no entanto, em geral, um estado $|\chi_n\rangle$ determinado pelo funcionamento do dispositivo de medida, que poderá em particular ser até mesmo ortogonal ao estado inicial $|\varphi\rangle$. Neste caso, o estado inicial para uma nova medição de g será $|\chi_n\rangle \otimes |\Phi_0\rangle$ e o seu resultado será em geral diferente do obtido na primeira medição. A “redução” se dará da forma prescrita pela forma padrão do “postulado” apenas nos casos particulares em que os estados $|\chi_n\rangle$ forem na realidade idênticos aos estados $|\phi_n\rangle$.

Um caso efetivamente encontrado na literatura experimental recente (M. Brune et al, *Phys. Rev. Letters* **77**, 4887 (1996)) ilustra concretamente a ocorrência de tais medições *não seqüên-*

⁶O processo descrito é na realidade uma *hipótese dinâmica* (no sentido que *presume* efeitos da evolução temporal do sistema composto constituído pelo sistema submetido à medição e pelo dispositivo de medida) cuja consistência com a dinâmica quântica foi demonstrada por von Neumann (*op. cit.*, seção VI-3).

cialmente reprodutíveis. Trata-se aqui, de fato, essencialmente de medir a energia de um átomo cujo vetor de estado é uma combinação linear de dois estados estacionários de energias diferentes, $|\varphi\rangle = c_1|e_1\rangle + c_2|e_2\rangle$, com $e_1 < e_2$. O dispositivo de medida é formado por um par de detetores de ionização, que consiste basicamente de duas regiões I_1 e I_2 em que há campos elétricos \vec{E}_1 e \vec{E}_2 respectivamente. A intensidade do campo \vec{E}_1 é suficiente para ionizar o estado $|e_1\rangle$ e portanto também o estado *menos ligado* $|e_2\rangle$, ao passo que \vec{E}_2 é suficiente para ionizar o estado $|e_2\rangle$ mas não o estado *mais ligado* $|e_1\rangle$. O processo de medida consiste em fazer com que o átomo atravesse as regiões I_2 e I_1 *nessa ordem*, detectando o electron resultante da ionização onde ela ocorra. Desse modo, a detecção de um electron em I_2 , que ocorre com probabilidade $|c_2|^2$, identifica o estado *anterior* do átomo como $|e_2\rangle$, ao passo que a detecção de um electron em I_1 (cuja probabilidade é $|c_1|^2$) o identifica como $|e_1\rangle$. *Em qualquer dos dois casos* o átomo é ionizado durante o processo de medição, de modo que seu estado *posterior* à medição é certamente ortogonal à combinação linear inicial. Uma reconfirmação do resultado da primeira medição está portanto excluída neste tipo de procedimento, ao qual portanto a descrição padrão da redução do estado *não* se aplica.

Finalmente, convém notar explicitamente que, apesar de o “postulado da redução” não ser válido *em geral* na sua forma padrão, existem situações numerosas e importantes em que sua utilização nessa forma é válida e conveniente. A conveniência decorre especialmente de que, quando a aplicação do postulado é válida, o processo de medição pode ser visto como um mecanismo eficaz para a *preparação de estados quânticos*, pois o resultado obtido na medição determina, nesse caso, o estado subsequente do sistema que foi submetido à medição. A mesma referência citada acima, na qual são usados os detetores de ionização aos quais o postulado claramente não se aplica, faz um uso essencial desse postulado numa situação algo diferente e precisamente com esse intuito. De uma forma esquemática, nesta outra situação o sistema submetido à medição é por sua vez constituído de dois subsistemas que interagiram entre si durante um certo tempo e desse modo se tornaram correlacionados um ao outro. No momento da medição, no entanto, já não há interação entre os subsistemas e um deles é submetido a um processo de medida que, em particular, é não sequencialmente reprodutível no sentido discutido acima. O resultado desta medida, no entanto, pode ser considerado, através do postulado de redução (K. Fonseca Romero et al. *Phys. Rev.* **A58**, 3205 (1998)), como uma *preparação do estado do segundo subsistema*, devido às correlações estabelecidas durante o tempo de interação entre eles.

2.4.1 Medidas reprodutíveis, o “efeito Zeno” e a redução do pacote de ondas.

Medidas reprodutíveis são medidas que correlacionam estados selecionados $\{|\phi_n\rangle\}$ do sistema medido com estados respectivamente correspondentes $\{|\Phi_n\rangle\}$ do medidor *sem alterar os estados selecionados do sistema medido*, isto é, medidas em que a relação geral

$$|\varphi\rangle = \sum_n c_n |\phi_n\rangle \otimes |\Phi_0\rangle \longrightarrow \sum_n c_n |\chi_n\rangle \otimes |\Phi_n\rangle$$

assume na realidade a forma particular

$$|\varphi\rangle = \sum_n c_n |\phi_n\rangle \otimes |\Phi_0\rangle \longrightarrow \sum_n c_n |\phi_n\rangle \otimes |\Phi_n\rangle, \quad (2.36)$$

onde os estados finais $|\chi_n\rangle$ permanecem idênticos aos $|\phi_n\rangle$ iniciais. Como será visto pouco mais adiante, até mesmo nesses casos o processo de medida pode interferir de forma importante na evolução posterior do sistema medido. Independentemente das dificuldades que possa haver em realizar experimentalmente, e com precisão suficiente, processos de medida desse tipo, interações entre subsistemas que se comportem de modo a produzir resultados concordantes com a relação (2.36) podem ser escritas formalmente sem dificuldade. Um exemplo simples nesse sentido envolve um sistema de dois níveis como “sistema medido”, o qual, quando considerado isoladamente, é descrito pelo hamiltoniano

$$H_0 = |+\rangle \frac{\epsilon}{2} \langle +| + |-\rangle \frac{-\epsilon}{2} \langle -|$$

e um “medidor” que consiste de um sistema de *três* níveis descrito, também quando isolado, pelo hamiltoniano

$$H_m = \sum_{j=-1}^{+1} |\Phi_j\rangle j \delta \langle \Phi_j|.$$

Os autovalores associados aos autovetores $|\pm\rangle$ de H_0 são portanto $\pm\epsilon/2$, enquanto os autovalores associados aos três autovetores $|\Phi_j\rangle$, $j = -1, 0, +1$ de H_m são respectivamente $-\delta$, 0 e $+\delta$. Esses dois sistemas podem ser acoplados através de um hamiltoniano adicional H' de modo a efetuar uma medida reprodutível, por exemplo, dos estados

$$|\phi_{\pm}\rangle \equiv \frac{|+\rangle \pm |-\rangle}{\sqrt{2}}.$$

Esses estados estão sendo escolhidos por constituírem uma particular base de estados não estacionários do sistema descrito por H_0 . O operador de evolução correspondente a esse hamiltoniano, representado em termos da base de seus autovetores normalizados $|\pm\rangle$, é dado por

$$U(t_f, t_i) = e^{\frac{i}{\hbar} H_0(t_f - t_i)} = |-\rangle e^{\frac{i}{2\hbar} \epsilon(t_f - t_i)} \langle -| + |+\rangle e^{-\frac{i}{2\hbar} \epsilon(t_f - t_i)} \langle +|$$

de modo que a evolução temporal dos estados $|\phi_{\pm}\rangle$, determinada por H_0 , a partir do tempo inicial t_i e até o tempo final t_f , leva a

$$U(t_f, t_i)|\phi_{\pm}\rangle = \frac{1}{\sqrt{2}} \left(e^{-\frac{i}{2\hbar} \epsilon(t_f - t_i)} |+\rangle \pm e^{\frac{i}{2\hbar} \epsilon(t_f - t_i)} |-\rangle \right).$$

Sempre que conveniente, esse resultado pode ser facilmente re-expresso em termos dos estados selecionados $|\phi_{\pm}\rangle$. A evolução temporal de um estado qualquer do sistema de três níveis, determinada pelo hamiltoniano H_m , pode, evidentemente, ser obtida de forma inteiramente análoga a essa.

A idéia da medida reprodutível é a de que a interação entre os dois subsistemas transforme um estado que, antes da medida, é dado por

$$|\varphi\rangle \otimes |\Phi_0\rangle = (c_+|\phi_+\rangle + c_-|\phi_-\rangle) \otimes |\Phi_0\rangle \quad (2.37)$$

no estado dado, a menos de uma fase global, por

$$c_+|\phi_+\rangle \otimes |\Phi_{+1}\rangle + c_-|\phi_-\rangle \otimes |\Phi_{-1}\rangle. \quad (2.38)$$

Desse modo, o estado $|\Phi_0\rangle$ funciona como um estado “neutro” do medidor, que se transforma em $|\Phi_{+1}\rangle$ ou em $|\Phi_{-1}\rangle$ conforme o sistema medido esteja no estado $|\phi_+\rangle$ ou $|\phi_-\rangle$, respectivamente, sem que *as componentes selecionadas* do estado do sistema de dois níveis sofram qualquer alteração. Não é difícil ver que uma interação entre os dois subsistemas que tem esse tipo de efeito é a que é dada por

$$\begin{aligned} H'(t) &= \frac{\pi}{2} \hbar \delta(t - t_0) (|\Phi_{+1}\rangle \Pi_+ \langle \Phi_0| + |\Phi_{-1}\rangle \Pi_- \langle \Phi_0| + \text{hermiteano conjugado}) \\ &\equiv \frac{\pi}{2} \hbar \delta(t - t_0) \Gamma \end{aligned}$$

onde Π_{\pm} são os operadores de projeção $\Pi_{\pm} = |\phi_{\pm}\rangle \langle \phi_{\pm}|$. Essa interação é uma interação dependente do tempo, *impulsiva*, cuja ação se limita ao instante $t = t_0$ (o “instante em que é feita a medida”). De fato, a evolução do sistema composto dos dois subsistemas sob a ação de H' se restringe ao intervalo de tempo infinitesimal $t_0 - \eta \rightarrow t_0 + \eta$, com $0 < \eta \rightarrow 0$, sendo dada pelo operador unitário

$$U'(t_0 - \eta, t_0 + \eta) = e^{-i\frac{\pi}{2} \int_{t_0 - \eta}^{t_0 + \eta} dt' \delta(t' - t_0) \Gamma} = e^{-i\frac{\pi}{2} \Gamma}.$$

A exponencial envolvendo o operador Γ pode ser calculada notando que

$$\Gamma^2 = |\Phi_0\rangle \hat{1}_0 \langle \Phi_0| + |\Phi_{+1}\rangle \Pi_+ \langle \Phi_{+1}| + |\Phi_{-1}\rangle \Pi_- \langle \Phi_{-1}| \quad \text{e} \quad \Gamma^3 = \Gamma \quad (2.39)$$

de modo que, colecionando potências pares e ímpares de Γ na expansão da exponencial, resulta

$$U'(t_0 - \eta, t_0 + \eta) = (\hat{1} - \Gamma^2) + \Gamma^2 \cos \frac{\pi}{2} - i\Gamma \sin \frac{\pi}{2} = (\hat{1} - \Gamma^2) - i\Gamma.$$

A unitariedade desse resultado pode ser verificada explicitamente usando também as propriedades (2.39)⁷. Portanto, se o estado do sistema conjunto em $t_0 - \eta$ é dado por (2.37), ele será, em $t_0 + \eta$,

$$U'(t_0 - \eta, t_0 + \eta)|\varphi\rangle \otimes |\Phi_0\rangle = -i(c_+|\phi_+\rangle \otimes |\Phi_{+1}\rangle + c_-|\phi_-\rangle \otimes |\Phi_{-1}\rangle)$$

que corresponde às propriedades desejadas de “não demolição”.

Um exercício relativamente simples e interessante consiste em calcular a evolução temporal do sistema conjunto, descrito pelo hamiltoniano completo $H_0 + H_m + H'(t)$, a partir do estado inicial $|\Psi_{0,m}(t=0)\rangle = |\phi_+\rangle \otimes |\Phi_0\rangle$ até um tempo $T > t_0$. A evolução temporal de $t=0$ até $t=t_0 - \eta$ é dada por $H_0 + H_m$ apenas, pois nesse intervalo de tempo $H'(t) = 0$. Um cálculo simples mostra que

$$\begin{aligned} \lim_{\eta \rightarrow 0} |\Psi_{0,m}(t_0 - \eta)\rangle &= e^{-\frac{i}{\hbar}(H_0 + H_m)t_0} |\phi_+\rangle \otimes |\Phi_0\rangle \\ &= \left(\cos \frac{\epsilon t_0}{2\hbar} |\phi_+\rangle - i \sin \frac{\epsilon t_0}{2\hbar} |\phi_-\rangle \right) \otimes |\Phi_0\rangle. \end{aligned}$$

Para calcular a evolução temporal desse estado de $t_0 - \eta$ a $t_0 + \eta$, os efeitos devidos a H_0 e a H_m podem ser ignorados, devido ao caráter impulsivo da interação $H'(t)$. De fato o operador de evolução completo, correspondente a esse intervalo de tempo, é

$$U(t_0 + \eta, t_0 - \eta) = \exp \left[-\frac{i}{\hbar}(H_0 + H_m) \times 2\eta - i\frac{\pi}{2}\Gamma \right]$$

⁷O valor particular $\pi\hbar/2$ escolhido para o coeficiente de $H'(t)$ tem o efeito de anular o termo proporcional a Γ^2 do operador de evolução. O não anulamento deste termo implica na existência de uma componente proporcional a $|\Phi_0\rangle$ no estado posterior à interação, o que pode ser interpretado em termos de um processo de medida cuja eficiência é menor que 100 %.

de modo que, no limite $\eta \rightarrow 0$, apenas o último termo do expoente sobrevive, e $U(t_0 + \eta, t_0 - \eta)$ se reduz a $U'(t_0 + \eta, t_0 - \eta)$. Desse modo o estado será, em $t = t_0 + \eta$,

$$\lim_{\eta \rightarrow 0} |\Psi_{0,m}(t_0 + \eta)\rangle = -i \left(\cos \frac{\epsilon t_0}{2\hbar} |\phi_+\rangle \otimes |\Phi_{+1}\rangle - i \sin \frac{\epsilon t_0}{2\hbar} |\phi_-\rangle \otimes |\Phi_{-1}\rangle \right).$$

A evolução subsequente, de $t_0 + \eta$ a $T > t_0$, é novamente determinada por $H_0 + H_m$ apenas, e leva o estado finalmente a

$$\begin{aligned} |\Psi_{0,m}(T)\rangle &= -i e^{-\frac{i}{\hbar}(H_0 + H_m)(T-t_0)} \left(\cos \frac{\epsilon t_0}{2\hbar} |\phi_+\rangle \otimes |\Phi_{+1}\rangle - i \sin \frac{\epsilon t_0}{2\hbar} |\phi_-\rangle \otimes |\Phi_{-1}\rangle \right) = \\ &= -i \cos \frac{\epsilon t_0}{2\hbar} \left(\cos \frac{\epsilon(T-t_0)}{2\hbar} |\phi_+\rangle - i \sin \frac{\epsilon(T-t_0)}{2\hbar} |\phi_-\rangle \right) \otimes |\Phi_{+1}\rangle e^{-i\delta(T-t_0)} - \\ &\quad - \sin \frac{\epsilon t_0}{2\hbar} \left(-i \sin \frac{\epsilon(T-t_0)}{2\hbar} |\phi_+\rangle + \cos \frac{\epsilon(T-t_0)}{2\hbar} |\phi_-\rangle \right) \otimes |\Phi_{-1}\rangle e^{i\delta(T-t_0)}. \end{aligned} \quad (2.40)$$

Esse resultado permite calcular a probabilidade $P_+(T)$ de que o sistema medido esteja no estado $|\phi_+\rangle$ no tempo T (para $T > t_0$, e independentemente do estado do “medidor”) como

$$P_+(T) = \langle \Psi_{0,m}(T) | (|\phi_+\rangle \langle \phi_+|) \otimes \hat{1}_m | \Psi_{0,m}(T) \rangle = \quad (2.41)$$

$$= \cos^2 \frac{\epsilon t_0}{2\hbar} \cos^2 \frac{\epsilon(T-t_0)}{2\hbar} + \sin^2 \frac{\epsilon t_0}{2\hbar} \sin^2 \frac{\epsilon(T-t_0)}{2\hbar} = \frac{1}{2} \left(1 + \cos \frac{\epsilon t_0}{\hbar} \cos \frac{\epsilon(T-t_0)}{\hbar} \right).$$

Caso a interação representada por $H'(t)$ não tivesse ocorrido, o estado do sistema composto no tempo T (e para qualquer $T > 0$) seria dado simplesmente por

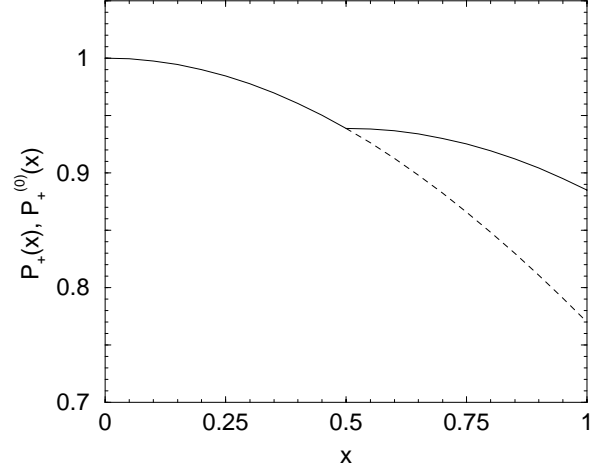
$$|\Psi_{0,m}^{(0)}\rangle = \left(\cos \frac{\epsilon T}{2\hbar} |\phi_+\rangle - i \sin \frac{\epsilon T}{2\hbar} |\phi_-\rangle \right) \otimes |\Phi_0\rangle$$

e a probabilidade correspondente a $P_+(T)$ seria

$$P_+^{(0)}(T) = \cos^2 \frac{\epsilon T}{2\hbar} = \frac{1}{2} \left(1 + \cos \frac{\epsilon T}{\hbar} \right). \quad (2.42)$$

A comparação das expressões (2.41) e (2.42) mostra imediatamente que a intervenção da interação $H'(t)$ associada à medida reprodutível no instante $t = t_0$ não é inócua no que se refere à evolução temporal ulterior do sistema quântico composto. A diferença resultante dos efeitos da interação pode ser vista na figura 2.1, onde as duas probabilidades, $P_+(T)$ e $P_+^{(0)}(T)$ são mostradas como função de $T > t_0$ no caso em que $t_0 = \hbar/2\epsilon$, e portanto $\epsilon t_0/\hbar = 0.5$, e consiste numa clara *inibição* da evolução temporal com relação à evolução livre $P_+^{(0)}(T)$ da probabilidade.

Figura 2.1: $P_+(x)$ (linha cheia, $x > 0.5$) e $P_+^{(0)}(x)$ (linha cheia, $x < 0.5$ e tracejada, $x > 0.5$), com $x = \epsilon T/\hbar$. A evolução com a intervenção de $H'(t)$ em $x = 0.5$ corresponde à linha cheia, de $x = 0$ a $x = 1$.



O mecanismo responsável pelo efeito da interação associada à medida reprodutível, $H'(t)$, sobre a evolução ulterior da probabilidade $P_+(T)$ pode ser identificado mais explicitamente notando que a taxa de variação dessa probabilidade com o tempo é dada em geral por

$$\frac{dP_+}{dt} = \frac{i}{\hbar} \langle \Psi_{0,m} | [H(t), |\phi_+\rangle \langle \phi_+| \otimes \hat{1}_m] | \Psi_{0,m} \rangle$$

com $H(t) = H_0 + H_m + H'(t)$. Para qualquer $t \neq t_0$ o termo impulsivo de interação é nulo e a expressão se reduz a

$$\begin{aligned} \frac{dP_+}{dt} &= \frac{i\epsilon}{2\hbar} \langle \Psi_{0,m} | (|\phi_-\rangle \langle \phi_+| - |\phi_+\rangle \langle \phi_-|) \otimes \hat{1}_m | \Psi_{0,m} \rangle = \\ &= -\frac{\epsilon}{2\hbar} \text{Im} \langle \Psi_{0,m} | (|\phi_-\rangle \langle \phi_+|) \otimes \hat{1}_m | \Psi_{0,m} \rangle, \quad t \neq t_0. \end{aligned} \quad (2.43)$$

Como o vetor de estado $|\Psi_{0,m}\rangle$ envolve os dois subsistemas, isto é, o sistema “medido” de dois, e o sistema “medidor” de três níveis, essa expressão precisa ser analisada com algum cuidado. Quando esse vetor de estado é do tipo (2.37), o que quer dizer um estado produto de uma combinação linear de $|\phi_+\rangle$ e $|\phi_-\rangle$ com um estado normalizado qualquer do sistema de três níveis, a taxa de variação da probabilidade se reduz simplesmente a

$$\text{estado (2.37)} \quad \longrightarrow \quad \frac{dP_+}{dt} = -\frac{\epsilon}{2\hbar} \text{Im} c_-^* c_+.$$

Essa expressão se anula para $t = 0$ (pois então $c_- = 0$) e cresce para $0 < t < t_0$, à medida que c_- cresce devido à evolução temporal do estado inicial. É importante aqui perceber que a taxa de variação da probabilidade envolve um “termo cruzado” dos coeficientes de

expansão do estado do sistema de dois níveis na base $\{|\phi_{\pm}\rangle\}$. Por outro lado, para um vetor de estado do tipo (2.38), no qual as componentes $|\phi_{+}\rangle$ e $|\phi_{-}\rangle$ se correlacionam com estados do sistema de três níveis que são *ortogonais entre si*, os termos cruzados necessários para levar a um valor não nulo da taxa de variação da probabilidade são excluídos devido ao fato de que $\langle\Phi_{\pm 1}|\hat{1}_m|\Phi_{\mp 1}\rangle = 0$. Desse modo

$$\text{estado (2.38)} \quad \longrightarrow \quad \frac{dP_{+}}{dt} = 0.$$

Como a interação $H'(t)$ transforma um estado tipo (2.37) em um estado tipo (2.38), pode-se ver que ela na realidade leva ao anulamento da taxa de variação da probabilidade através da destruição dos termos cruzados necessários para que seu valor seja diferente de zero. O mecanismo responsável por esse anulamento é justamente o estabelecimento de correlações entre as componentes $|\phi_{+}\rangle$ e $|\phi_{-}\rangle$ do estado (2.37) com estados do sistema medidor que são ortogonais entre si. O processo sofrido pelo sistema de dois níveis através da interação $H'(t)$ é chamado *decoerência*, e pode ser portanto apontado como o ingrediente básico que leva ao comportamento de $P_{+}(T)$ mostrado na figura 2.1. A ocorrência desse processo mostra que a interação associada à medida reproduzível na realidade produz estragos apreciáveis no estado do sistema medido.

Cabe observar ainda que, para $T > t_0$, a evolução unitária devida a H_0 volta a produzir componentes $|\phi_{+}\rangle$ e $|\phi_{-}\rangle$ para *cada uma das componentes* do estado (2.38), como pode ser visto explicitamente na expressão (2.40), levando novamente a valores diferentes de zero da taxa de variação da probabilidade.

A inibição da evolução livre $P_{+}^{(0)}(T)$ da probabilidade pela interação H' , como mostrada na figura 2.1, pode ser vista, por assim dizer, como uma forma embrionária do chamado “efeito Zeno quântico”, apontado e assim denominado por Misra e Sudarshan em 1977⁸. Na sua versão original, esse efeito foi apresentado como uma consequência do postulado de redução do pacote de ondas para sistemas quânticos cuja evolução temporal para tempos curtos tem um comportamento semelhante ao obtido acima para o estado $|\phi_{+}\rangle$, isto é, a probabilidade $P(t)$ de encontrar o sistema no estado inicial após um intervalo de tempo t tem derivada nula em $t = 0$. No caso do modelo tratado acima $P(t)$ depende quadraticamente de t para tempos curtos δt , isto é $P(t) \propto 1 - (\delta t/\tau)^2$ (v. Eq. (2.42)). Nessas condições, uma sucessão rápida de medidas capazes de decidir sobre a permanência ou não do sistema no estado inicial inibe a evolução temporal, e a impede completamente no limite de “observação contínua”. De fato, utilizando o postulado da redução, o estado do sistema após uma medida cujo resultado indica que o estado é $|\phi_{+}\rangle$ será novamente $|\phi_{+}\rangle$. Como a probabilidade de obter esse resultado um tempo δt após

⁸B. Misra e E. C. G. Sudarshan, *J. Math. Phys.* **18**, 756 (1977). V. também C. B. Chiu, B. Misra e E. C. G. Sudarshan, *Phys. Rev.* **D16**, 520 (1977) e *Phys. Lett.* **B117**, 34 (1982).

a preparação do estado é $1 - (\delta t/\tau)^2$, a probabilidade de obter o mesmo resultado ao longo de uma série de N medidas separadas por intervalos de tempo δt será

$$P_+^{(N)} = \left[1 - \left(\frac{\delta t}{\tau} \right)^2 \right]^N = 1 - N \left(\frac{\delta t}{\tau} \right)^2 + \frac{N(N-1)}{2} \left(\frac{\delta t}{\tau} \right)^4 - \dots \quad (2.44)$$

Se a sequência de N medidas é efetuada em um tempo total T , $\delta t = T/N$ e então

$$P_+^{(N)} = 1 - N \frac{T^2}{N^2 \tau^2} + \dots = 1 - \frac{1}{N} \left(\frac{T}{\tau} \right)^2 + \dots$$

o que, mantendo T fixo e tomando o limite $N \rightarrow \infty$ (que corresponde à “observação contínua”), leva a

$$\lim_{N \rightarrow \infty} P_+^{(N)} = 1,$$

isto é, o estado do sistema permanece igual ao estado inicial com probabilidade 1. É claro que o postulado da redução desempenha um papel central nessa formulação, pois é usado de forma essencial para escrever a relação (2.44).

A reavaliação do efeito Zeno como um processo puramente dinâmico, independente da prescrição dada pelo postulado da redução, seguiu-se à realização, por Itano et al., de uma experiência ao longo de linhas propostas pouco antes por Cook⁹. O cálculo que leva ao resultado da figura 2.1 é evidentemente um cálculo puramente dinâmico, que prescinde, em particular, da informação registrada pelo “aparelho de medida”, dado que a probabilidade $P_+(T)$ é calculada de forma *inclusiva*, somando sobre todos os estados possíveis do sistema descrito por H_m . Com algum trabalho algébrico adicional, é possível calcular a probabilidade inclusiva $P_+(N \delta t)$ correspondente a uma série de N interações H' agindo em tempos $n \delta t$, $n = 1, 2, \dots, N$ e envolvendo sequencialmente N sistemas *independentes* de três níveis. O resultado desse cálculo é

$$P_+(N \delta t) = \frac{1}{2} \left(1 + \cos^N \frac{\epsilon \delta t}{\hbar} \right) \quad (2.45)$$

que se reduz também a 1 nas mesmas condições que (2.44).

É preciso comentar ainda o fato de que as probabilidades (2.44) e (2.45) levam a resultados diferentes no caso de uma sucessão *finita* de N medidas ou interações, separadas por intervalos

⁹W. H. Itano, D. J. Heinzen, J. J. Bollinger e D. J. Wineland, *Phys. Rev.* **A41**, 2295 (1990); R. J. Cook, *Phys. Scr.* **T21**, 49 (1988). A respeito da elaboração desse outro ponto de vista, v. S. Pascazio e M. Namiki, *Phys. Rev.* **A50**, 4582 (1994) e as referências aí citadas, especialmente L. E. Ballentine, *Phys. Rev.* **A43**, 5165 (1991).

de tempo curtos mas também finitos δt . De fato, utilizando a evolução temporal definida para o modelo, a equação (2.44) aparece sob a forma

$$P_+^{(N)} = \cos^{2N} \frac{\epsilon \delta t}{2\hbar} \quad (2.46)$$

que difere de (2.45) para $N \geq 2$. Essa diferença se deve ao caráter inclusivo da probabilidade dada por (2.45), em que são computadas também as probabilidades associadas às componentes $|\phi_+\rangle$ resultantes da evolução unitária da componente $|\phi_-\rangle$ presente imediatamente após a interação precedente, e que são excluídas pelo argumento que envolve o postulado da redução. Isso faz com que se tenha $P_+(N \delta t) > P_+^{(N)}$ para $N \geq 2$ e, quando $N \rightarrow \infty$ (com δt fixo), $P_+(N \delta t) \rightarrow 1/2$ enquanto $P_+^{(N)} \rightarrow 0$. A partir de (2.45) e de (2.46) pode-se verificar também que o efeito de uma medida ou interação adicional sobre as probabilidades, após uma sequência de N medidas ou interações anteriores, é o de alterar essas probabilidades de acordo com as relações

$$\frac{P_+[(N+1)\delta t]}{P_+(N\delta t)} = \cos^2 \frac{\epsilon \delta t}{2\hbar} + \frac{1 - P_+(N\delta t)}{P_+(N\delta t)} \sin^2 \frac{\epsilon \delta t}{2\hbar} > \frac{P_+^{(N+1)}}{P_+^{(N)}} = \cos^2 \frac{\epsilon \delta t}{2\hbar}.$$

Problemas.

2.1 - a) Mostre que os coeficientes c_n que aparecem na Eq. (2.2) são dados por $c_n = (\phi_n, \varphi)$ quando o conjunto de vetores $\{\phi_n\}$, $n = 1, 2, \dots$, formam uma base ortonormal. Como deve ser escrita a expansão de φ em termos de uma base $\{\hat{\phi}_n\}$, ortogonal mas cujos vetores não são normalizados, isto é, tal que $(\hat{\phi}_n, \hat{\phi}_m) = \nu_n \delta_{nm}$ com os ν_n não nulos, dependentes de n e diferentes de 1?

b) Mostre que os elementos g_{nm} da matriz que representa um operador hermiteano g na base ortonormal $\{\phi_n\}$ são $g_{nm} = (\phi_n, g\phi_m)$, que $g_{nm} = g_{mn}^*$, e que os coeficientes da expansão da imagem $g|\varphi\rangle$ de um vetor qualquer $|\varphi\rangle$ nessa base são dados por $\langle \phi_m | g|\varphi \rangle = \sum_n g_{mn} c_n$, com $c_n = \langle \phi_n | \varphi \rangle$. Como se modifica esta última relação no caso em que se use a base ortogonal de vetores não normalizados $\{\hat{\phi}_n\}$?

c) Mostre que a matriz que representa o produto gh de dois operadores g e h numa base ortonormal é o produto (matricial) das matrizes que representam cada um deles na mesma base, isto é

$$(gh)_{nm} = \sum_p g_{np} h_{pm}.$$

d) Verifique a expressão (2.5) usando as expansões em série das exponenciais e a hermiticidade de H .

e) Usando as expansões em série de potências das exponenciais em (2.4) mostre que $g(t)$ pode ser escrito em termos da série de potências de t envolvendo múltiplos comutadores

$$g(t) = g(0) + \frac{i}{\hbar}[H, g(0)]t + \frac{1}{2!} \left(\frac{i}{\hbar}\right)^2 [H, [H, g(0)]]t^2 + \dots$$

Calcule também o termo seguinte dessa série.

f) Verifique a equação (2.8) mostrando que $e^{\frac{i}{\hbar}Ht}\phi_n = e^{\frac{i}{\hbar}E_nt}\phi_n$, dado que $H\phi_n = E_n\phi_n$. Use para isso a expansão da exponencial em série de potências. Esse resultado depende da norma de ϕ_n ? Qual é a relação entre a norma de ϕ_n e a de $e^{\frac{i}{\hbar}Ht}\phi_n$?

2.2 - Usando a definição de norma (2.9) mostre que

a) dados dois vetores $|\varphi\rangle$ e $|\psi\rangle$ quaisquer, vale a *igualdade do paralelogramo*

$$2\|\varphi\|^2 + 2\|\psi\|^2 = \|\varphi + \psi\|^2 + \|\varphi - \psi\|^2$$

onde $\|\varphi \pm \psi\|$ denota a norma dos vetores $|\varphi\rangle \pm |\psi\rangle$.

b) dada uma família ortonormal de N vetores, $\{|\phi_n\rangle\}$, $n = 1, \dots, N$, e um vetor qualquer $|\varphi\rangle$ em um espaço de dimensão maior que N (possivelmente infinita), vale o *teorema de Pitágoras*

$$\|\varphi\|^2 = \sum_{n=1}^N |\langle\phi_n|\varphi\rangle|^2 + \|\varphi - \sum_{n=1}^N \phi_n \langle\phi_n|\varphi\rangle\|^2$$

onde o último termo denota o quadrado da norma do vetor $|\varphi\rangle - \sum_{n=1}^N |\phi_n\rangle \langle\phi_n|\varphi\rangle$. **Sugestão:** escreva o vetor $|\varphi\rangle$ como a soma de dois vetores ortogonais entre si usando o operador de projeção $P \equiv \sum_{n=1}^N |\phi_n\rangle \langle\phi_n|$ e seu complementar $Q \equiv \hat{1} - P$.

Use as propriedades da norma para obter, a partir do teorema de Pitágoras, a *desigualdade de Bessel*

$$\|\varphi\|^2 \leq \sum_{n=1}^N |\langle\phi_n|\varphi\rangle|^2.$$

Em que condições essa desigualdade vale como igualdade?

2.3 - Descreva os autovetores do operador de projeção $Q_{\phi_n} = \hat{1} - P_{\phi_n}$, com $P_{\phi_n} \equiv |\phi_n\rangle \langle\phi_n|$, $\langle\phi_n|\phi_n\rangle = 1$, associados a cada um dos autovalores possíveis, 1 e 0 bem como o grau de degenerescência de cada um dos autovalores. Compare o resultado com o que corresponde ao operador de projeção complementar P_{ϕ_n} **Sugestão:** use uma base ortonormal do espaço vetorial que contém $|\phi_n\rangle$ como um de seus vetores.

2.4 - Considere uma família de N vetores *não ortogonais* mas linearmente independentes $|\varphi_j\rangle$; então

a) mostre que a matriz $N \times N$ (não diagonal) de produtos escalares $s_{ij} \equiv \langle \varphi_i | \varphi_j \rangle$ é uma matriz hermiteana positiva definida, isto é, $\sum_{ij} c_i^* s_{ij} c_j \geq 0$ para qualquer conjunto de coeficientes $\{c_i\}$, $i = 1, \dots, N$, sendo que a igualdade a zero implica que todos os c_i são nulos (**sugestão:** use a independência linear dos vetores $|\varphi_j\rangle$ e as propriedades de positividade da norma);

b) considere o problema de autovalores

$$\sum_{j=1}^N s_{ij} f_j^{(n)} = \sigma_n f_i^{(n)} \quad (2.47)$$

em que $\{f_j^{(n)}\}$, $j = 1, \dots, N$ são autovetores da matriz $\{s_{ij}\}$ com autovalores σ_n ; mostre que os autovalores podem ser obtidos como as N raízes da *equação característica* $\det(s_{ij} - \lambda \delta_{ij}) = 0$; use a hermiticidade da matriz para mostrar que os autovalores são reais (a positividade da matriz implica que todos os autovalores são positivos);

c) supondo todas as raízes distintas (isto é, supondo que não haja raízes múltiplas), mostre que os autovetores satisfazem a condição de ortogonalidade

$$\sum_{j=1}^N f_j^{*(m)} f_j^{(n)} = 0 \quad \text{se } m \neq n;$$

d) supondo também que os autovetores sejam normalizados como $\sum_j |f_j^{(n)}|^2 = 1$, mostre que os N vetores

$$|\phi_n\rangle \equiv \sum_j \frac{1}{\sqrt{\sigma_n}} f_j^{(n)} |\varphi_j\rangle$$

são ortonormais, isto é $\langle \phi_m | \phi_n \rangle = \delta_{mn}$;

e) mostre finalmente que, se $|\psi\rangle = \sum_j c_j |\varphi_j\rangle$ é uma combinação linear qualquer dos vetores de partida e $P \equiv \sum_n |\phi_n\rangle \langle \phi_n|$, então P é um operador de projeção e $P|\psi\rangle = |\psi\rangle$.

Observação: no caso em que a equação característica tenha raízes múltiplas, a hermiticidade de $\{s_{ij}\}$ garante que a indeterminação do sistema de equações lineares (2.47) é sempre suficiente para que haja um número de autovetores $\{f_j^{(n,k)}\}$ independentes igual à multiplicidade μ_n da raiz múltipla correspondente σ_n (isto é, $k = 1, \dots, \mu_n$), os quais podem ser escolhidos de forma a serem ortogonais entre si. Essa escolha não é porem única, ao contrário do que ocorre no caso em que todas as raízes da equação característica sejam distintas. No entanto, vale a pena observar também que o procedimento usado acima para gerar uma família ortonormal da qual resulte um operador de projeção com a propriedade expressa em **e)** por sua vez não é o

único possível, e que procedimentos diferentes fornecem em geral famílias ortonormais também diferentes. A existência de uma infinidade de famílias ortonormais correspondendo ao mesmo projetor P resulta de que transformando qualquer família particular (por exemplo, a obtida em **d)** acima) por meio de uma matriz *unitária* qualquer de dimensão N resulta outra família que leva ao mesmo P .

f) Verifique a última afirmação da observação acima: se $\{u_{im}\}$, $i, m = 1, \dots, N$ é uma matriz unitária, isto é, $(u^{-1})_{nj} = (u^\dagger)_{nj} = u_{jn}^*$ e portanto $\sum_i u_{im} u_{in}^* = \delta_{mn}$, então os N vetores $|\tilde{\phi}_i\rangle \equiv \sum_m u_{im} |\phi_m\rangle$ são ortonormais (supondo que os $|\phi_n\rangle$ também sejam, tendo sido obtidos, por exemplo, como em **d)** acima) e

$$\sum_{i=1}^N |\tilde{\phi}_i\rangle \langle \tilde{\phi}_i| = \sum_{n=1}^N |\phi_n\rangle \langle \phi_n| \equiv P.$$

2.5 - a) Se $\{|\phi_n\rangle\}$, $n = 1, 2, \dots$ é uma base em um espaço de Hilbert e g é o operador linear, representado nessa base pela matriz $g_{mn} = \langle \phi_m | g | \phi_n \rangle$, mostre que o g pode ser escrito como

$$g = \sum_{mn} |\phi_m\rangle g_{mn} \langle \phi_n|.$$

b) Considere *outra* base ortonormal $\{|\chi_n\rangle\}$, $n = 1, 2, \dots$. Mostre que a matriz \mathbf{u} tal que $u_{mn} \equiv \langle \chi_m | \phi_n \rangle$ é unitária, isto é

$$\mathbf{u} \mathbf{u}^\dagger = \mathbf{u}^\dagger \mathbf{u} = \mathbf{1} \quad \text{ou seja} \quad \sum_l u_{ml} u_{nl}^* = \sum_l u_{lm}^* u_{ln} = \delta_{mn}.$$

c) O *traço* de um operador linear g , $\text{Tr}[g]$, pode ser definido como a soma dos elementos de matriz diagonais $g_{nn} = \langle \phi_n | g | \phi_n \rangle$, isto é

$$\text{Tr}[g] = \sum_n g_{nn}.$$

Mostre que $\text{Tr}[g]$ é independente da base ortonormal escolhida.

d) Mostre que o elemento de matriz $g_{mn} = \langle \phi_m | g | \phi_n \rangle$ pode ser obtido como

$$g_{mn} = \text{Tr} [g(U_{mn})^\dagger]$$

onde $U_{mn} \equiv |\phi_m\rangle \langle \phi_n|$.

2.6 - Um operador linear σ_1 agindo sobre um espaço vetorial de duas dimensões é representado numa certa base ortonormal $\{|\phi_1\rangle, |\phi_2\rangle\}$ pela matriz

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

a) Esse operador é hermiteano? Obtenha seus autovalores s_1 e s_2 e também os autovetores normalizados correspondentes nessa mesma base $\{|\phi_1\rangle, |\phi_2\rangle\}$.

b) Usando ainda a mesma base, escreva as matrizes que representam os operadores de projeção P_{s_i} , $i = 1, 2$ sobre cada um dos dois autovetores de σ_1 . Usando essas matrizes, verifique explicitamente as relações $P_{s_1}P_{s_2} = P_{s_2}P_{s_1} = \mathbf{0}$ e $P_{s_1} + P_{s_2} = \mathbf{1}$ onde $\mathbf{0}$ é a matriz nula e $\mathbf{1}$ é a matriz identidade 2×2 .

c) Mostre que a matriz correspondente ao operador $e^{i\alpha\sigma_1}$, onde α é um número real, pode ser escrita como

$$e^{i\alpha\sigma_1} = \mathbf{1} \cos \alpha + i\sigma_1 \sin \alpha.$$

Sugestão: Verifique que $\sigma_1^2 = \mathbf{1}$ e use esse fato na expansão da exponencial em série de potências.

2.7 - Seja o hamiltoniano de oscilador unidimensional

$$H = \frac{p^2}{2M} + \lambda q^k = H^\dagger, \quad \lambda > 0, \quad k \text{ par } \geq 2,$$

e sejam E_n e $|\phi_n\rangle$ seus autovalores e autovetores normalizados, isto é

$$H|\phi_n\rangle = E_n|\phi_n\rangle, \quad \langle\phi_n|\phi_n\rangle = 1.$$

a) Se g é uma variável dinâmica qualquer, mostre que

$$\langle\phi_n|[g, H]|\phi_n\rangle = 0$$

para qualquer autovetor $|\phi_n\rangle$ de H .

b) Calcule os comutadores $[q, H]$, $[p, H]$ e $[pq + qp, H]$ em termos de p e q , usando $[p, q] = \hbar\hat{1}/i$. Verifique em particular que o último deles pode ser expresso em termos de $K = p^2/2M$ e de $V = \lambda q^k$. Use então esses comutadores e o resultado **a)** para mostrar as seguintes relações, válidas para qualquer autovetor $|\phi_n\rangle$:

$$\textbf{i)} \quad \langle\phi_n|p|\phi_n\rangle = 0;$$

$$\textbf{ii)} \quad \langle\phi_n|q^{k-1}|\phi_n\rangle = 0;$$

$$\textbf{iii)} \quad \langle\phi_n|\frac{p^2}{2M}|\phi_n\rangle = \frac{k}{2}\langle\phi_n|\lambda q^k|\phi_n\rangle \quad (\text{teorema do virial});$$

$$\textbf{iv)} \quad \langle\phi_n|\frac{p^2}{2M}|\phi_n\rangle = \frac{kE_n}{k+2} \quad \text{e} \quad \langle\phi_n|\lambda q^k|\phi_n\rangle = \frac{2E_n}{k+2}.$$

c) No caso de um oscilador anarmônico cujo potencial pode ser expandido em série de potências $V(q) = \sum_k v_k q^k$, mostre que o teorema do virial aparece sob a forma mais geral

$$\langle \phi_n | \frac{p^2}{2M} | \phi_n \rangle = \frac{1}{2} \langle \phi_n | q \frac{\partial V(q)}{\partial q} | \phi_n \rangle.$$

2.8 - Usando os resultados do problema **2.7** no caso $k = 2$, calcule o valor do produto de incertezas $\sigma_p \sigma_q$ para o estado estacionário de energia $E_n = (n+1/2)\hbar\omega_0$ do oscilador harmônico. Verifique, em particular, que o estado de menor energia é um estado de incerteza mínima..

2.9 - Um “sistema de dois níveis” é descrito por um hamiltoniano cuja matriz, numa certa base ortonormal $\{|\phi_1\rangle, |\phi_2\rangle\}$, é

$$H = \begin{pmatrix} \epsilon & \gamma \\ \gamma & \epsilon \end{pmatrix}.$$

a) Determine as energias dos estados estacionários desse sistema e os vetores de estado normalizados que os representam, em termos da mesma base ortonormal.

b) Tratando a dinâmica do sistema na descrição de Schrödinger, determine a forma do vetor de estado num instante genérico $t > 0$ a partir da condição inicial de que em $t = 0$ o vetor de estado é $|\phi_1\rangle$.

c) Calcule a probabilidade de que o sistema esteja no mesmo estado $|\phi_1\rangle$ no tempo $t > 0$. Existem valores de t para os quais essa probabilidade se anula? Em caso positivo, qual o vetor de estado correspondente à descrição de Schrödinger nesses tempos?

d) Para resolver esse mesmo problema na descrição de Heisenberg, obtenha o operador $P_1(t)$ dessa descrição que em $t = 0$ é o operador de projeção $P_1(t = 0) \equiv |\phi_1\rangle\langle\phi_1|$, calcule o valor médio desse operador no vetor de estado independente do tempo $|\phi_1\rangle$ e compare o resultado com o obtido em c).

e) O operador $P_1(t)$ é ainda um operador de projeção para $t > 0$? Obtenha os autovetores desse operador e determine a sua relação com o resultado do item b).

2.10 - Mostre que o hamiltoniano não conecta dois autovetores de uma constante de movimento associados a autovalores diferentes, isto é: se g é uma constante de movimento e se $g|\phi_i\rangle = \gamma_i|\phi_i\rangle$, $i = 1, 2$ com $\gamma_1 \neq \gamma_2$, então $\langle\phi_1|H|\phi_2\rangle = 0$.

2.11 - O hamiltoniano de um sistema de três níveis é representado numa certa base ortonormal $\{|\phi_i\rangle\}$, $i = 1, 2, 3$ pela matriz

$$H = e_0 \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & \sqrt{3} \\ 0 & \sqrt{3} & 2 \end{pmatrix}$$

onde e_0 é uma constante real com dimensões de energia.

a) Verifique que o observável representado na mesma base pela matriz

$$g = \gamma_0 \begin{pmatrix} 4 & -2\sqrt{3} & 2 \\ -2\sqrt{3} & 5 & \sqrt{3} \\ 2 & \sqrt{3} & 7 \end{pmatrix}$$

onde γ_0 é também uma constante real com dimensões apropriadas, é uma *constante de movimento*, e portanto compatível com H .

b) Obtenha um conjunto de autovetores comuns às duas matrizes (sempre em termos da base ortonormal $\{|\phi_i\rangle\}$, $i = 1, 2, 3$) e os autovalores correspondentes, para cada uma delas. Esses dois observáveis formam um *conjunto completo de observáveis compatíveis* para o sistema de três níveis?

2.12 - a) Com referência à discussão do “efeito Zeno” quântico na seção 2.4.1, calcule a probabilidade inclusiva

$$P_+(2\delta t) = \lim_{\eta \rightarrow \infty} \langle \Psi_{0,m}(2\delta t + \eta) | (|\phi_+\rangle\langle\phi_-|) \otimes \hat{1}_m | \Psi_{0,m}(2\delta t + \eta) \rangle$$

onde $|\Psi_{0,m}(2\delta t + \eta)\rangle$ é o estado que evolui a partir do estado inicial

$$|\Psi_{0,m}(t=0)\rangle = |\phi_+\rangle \otimes |\Phi_0^{(1)}\rangle \otimes |\Phi_0^{(2)}\rangle$$

através do hamiltoniano $H = H_0 + H_m + H'$, com

$$H_0 = |+\rangle \frac{\epsilon}{2} \langle + | + |-\rangle \frac{-\epsilon}{2} \langle - |, \quad H_m = \sum_{n=1}^2 \sum_{j=-1}^{+1} |\Phi_j^{(n)}\rangle j \delta \langle \Phi_j^{(n)}|$$

e

$$H' = \frac{\pi\hbar}{2} \sum_{n=1}^2 \delta(t - n\delta t) \Gamma^{(n)}, \quad \Gamma^{(n)} = (|\Phi_{+1}^{(n)}\rangle \Pi_+ \langle \Phi_0^{(n)}| + |\Phi_{-1}^{(n)}\rangle \Pi_- \langle \Phi_0^{(n)}|) + \text{h. c. .}$$

Sugestão: Note que existem *dois* sistemas “medidores” independentes, distinguidos pelo índice superior $n = 1, 2$, que são *respectivamente* acoplados ao sistema de dois níveis nos tempos $n\delta t$. Nessas condições, a evolução temporal até o instante $t = 2\delta t - \eta$ é dada pela expressão (2.40), o que permite obter rapidamente o resultado desejado.

b) Calcule a probabilidade *exclusiva*

$$P_+^{(N=2)} \equiv \lim_{\eta \rightarrow \infty} \langle \Psi_{0,m}(2\delta t + \eta) | \left(\sum_{n=1}^2 |\Phi_{+1}^{(n)}\rangle \langle \Phi_{+1}^{(n)}| \right) \otimes \hat{1}_0 | \Psi_{0,m}(2\delta t + \eta) \rangle$$

que corresponde a encontrar *os sistemas medidores* nos estados $|\Phi_{+1}^{(n)}\rangle$ no instante $t = 2\delta t + \eta$. Compare essa probabilidade com o resultado correspondente obtido utilizando o postulado da redução como na seção 2.4.1.

Referências gerais.

O formalismo geral e a interpretação da mecânica quântica são tratados em praticamente todos os numerosos textos sobre o assunto, com variações eventuais de ênfase e sofisticação. Tratamentos que podem ser vistos como “padrão” se encontram, por exemplo, em

★ **A. Messiah**, *Quantum Mechanics*, John Wiley, 1961, Capítulo VII.

★ **C. Cohen-Tanoudji, B. Diu e F. Laloë**, *Mécanique Quantique*, Hermann, 1973, Capítulos II e III.

★ **J. J. Sakurai**, *Modern Quantum Mechanics*, Addison-Wesley Publ. Co., 1985, Capítulos 1 e 2.

Um texto padrão dedicado à formulação de Feynman é

★ **R. P. Feynman e A. R. Hibbs**, *Quantum Mechanics and Path Integrals*, McGraw-Hill Book Co., 1965.

O tratamento clássico (e muito legível), cuja influência pode ser facilmente percebida nos textos posteriores, se encontra no entanto em

★ **P. A. M. Dirac**, *The Principles of Quantum Mechanics*, 4th. Edition, Oxford U. Press, 1958, Capítulos I a IV.

Um tratamento compacto mas profundo, que inclui alguns pontos raramente abordados, como a não reprodutibilidade, em geral, de medições feitas sobre sistemas quânticos, encontra-se em

★ **L. Landau e E. Lifshitz**, *Quantum Mechanics*, Pergamon Press, 1958

do qual existem também edições posteriores aumentadas, notadamente em traduções francesas publicadas pela Editora MIR, Moscou.

O tratamento matematicamente mais preciso do problema espectral para operadores lineares em espaços de dimensão infinita (isto é, um tratamento “na linha von Neumann”, v. nota à pag. 59) pode ser encontrado, por exemplo, em

★ **A. N. Kolmogorov e S. V. Fomin**, *Elementos de la teoria de funciones y del analisis funcional*, Editorial MIR, Moscou, 1972, Capítulo IV

e especialmente em

★ **M. Reed e B. Simon**, *Methods of Modern Mathematical Physics*, Volume I, Academic Press, 1972

do qual vale a pena citar a epígrafe da seção preliminar: “The beginner...should not be discouraged if...he finds that he does not have the prerequisites for reading the prerequisites. (P. Halmos)”

Ao longo da última década houve um recrudescimento do interesse em questões relativas à interpretação da mecânica quântica, fortemente apoiado pela realização e análise experimental de algumas situações mais exóticas previstas pela teoria. Um tratamento bastante extenso dessas questões, bem como um rol conveniente de referências à literatura periódica, pode ser encontrada em

★ **Roland Omnès**, *The Interpretation of Quantum Mechanics*, Princeton Series in Physics, Princeton U. Press, 1994.

Capítulo 3

Sistemas quânticos simples.

Este capítulo trata do uso da teoria desenvolvida no capítulo precedente para o tratamento de sistemas quânticos simples e completamente solúveis. A solubilidade completa não pode chegar a ser vista como uma característica “típica” de sistemas quânticos, o que aliás é revelado pela grande importância e considerável sofisticação de métodos de aproximação desenvolvidos no contexto da mecânica quântica. Ela permite, no entanto, ilustrar o funcionamento efetivo da teoria e alguns aspectos cruciais do comportamento de sistemas quânticos independentemente das complicações e incertezas de ordem mais técnica ligadas ao uso de métodos aproximados de solução. As duas primeiras seções discutem soluções de problemas em termos “abstratos”, isto é, em termos de operadores e vetores de estado tratados independentemente de representações particulares. As seções seguintes introduzem e usam a chamada “representação de Schrödinger”, na qual os vetores de estado são representados por funções complexas da posição. Uma característica importante desta representação particular é a de permitir uma análise “local” (em termos de posição) de propriedades gerais, como a unitariedade da dinâmica quântica. Esse tipo de análise é particularmente relevante para compreender alguns aspectos do chamado regime semi-clássico da teoria quântica.

3.1 O oscilador harmônico.

O tratamento dado na seção 1.3.2 ao problema do oscilador harmônico baseou-se fortemente no esquema geral da mecânica das matrizes tal como formulado por Born e Jordan. Nesse esquema os vetores de estado correspondentes aos estados estacionários do oscilador (isto é, os autovetores do hamiltoniano do oscilador) não aparecem explicitamente, embora estejam implicitamente presentes no caráter diagonal exigido para a matriz que representa o hamiltoniano. Esta seção trata em primeiro lugar, portanto, de explicitar a natureza e propriedades dos

estados estacionários de um oscilador harmônico unidimensional.

Um problema mais geral, e que ilustra o uso de hamiltonianos que dependem explicitamente do tempo, é o do oscilador harmônico *forçado*, isto é, sujeito à ação de uma força externa dependente do tempo. O hamiltoniano a ser considerado nesse caso é

$$H(t) = \frac{p^2}{2M} + \frac{M\omega_0^2}{2}q^2 - f(t)q \quad (3.1)$$

onde p e q são entendidos como operadores satisfazendo a relação de comutação $[p, q] = \hbar/i$ e $f(t)$ é uma função real do tempo que representa uma força externa agindo sobre o oscilador. A realidade de $f(t)$ garante o caráter hermiteano de $H(t)$ em qualquer tempo, e o significado dessa função, que é o de uma força externa dependente do tempo que age sobre o oscilador, se torna claro calculando explicitamente a equação de movimento de Heisenberg para p (cf. Eq. (1.18))

$$\frac{dp}{dt} = -\frac{\partial H(t)}{\partial q} = -M\omega_0^2 q + f(t).$$

3.1.1 Espectro e autovetores do hamiltoniano do oscilador.

O problema a ser tratado aqui consiste em obter os autovalores e autovetores do hamiltoniano (3.1) no caso em que $f(t) \equiv 0$. Para fixar idéias, isso será feito numa descrição de Schrödinger, em que os vários operadores a serem considerados serão tratados como sendo independentes do tempo, sendo a dinâmica finalmente implementada através de uma dependência temporal do vetor de estado prescrita pela equação de movimento (2.28). O contacto com a descrição de Heisenberg será re-estabelecido posteriormente, fazendo uso do operador de evolução do oscilador livre.

Através das transformações (1.20) o hamiltoniano (3.1) pode novamente ser posto sob a forma (tomando aqui $f(t) \equiv 0$)

$$H = \hbar\omega_0 \left(a^\dagger a + \frac{1}{2} \right)$$

onde os operadores a e a^\dagger satisfazem a relação de comutação $[a, a^\dagger] = 1$. Desse modo, o problema de autovalores para o hamiltoniano fica reduzido ao problema de autovalores para o operador hermiteano $n \equiv a^\dagger a$, isto é, ao problema de identificar números ν e vetores $|\nu\rangle$ normalizados de forma que $\langle \nu | \nu \rangle = 1$ tais que

$$n|\nu\rangle \equiv a^\dagger a|\nu\rangle = \nu|\nu\rangle.$$

Supondo inicialmente que algum vetor com essa propriedade tenha sido identificado, usando a relação de comutação de a e a^\dagger pode-se ver facilmente que

$$a^\dagger a(a^\dagger|\nu\rangle) = a^\dagger(a^\dagger a + 1)|\nu\rangle = (\nu + 1)(a^\dagger|\nu\rangle)$$

e também que

$$a^\dagger a(a|\nu\rangle) = (aa^\dagger - 1)a|\nu\rangle = (\nu - 1)(a|\nu\rangle)$$

de modo que os vetores $a^\dagger|\nu\rangle$ e $a|\nu\rangle$ são também autovetores de n com autovalores $\nu + 1$ e $\nu - 1$ respectivamente. Supondo que o vetor $|\nu\rangle$ seja devidamente normalizado, a forma normalizada dos autovetores deduzidos $|\nu \pm 1\rangle$ pode ser obtida calculando constantes de normalização $c_\pm(\nu)$ tais que

$$1 = |c_+(\nu)|^2 \langle \nu | aa^\dagger | \nu \rangle = |c_+(\nu)|^2 \langle \nu | (a^\dagger + 1) | \nu \rangle = |c_+(\nu)|^2 (\nu + 1)$$

e

$$1 = |c_-(\nu)|^2 \langle \nu | a^\dagger a | \nu \rangle = |c_-(\nu)|^2 \nu.$$

Como ocorre normalmente na determinação de constantes de normalização para autovetores, essas relações determinam $c_\pm(\nu)$ a menos de um fator de fase. Fazendo a escolha mais simples, que consiste em tomar essas constantes como reais e positivas (supondo $\nu > 0$), resulta que de um autovetor normalizado $|\nu\rangle$ associado a um autovalor ν pode-se deduzir os novos autovetores, também normalizados, $|\nu \pm 1\rangle$, associados a autovalores $\nu \pm 1$,

$$|\nu + 1\rangle = \frac{1}{\sqrt{\nu + 1}} a^\dagger |\nu\rangle \quad \text{e} \quad |\nu - 1\rangle = \frac{1}{\sqrt{\nu}} a |\nu\rangle.$$

Não é difícil ver, em seguida, que os valores possíveis de ν devem ser limitados inferiormente. De fato, se ν é um autovalor, então $\nu - 1$ também será um autovalor sempre que a norma do vetor $a|\nu\rangle$ seja *positiva*, pois o vetor nulo não é aceitável como autovetor. Por outro lado, essa norma pode ser calculada explicitamente com o resultado

$$\langle \nu | a^\dagger a | \nu \rangle = \nu.$$

A positividade da norma exige então que $\nu \geq 0$, e para que isso seja sempre verdade, os valores admissíveis de ν são os *inteiros não negativos*. De fato, qualquer autovalor ν *não inteiro* levaria a um vetor de norma negativa após um número suficiente de aplicações do operador a ao respectivo autovetor. Se, por outro lado ν é um inteiro positivo, a aplicação do operador

a ao autovetor $|\nu\rangle$, repetida $\nu + 1$ vezes, leva ao vetor nulo, e conseqüentemente à terminação da seqüência descendente de autovalores e autovetores. O menor autovalor do oscilador é claramente aquele em que $\nu = 0$, pois o vetor $a|\nu = 0\rangle$ tem norma nula. O autovetor $|\nu = 0\rangle$ correspondente ao estado fundamental é portanto *definido* pela condição $a|\nu = 0\rangle = 0$.

Dessa forma, o espectro de $n \equiv a^\dagger a$ consiste dos inteiros não negativos $\nu = 0, 1, 2, \dots$ e o espectro do hamiltoniano é dado por

$$E_\nu = \hbar\omega_0 \left(\nu + \frac{1}{2} \right) ;$$

os autovetores normalizados correspondentes podem ser designados como $|\nu\rangle$, $\nu = 0, 1, 2, \dots$, são ortogonais (pois são autovetores de um operador hermiteano associados a autovalores diferentes) e satisfazem às propriedades

$$a^\dagger |\nu\rangle = \sqrt{\nu + 1} |\nu + 1\rangle \quad \text{e} \quad a |\nu\rangle = \sqrt{\nu} |\nu - 1\rangle$$

que de fato permitem calcular elementos de matriz para qualquer variável dinâmica. Tomando como exemplo o operador q , usando a transformação que o exprime em termos de a e a^\dagger

$$q = \sqrt{\frac{\hbar}{2M\omega_0}} (a^\dagger + a)$$

é imediato que

$$\langle \nu' | q | \nu \rangle = \sqrt{\frac{\hbar}{2M\omega_0}} \left(\sqrt{\nu + 1} \delta_{\nu', \nu+1} + \sqrt{\nu} \delta_{\nu', \nu-1} \right) .$$

Em particular, elementos de matriz diagonais $\langle \nu | q | \nu \rangle$ são nulos. Resultados semelhantes valem também para os elementos de matriz de p , dado que este operador é também uma combinação de a e a^\dagger .

O anulamento dos valores médios de q e p nos autovetores do hamiltoniano permite calcular os produtos de incerteza desses dois operadores, nesses estados, simplesmente a partir dos valores médios dos quadrados dos operadores p e q

$$p^2 = -\frac{\hbar M\omega_0}{2} (a^{\dagger 2} + a^2 - 2a^\dagger a - 1) \quad \text{e} \quad q^2 = \frac{\hbar}{2M\omega_0} (a^{\dagger 2} + a^2 + 2a^\dagger a + 1)$$

o que leva a

$$\sigma_p^{2(\nu)} \sigma_q^{2(\nu)} = \langle \nu | p^2 | \nu \rangle \langle \nu | q^2 | \nu \rangle = \hbar^2 \left(\nu + \frac{1}{2} \right)^2$$

mostrando que os produtos de incerteza crescem com ν e que o estado fundamental $\nu = 0$ é na realidade um *estado de incerteza mínima*.

Finalmente, é útil discutir o uso dos autovalores e autovetores em conexão com a dinâmica do oscilador harmônico. O elemento essencial, comum às diferentes realizações da dinâmica do oscilador harmônico, é o operador de evolução

$$U(t, t_0) \equiv e^{-\frac{i}{\hbar}H(t-t_0)} = e^{-i\omega_0(t-t_0)(a^\dagger a + \frac{1}{2})}.$$

Como observado na seção 2.3.1, introduzindo a resolução da unidade em termos dos autovetores de H

$$\hat{1} = \sum_{\nu=0}^{\infty} |\nu\rangle\langle\nu|$$

esse operador pode ser escrito sob a forma

$$U(t, t_0) = \sum_{\nu=0}^{\infty} |\nu\rangle e^{-i(\nu+\frac{1}{2})\omega_0(t-t_0)} \langle\nu|$$

que é muito útil para exprimir os mais diversos objetos dinâmicos, seja na descrição de Schrödinger seja na de Heisenberg. Por exemplo, a *solução* do problema de condições iniciais na descrição de Schrödinger, que consiste em determinar a solução da equação de movimento (2.28) que em $t = 0$ se reduz a um vetor de estado $|\varphi(0)\rangle$ dado, pode ser escrita imediatamente como

$$|\varphi(t)\rangle = U(t, 0)|\varphi(0)\rangle = \sum_{\nu=0}^{\infty} |\nu\rangle e^{-i(\nu+\frac{1}{2})\omega_0 t} \langle\nu|\varphi(0)\rangle.$$

Por outro lado, os observáveis dependentes do tempo da descrição de Heisenberg podem ser obtidos a partir de $a(t)$ e de seu hermiteano adjunto $a^\dagger(t)$, e estes por sua vez podem ser obtidos facilmente a partir do operador de evolução. De fato,

$$\begin{aligned} a(t) = U^\dagger(t, 0)a(0)U(t, 0) &= \sum_{\nu'=0}^{\infty} \sum_{\nu=0}^{\infty} |\nu'\rangle e^{i\nu'\omega_0 t} \langle\nu'|a(0)|\nu\rangle e^{-i\nu\omega_0 t} \langle\nu| \\ &= e^{-i\omega_0 t} \hat{1} a(0) \hat{1} = e^{-i\omega_0 t} a(0) \end{aligned}$$

donde resulta que $a^\dagger(t) = e^{i\omega_0 t} a^\dagger(0)$ e

$$q(t) = \sqrt{\frac{\hbar}{2M\omega_0}} (e^{i\omega_0 t} a^\dagger + e^{-i\omega_0 t} a), \quad p(t) = i\sqrt{\frac{\hbar M\omega_0}{2}} (e^{i\omega_0 t} a^\dagger - e^{-i\omega_0 t} a).$$

Quando os operadores $p(t)$ e $q(t)$ (ou, equivalentemente, $a(t)$ e $a^\dagger(t)$) são substituídos no hamiltoniano, a dependência temporal, associada à representação de Heisenberg, se cancela, ilustrando o fato geral de que um hamiltoniano que não depende explicitamente do tempo é independente do tempo, isto é, uma constante do movimento.

3.1.2 Dinâmica do oscilador forçado.

O problema da dinâmica de oscilador quântico *forçado*, envolve o hamiltoniano dependente do tempo (3.1), agora com $f(t)$ diferente de zero. A sua solução é dada, portanto, em termos de um operador de evolução que satisfaz a equação diferencial (2.31) com a respectiva condição inicial ou, equivalentemente, a equação integral (2.32). Na descrição de Schrödinger, esse operador de evolução determina a evolução temporal do vetor de estado a partir de um vetor de estado inicial dado em $t = 0$, isto é

$$|\varphi(t)\rangle = U(t, 0)|\varphi(t = 0)\rangle.$$

Esse vetor de estado dependente do tempo permite em seguida calcular as propriedades dos vários observáveis que sejam de interesse. A substituição de p e q pelos operadores a e a^\dagger é também útil para simplificar a forma do hamiltoniano, que aparece então sob a forma

$$H(t) = \hbar\omega_0 \left(a^\dagger a + \frac{1}{2} \right) - \sqrt{\frac{\hbar}{2M\omega_0}} f(t)(a^\dagger + a) \equiv \hbar\omega_0 \left(a^\dagger a + \frac{1}{2} \right) - F(t)(a^\dagger + a). \quad (3.2)$$

Alternativamente, o problema pode ser tratado (e com maior simplicidade, neste caso) na descrição de Heisenberg, em que o vetor de estado é sempre aquele correspondente à condição inicial dada, enquanto os observáveis de interesse são dependentes do tempo, satisfazendo à equação de movimento (2.33). Como os vários observáveis são exprimíveis em termos de $p(t)$ e $q(t)$, e portanto também em termos de $a(t)$ e $a^\dagger(t)$, basta considerar a equação de movimento para $a(t)$, que pode ser escrita explicitamente como

$$i\hbar \frac{da(t)}{dt} = [a(t), H_H(t)] = \hbar\omega_0 a(t) - F(t)\hat{1} \quad (3.3)$$

cuja solução também pode ser expressa em termos do operador de evolução como

$$a(t) = U^\dagger(t, 0)a(t = 0)U(t, 0) \quad (3.4)$$

onde $a(t = 0)$ pode ser identificado com o operador que aparece na descrição de Schrödinger. Nessa equação, o hamiltoniano $H_H(t)$ tem a mesma forma que $H(t)$ em (3.2), mas com a e a^\dagger substituídos pelos correspondentes operadores de Heisenberg $a(t)$ e $a^\dagger(t)$, que também satisfazem a relação de comutação em tempos iguais $[a(t), a^\dagger(t)] = 1$. Isso corresponde à definição geral para as variáveis dinâmicas na descrição de Heisenberg

$$H_H(t) = U^\dagger(t, 0)H(t)U(t, 0)$$

que leva agora a um resultado que *difere* do hamiltoniano $H(t)$ correspondente à descrição de Schrödinger devido à não comutatividade de $H(t)$ com $U(t)$. De fato, o operador de evolução, definido pelas equações (2.31) ou (2.32), envolve o hamiltoniano em tempos $t' \neq t$, e $[H(t), H(t')] \neq 0$ sempre que $f(t) \neq f(t')$.

Um modo simples de resolver esse problema dinâmico, sem necessidade de obter o operador de evolução explicitamente, consiste em considerar diretamente a equação (3.3) para o operador $a(t)$. Esta equação pode ser facilmente resolvida observando que

$$\frac{d}{dt} \left(e^{i\omega_0 t} a(t) \right) = i\omega_0 e^{i\omega_0 t} a(t) + e^{i\omega_0 t} \frac{da(t)}{dt}$$

o que permite reescreve-la sob a forma

$$\frac{d}{dt} \left(e^{i\omega_0 t} a(t) \right) = \frac{i}{\hbar} e^{i\omega_0 t} F(t) \hat{1} \quad (3.5)$$

que pode ser imediatamente integrada com o resultado

$$a(t) = e^{-i\omega_0 t} a(0) + \frac{i}{\hbar} e^{-i\omega_0 t} \int_0^t dt' e^{i\omega_0 t'} F(t') \hat{1} \equiv e^{-i\omega_0 t} \left(a(0) + v(t) \hat{1} \right) \quad (3.6)$$

onde foi introduzida por conveniência a abreviação

$$v(t) \equiv \frac{i}{\hbar} \int_0^t dt' e^{i\omega_0 t'} F(t').$$

Várias observações devem ser feitas com relação a este resultado. Em primeiro lugar, o primeiro termo da solução para $a(t)$ provém da consideração da *condição inicial* de que para $t = 0$ o operador de Heisenberg $a(t)$ deve se reduzir a $a(0)$, ao resolver a equação diferencial (3.5). Esse termo é independente da força externa $F(t)$ e de fato coincide com o operador de Heisenberg $a(t)$ para o oscilador *livre*. Em segundo lugar, deve ser observado que todos os efeitos da força externa estão contidos no último termo da solução, que tem a forma bastante especial de um múltiplo (dependente do tempo) do operador unidade. Essa forma garante, em particular, que a relação de comutação $[a(t), a^\dagger(t)] = 1$ valha sempre, isto é, qualquer que seja a força externa aplicada ao oscilador. De fato, a equação (3.6) resolve *completamente* o problema do oscilador forçado, no sentido de que permite calcular propriedades relativas a quaisquer observáveis, para qualquer vetor de estado que seja especificado em $t = 0$. Como exemplos simples do uso dessa solução, os valores médios da posição e do momento do oscilador forçado, quando ele está inicialmente no estado fundamental $|\nu = 0\rangle$ são dados respectivamente por

$$\langle q(t) \rangle = \sqrt{\frac{\hbar}{2M\omega_0}} \langle 0 | (a^\dagger(t) + a(t)) | 0 \rangle = \sqrt{\frac{2\hbar}{M\omega_0}} \operatorname{Re} \left[e^{-i\omega_0 t} v(t) \right]$$

e

$$\langle p(t) \rangle = i\sqrt{\frac{\hbar M\omega_0}{2}} \langle 0 | (a^\dagger(t) - a(t)) | 0 \rangle = \sqrt{2\hbar M\omega_0} \operatorname{Im} [e^{-i\omega_0 t} v(t)] .$$

Para essa condição inicial particular (oscilador inicialmente no estado $|\nu = 0\rangle$) é possível obter explicitamente a solução do problema na descrição de Schrödinger, a partir da equação (3.6), do seguinte modo. Como $a(0)|\nu = 0\rangle = 0$, aplicando o operador $a(t)$ sobre o estado fundamental obtém-se

$$a(t)|\nu = 0\rangle = e^{-i\omega_0 t} v(t)|\nu = 0\rangle .$$

Por outro lado, o operador de Heisenberg $a(t)$ pode ser escrito em termos de $a(0)$ e do operador de evolução $U(t, 0)$ como na equação (3.4), de modo que, usando a unitariedade de $U(t, 0)$,

$$U(t, 0)a(t)|\nu = 0\rangle = a(0)U(t, 0)|\nu = 0\rangle = e^{-i\omega_0 t} v(t)U(t, 0)|\nu = 0\rangle . \quad (3.7)$$

No entanto, o estado $U(t, 0)|\nu = 0\rangle$ é a *solução* da equação de movimento, correspondente à condição inicial considerada, para o estado do oscilador forçado na descrição de Schrödinger, e a última equação mostra que esse estado é, no tempo t , um autovetor do operador $a(0)$ com autovalor $e^{-i\omega_0 t} v(t)$. Vetores de estado com essa propriedade são conhecidos como *estados coerentes* do oscilador harmônico e podem ser facilmente expressos, por exemplo, em termos dos autovetores $\{|\nu\rangle\}$ do hamiltoniano do oscilador livre. De fato, impondo a relação

$$a|z\rangle = z|z\rangle$$

onde z é um número complexo, e introduzindo a expansão

$$|z\rangle = \sum_{\nu=0}^{\infty} c_\nu |\nu\rangle$$

resulta, usando a forma explícita da ação de a sobre os autovetores normalizados $|\nu\rangle$, que se deve ter

$$\sum_{\nu=1}^{\infty} c_\nu \sqrt{\nu} |\nu - 1\rangle = z \sum_{\nu=0}^{\infty} c_\nu |\nu\rangle .$$

Comparando os coeficientes de um dado autovetor $|\nu\rangle$ das duas expansões resulta então que

$$c_{\nu+1} = \frac{z}{\sqrt{\nu+1}} c_\nu \quad \text{ou ainda} \quad c_\nu = \frac{z^\nu}{\sqrt{\nu!}} c_0$$

de modo que se obtém para o autovetor de a

$$|z\rangle = c_0 \sum_{\nu=0}^{\infty} \frac{z^{\nu}}{\sqrt{\nu!}} |\nu\rangle.$$

O valor de c_0 pode ser determinado pela condição de normalização $\langle z|z\rangle = 1$ que dá, em vista da ortonormalidade dos $|\nu\rangle$,

$$1 = |c_0|^2 \sum_{\nu=0}^{\infty} \frac{|z|^{2\nu}}{\nu!} = |c_0|^2 e^{|z|^2}.$$

Com a escolha mais simples da fase de c_0 , a forma final normalizada dos estados coerentes é portanto (*cf.* problema 1.5)

$$|z\rangle = e^{-\frac{|z|^2}{2}} \sum_{\nu=0}^{\infty} \frac{z^{\nu}}{\sqrt{\nu!}} |\nu\rangle.$$

Esse resultado mostra que *qualquer número complexo* z é autovalor do operador (não hermiteano) a , tendo como autovetor *normalizado* o estado coerente construído com esse autovalor. É fácil verificar que estados coerentes com valores diferentes de z *não são ortogonais*. Em vista do resultado (3.7), a solução da equação de movimento, na descrição de Schrödinger, para o vetor de estado de um oscilador forçado que se encontra inicialmente em seu estado fundamental é na realidade o estado coerente

$$|z = e^{-i\omega_0 t} v(t)\rangle = e^{-\frac{|v(t)|^2}{2}} \sum_{\nu=0}^{\infty} \frac{e^{-i\nu\omega_0 t} v(t)^{\nu}}{\sqrt{\nu!}} |\nu\rangle.$$

Os estados coerentes têm uma série de outras propriedades notáveis, algumas das quais são tratadas nos problemas deste capítulo. Em particular, eles são estados de incerteza mínima para qualquer valor de z , cuja evolução temporal, prescrita pela equação de movimento da descrição de Schrödinger para o oscilador *livre*, permanece um estado coerente. Nesse caso, a dependência das partes real e imaginária de z com o tempo obedece às equações de Hamilton para um oscilador harmônico *clássico*.

3.2 Sistemas de dois níveis.

Sistemas quânticos complexos com espectro de energia discreto são freqüentemente encontrados em situações dinâmicas especiais nas quais apenas dois estados estacionários têm participação significativa. Em tais casos é freqüente o uso de uma aproximação esquemática em que todos os

outros estados estacionários são ignorados e o sistema fica portanto reduzido a um “sistema de dois níveis”. O interesse desse tipo de aproximação vem de que ela permite tratar a dinâmica quântica resultante de forma exata e completa, em termos analíticos. Um exemplo frequentemente esquematizado dessa forma é o da molécula de amônia (v. *The Feynman Lectures on Physics*, Addison-Wesley Publ. Co., 1965, vol. III, Lecture 9). Outra situação que se presta, e com razoável realismo, a um tratamento nesses termos é a que envolve a interação ressonante de dois estados estacionários atômicos com o campo eletromagnético de uma cavidade, permitindo ignorar a participação de outros níveis não ressonantes (v. S. Haroche e P. Raymond, em *Cavity Quantum Electrodynamics*, P. R. Berman (ed.), Academic Press, N.Y., 1994).

A dinâmica quântica de um sistema de dois níveis se dá em um espaço vetorial complexo de duas dimensões, que pode portanto ser tratado em termos de uma base ortonormal de dois vetores $\{|\phi_1\rangle, |\phi_2\rangle\}$. Nesse espaço, qualquer operador linear pode ser escrito como combinação linear de *quatro* operadores que, em termos dessa base, podem ser tomados como sendo $|\phi_i\rangle\langle\phi_j|$, $i, j = 1, 2$. Desse modo, um operador linear qualquer g pode ser representado como

$$g = \sum_{i,j=1}^2 g_{ij} |\phi_i\rangle\langle\phi_j|$$

ou, equivalentemente, em termos da matriz 2×2

$$g \rightarrow \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

enquanto um vetor geral do espaço pode ser escrito como uma combinação linear dos dois vetores de base, $|\varphi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle$, que pode ser também expressa sob a forma

$$|\varphi\rangle \rightarrow \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Para que o estado $|\varphi\rangle$ esteja devidamente normalizado é preciso ainda impor sobre os coeficientes dessa expansão a condição $|c_1|^2 + |c_2|^2 = 1$.

Adotando como vetores de base os vetores associados aos dois estados estacionários relevantes, o hamiltoniano H_0 do sistema de dois níveis é representado por uma matriz diagonal que, definindo a escala de energia de forma que o estado de energia menor tenha autovalor zero, pode ser escrita como

$$H_0 \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & \Delta \end{pmatrix}$$

onde $\Delta = \langle\phi_2|H_0|\phi_2\rangle \geq 0$ é a energia do segundo estado. O caso $\Delta = 0$ corresponde evidentemente ao caso de dois níveis degenerados. A matriz correspondente ao operador de evolução,

definido pela equação (2.26) com a condição inicial $U_0(t_0, t_0) = 1$, tem elementos $\langle \phi_i | U_0(t, t_0) | \phi_j \rangle$ e é portanto dada por

$$U_0(t, t_0) \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & e^{-\frac{i}{\hbar} \Delta(t-t_0)} \end{pmatrix}$$

que de fato resolve completamente o problema dinâmico do sistema de dois níveis com o hamiltoniano H_0 .

Uma modificação da dinâmica do sistema de dois níveis (através da introdução de campos externos independentes do tempo, por exemplo) tal que ela continue restrita ao espaço de duas dimensões gerado pela base $\{|\phi_1\rangle, |\phi_2\rangle\}$, pode sempre ser representada adicionando a H_0 um termo H_1 que, para que a hermiticidade do hamiltoniano completo $H \equiv H_0 + H_1$ seja mantida, deve ser da forma

$$H_1 \rightarrow \begin{pmatrix} \delta_1 & \gamma \\ \gamma^* & \delta_2 \end{pmatrix}$$

onde em geral γ pode ser complexo, mas δ_1 e δ_2 são reais. Isso de fato define uma nova dinâmica, governada pelo hamiltoniano modificado

$$H = H_0 + H_1 \rightarrow \begin{pmatrix} \delta_1 & \gamma \\ \gamma^* & \Delta + \delta_2 \end{pmatrix}. \quad (3.8)$$

Como foi dito, essa dinâmica pode ser, por sua vez, tratada exatamente de forma analítica e, apesar de sua extrema simplicidade, usada para ilustrar uma série de propriedades quânticas presentes também em sistemas mais complexos. As seções seguintes são dedicadas a essas questões.

3.2.1 Estados estacionários e o operador de evolução.

O fato de que H não é diagonal na base $\{|\phi_1\rangle, |\phi_2\rangle\}$ significa que os vetores dessa base não são autovetores de H , e que novos estados estacionários devem ser determinados através da diagonalização da matriz (3.8). Os autovalores dessa matriz correspondem às energias dos estados estacionários do sistema de dois níveis modificado, e são as raízes da equação secular

$$E^2 - (\delta_1 + \delta_2 + \Delta)E + \delta_1(\Delta + \delta_2) - |\gamma|^2 = 0$$

que são dadas por

$$E_{\pm} = \frac{1}{2} \left(\delta_1 + \delta_2 + \Delta \pm \sqrt{(\delta_2 + \Delta - \delta_1)^2 + 4|\gamma|^2} \right).$$

Uma primeira observação que pode ser feita com relação a esse resultado é a de que a diferença de energia entre os novos estados estacionários é

$$E_+ - E_- = \sqrt{(\delta_2 + \Delta - \delta_1)^2 + 4|\gamma|^2}$$

que se anula *apenas se* $\gamma = 0$ e *também* $\delta_1 = \delta_2 + \Delta$, o que equivale a dizer que a degenerescência dos dois níveis só ocorrerá quando a matriz H for um múltiplo da matriz unidade. Tal propriedade é na realidade independente da base escolhida para representar H , e portanto em *qualquer* outro caso os autovalores *não* serão degenerados, independentemente do caráter diagonal (em que isso é na realidade óbvio) ou não diagonal de H na base escolhida. A “propensão à não degenerescência” representada pelo fato de que os autovalores de qualquer matriz hermiteana 2×2 diferente de um múltiplo da matriz unidade são distintos é de fato um caso particular de uma propriedade de hamiltonianos mais gerais conhecida como “repulsão de níveis”.

Conhecidos os autovalores de H , é fácil determinar analiticamente também os autovetores correspondentes. Embora não haja maiores dificuldades para se tratar o caso geral do hamiltoniano (3.8), para simplificar as expressões algébricas e também a discussão, os resultados explícitos dados a seguir serão limitados ao caso particular em que $\delta_1 = \delta_2 = 0$. Com uma escolha natural para os fatores de fase arbitrários na determinação dos autovetores, eles podem ser representados, na mesma base utilizada para a representação matricial de H , pelos vetores

$$|E_+\rangle \rightarrow N \begin{pmatrix} 2\gamma \\ \Delta + \sqrt{\Delta^2 + 4|\gamma|^2} \end{pmatrix} \quad \text{e} \quad |E_-\rangle \rightarrow N \begin{pmatrix} \Delta + \sqrt{\Delta^2 + 4|\gamma|^2} \\ -2\gamma^* \end{pmatrix}, \quad \delta_1 = \delta_2 = 0 \quad (3.9)$$

onde N é o coeficiente de normalização

$$N = \left[\left(\Delta + \sqrt{\Delta^2 + 4|\gamma|^2} \right)^2 + 4|\gamma|^2 \right]^{-\frac{1}{2}}.$$

Essas expressões são suficientemente pouco transparentes para que valha a pena estudar alguns limites em que elas se reduzem a formas mais simples, para avaliar melhor o seu conteúdo. No caso em que $\gamma \rightarrow 0$, a matriz H é diagonal e, como pode ser verificado, os estados estacionários $|E_-\rangle$ e $|E_+\rangle$ se reduzem respectivamente a $|\phi_1\rangle$ e a $|\phi_2\rangle$. Outro limite interessante é aquele em que $\Delta \rightarrow 0$ com $\gamma \neq 0$, isto é, os níveis de H_0 são degenerados. Nesse caso os novos estados estacionários se reduzem a

$$|E_+\rangle = \frac{1}{\sqrt{2}} \left(\frac{\gamma}{|\gamma|} |\phi_1\rangle + |\phi_2\rangle \right) \quad \text{e} \quad |E_-\rangle = \frac{1}{\sqrt{2}} \left(|\phi_1\rangle - \frac{\gamma^*}{|\gamma|} |\phi_2\rangle \right), \quad \delta_1 = \delta_2 = \Delta = 0. \quad (3.10)$$

É imediato verificar que eles são ortogonais para qualquer valor, em geral complexo, de γ . Vale a pena observar que a fase relativa das duas componentes desses estados na base escolhida depende da fase do elemento de matriz não diagonal γ de H . Em particular, para γ real e positivo, o estado de menor energia $|E_- \rangle$ corresponde à combinação *antissimétrica* de $|\phi_1 \rangle$ e $|\phi_2 \rangle$, ao passo que quando γ é real e negativo ele corresponde à combinação *simétrica*.

Da mesma forma que no caso simples envolvendo apenas o hamiltoniano H_0 , a determinação dos autovalores e autovetores de H resolve completamente a dinâmica do sistema descrito por esse hamiltoniano modificado. De fato a resolução da unidade em termos dos autovetores de H , $\hat{1} = |E_- \rangle \langle E_-| + |E_+ \rangle \langle E_+|$, permite escrever o operador de evolução sob a forma

$$U(t, 0) = e^{-\frac{i}{\hbar} H t} (|E_- \rangle \langle E_-| + |E_+ \rangle \langle E_+|) = |E_- \rangle e^{-\frac{i}{\hbar} E_- t} \langle E_-| + |E_+ \rangle e^{-\frac{i}{\hbar} E_+ t} \langle E_+|. \quad (3.11)$$

Por outro lado, como a expansão dos autovetores na base de referência $\{|\phi_1 \rangle, |\phi_2 \rangle\}$ é conhecida, as quantidades $\langle E_{\pm} | \varphi \rangle$ podem ser calculadas para qualquer vetor de estado $|\varphi \rangle$ especificado em termos de suas componentes nessa mesma base, o que permite obter, por exemplo, o vetor $|\varphi(t) \rangle = U(t, 0) |\varphi(0) \rangle$ que caracteriza completamente o comportamento dinâmico na descrição de Schrödinger.

Como exemplo concreto do uso desse procedimento, a probabilidade de observar o sistema no estado $|\phi_2 \rangle$ no tempo t , dado que ele se encontra no estado $|\phi_1 \rangle$ em $t = 0$, é dada por $p_{1 \rightarrow 2}(t) = |\langle \phi_2 | U(t, 0) | \phi_1 \rangle|^2$. A probabilidade $p_{1 \rightarrow 1}(t) = |\langle \phi_1 | U(t, 0) | \phi_1 \rangle|^2$ de que o sistema *permaneça* no estado $|\phi_1 \rangle$ no tempo t é evidentemente dada por $p_{1 \rightarrow 1}(t) = 1 - p_{1 \rightarrow 2}(t)$, pois a unitariedade de $U(t, 0)$ garante que o estado $U(t, 0) |\phi_1 \rangle$ tem a mesma normalização de $|\phi_1 \rangle$ e os dois vetores $|\phi_1 \rangle$ e $|\phi_2 \rangle$ constituem uma base ortonormal para o sistema de dois níveis. Utilizando a forma (3.11) do operador de evolução, $p_{1 \rightarrow 2}(t)$ pode ser calculada, fazendo uso da ortogonalidade de $|\phi_1 \rangle$ e $|\phi_2 \rangle$, da resolução da unidade em termos dos autovetores de H e da normalização dos estados $|E_{\pm} \rangle$, como

$$\begin{aligned} p_{1 \rightarrow 2}(t) &= \left| \langle \phi_2 | E_- \rangle e^{-\frac{i}{\hbar} E_- t} \langle E_- | \phi_1 \rangle + \langle \phi_2 | E_+ \rangle e^{-\frac{i}{\hbar} E_+ t} \langle E_+ | \phi_1 \rangle \right|^2 \\ &= \left| \langle \phi_2 | E_+ \rangle \left(e^{-\frac{i}{\hbar} (E_+ - E_-) t} - 1 \right) \langle E_+ | \phi_1 \rangle \right|^2 \\ &= 4 |\langle \phi_2 | E_+ \rangle \langle E_+ | \phi_1 \rangle|^2 \sin^2 \left(\frac{E_+ - E_-}{2\hbar} t \right) \\ &= 4 |\langle \phi_1 | E_+ \rangle|^2 \left(1 - |\langle \phi_1 | E_+ \rangle|^2 \right) \sin^2 \left(\frac{E_+ - E_-}{2\hbar} t \right). \end{aligned} \quad (3.12)$$

Essa probabilidade se anula em $t = 0$ como consequência da ortogonalidade de $|\phi_1 \rangle$ e $|\phi_2 \rangle$ e tem um comportamento oscilatório para $t > 0$ com a frequência de Bohr correspondente à

diferença de energia entre os dois estados estacionários. A amplitude das oscilações depende da constituição dos estados estacionários em termos dos estados $|\phi_1\rangle$ e $|\phi_2\rangle$ e de fato é máxima (e igual a 1) no caso em que $|\langle\phi_1|E_+\rangle|^2 = 1/2$, que corresponde ao limite (3.10), em que $\delta_1 = \delta_2 = \Delta = 0$ com $\gamma \neq 0$. No caso mais geral em que $\Delta \neq 0$ a amplitude pode ser calculada usando os autovetores (3.9), e é dada por

$$\mathcal{A}(|\gamma|, \Delta) \equiv 4 |\langle\phi_1|E_+\rangle|^2 \left(1 - |\langle\phi_1|E_+\rangle|^2\right) = \frac{16|\gamma|^2 \left(\Delta + \sqrt{\Delta^2 + 4|\gamma|^2}\right)^2}{\left[\left(\Delta + \sqrt{\Delta^2 + 4|\gamma|^2}\right)^2 + 4|\gamma|^2\right]^2}$$

de onde se pode extrair ainda a aproximação, válida no limite em que $|\gamma| \ll \Delta$,

$$\mathcal{A}(|\gamma|, \Delta) = 4 \frac{|\gamma|^2}{\Delta^2} + \mathcal{O}\left(\frac{|\gamma|^4}{\Delta^4}\right). \quad (3.13)$$

Lembrando também da relação simples existente entre $p_{1 \rightarrow 1}$ e $p_{1 \rightarrow 2}$, esses resultados mostram claramente que o caráter não estacionário, sob a dinâmica de H , dos autovetores $|\phi_1\rangle$ e $|\phi_2\rangle$ de H_0 tende a ser *inibido* por um aumento da diferença de energia Δ entre eles, para um dado valor do termo não diagonal γ , no sentido que a amplitude $\mathcal{A}(|\gamma|, \Delta)$ diminui quando Δ cresce com γ fixo. Isso é, na realidade, a contrapartida dinâmica do fato de que, nessas condições, as superposições de $|\phi_1\rangle$ e $|\phi_2\rangle$ que correspondem aos novos autovetores se tornam mais assimétricas, com predominância do autovetor de H_0 cujo autovalor é mais próximo do autovalor do novo autovetor considerado. A expansão (3.13) sugere que o uso de métodos aproximados de cálculo pode ser útil para descrever os efeitos da modificação de H_0 quando $|\gamma|/\Delta$ for suficientemente pequeno. Embora a existência de uma solução analítica exata torne tais métodos dispensáveis no caso simples de sistemas de dois níveis, eles são de utilidade considerável em situações mais complicadas nas quais soluções exatas não são práticas ou simplesmente não podem ser obtidas. Existe, por isso um conjunto extenso de métodos sistemáticos de aproximação aplicáveis em situações razoavelmente gerais, chamados genericamente *métodos perturbativos*, que serão tratados no Capítulo 4. Esses métodos permitem, por exemplo, obter resultados como o da equação (3.13) diretamente a partir dos autovetores $|\phi_1\rangle$ e $|\phi_2\rangle$ de H_0 e do hamiltoniano (3.8) (com $\delta_1 = \delta_2 = 0$), independentemente do conhecimento dos seus autovalores e autovetores exatos.

3.2.2 “Tunelamento” e interferência de estados estacionários.

Uma situação concreta que ilustra o uso dos resultados gerais obtidos acima é a um sistema que consiste em um grupo “molecular” de muitos átomos, que se comporta como um objeto magnético de spin S alto ($S = 10\hbar$ para $\text{Mn}_{12}\text{O}_{12}$) que, devido à sua interação com o meio

em que se encontra, prefere se orientar paralela ou anti-paralelamente a uma direção dada, que é identificada com o eixo z de um sistema de coordenadas¹. Uma primeira modelagem desse sistema consiste em considera-lo descrito por um hamiltoniano do tipo

$$h_0 = -A S_z^2$$

onde A é uma constante positiva e S_z é o operador que corresponde à componente do spin ao longo da direção z . Como será discutido mais adiante (v. seção 5.1.1), esse operador tem um espectro discreto com $2S + 1$ autovetores $|m\rangle$ e autovalores $m = -S\hbar, (-S + 1)\hbar, (-S + 2)\hbar, \dots, S\hbar$. Desse modo, os vetores $|m\rangle$ são também autovetores de h_0 com autovalores $-Am^2$, de modo que os pares de autovetores $|m\rangle$ e $|-m\rangle$ são degenerados. Em particular, o nível de menor energia de h_0 é também duplamente degenerado, com autovetores $|S\hbar\rangle$ e $|-S\hbar\rangle$. Esses autovetores descrevem o spin orientado paralela e anti-paralelamente ao eixo z , respectivamente.

Essa descrição do sistema é no entanto ainda excessivamente simplificada. Uma descrição melhor é obtida tendo em conta outros efeitos do meio em que o complexo molecular se encontra, adicionando a h_0 outro termo h_1 que tem a propriedade crucial de não ser diagonal na base dos autovetores de h_0 , isto é, $\langle m'|h_1|m\rangle \neq 0$ para $m' \neq m$, e de ter elementos de matriz diagonais nulos nessa base, $\langle m|h_1|m\rangle = 0$. Esse termo adicional é no entanto suficientemente pequeno para que o seu efeito seja desprezível sempre que considerada com relação a estados estacionários de h_0 com energias diferentes. Essa situação é semelhante à encontrada na equação (3.13), entendendo a quantidade Δ que aparece ali como a diferença de energia dos autovetores de h_0 e $|\gamma|$ como o elemento de matriz relevante envolvendo h_1 . Nessas condições, é possível estudar os efeitos de h_1 sobre cada um dos dubletos degenerados de h_0 separadamente, e em particular para o dubleto correspondente ao estado fundamental. Isso reduz o problema a um conjunto de problemas envolvendo apenas dois níveis. No caso do estado fundamental, que será considerado a seguir, os dois níveis correspondem aos autovetores $|\pm S\hbar\rangle$, degenerados para h_0 mas modificados por um termo não diagonal² $\gamma = \langle -S\hbar|h_1|S\hbar\rangle$. O modelo resultante é neste caso descrito por um hamiltoniano com a forma (3.8), com $\Delta = \delta_1 = \delta_2 = 0$.

Os resultados obtidos na seção anterior são suficientes para resolver completamente esse modelo. Os autovetores do hamiltoniano h_0 modificado por h_1 e restrito ao dubleto que constitui o estado fundamental de h_0 são dados por (3.10). Devido aos efeitos de h_1 (representados pelo elemento de matriz γ) eles *não* são degenerados, e seus autovalores são tais que $E_+ - E_- = 2|\gamma|$.

¹V. e.g. a seção “Search and Discovery” de *Physics Today*, Janeiro de 1997, pag. 17.

²De uma forma mais precisa, o termo não diagonal γ que deve ser acrescentado ao problema truncado de um dos dubletos de h_0 não envolve diretamente o elemento de matriz da modificação h_1 apropriada para o problema completo, mas um *hamiltoniano efetivo* \tilde{h}_1 que se anula quando h_1 se anula mas inclui efeitos “intermediários” que podem depender da parte truncada do espaço de fases de h_0 e h_1 . Dessa forma, γ deve ser entendido mais precisamente como um *elemento de matriz efetivo* ligando os dois estados do dubleto, e não apenas como o acoplamento “direto” entre eles através de h_1 .

Um ponto importante a ser notado é que tanto o novo estado fundamental $|E_-\rangle$ como o segundo estado $|E_+\rangle$ são *combinações lineares com coeficientes de mesmo módulo* dos dois estados $|\pm S\hbar\rangle$, que correspondem a *orientações opostas* do spin. Em outras palavras, a orientação do spin fica completamente “deslocalizada” dentro do espaço de dois níveis considerado. Um estado no qual a orientação é “localizada” na direção paralela a z , por exemplo, pode ser construído como

$$|S\hbar\rangle = \frac{1}{\sqrt{2}} \left(|E_-\rangle + \frac{\gamma^*}{|\gamma|} |E_+\rangle \right),$$

mas esse estado *não é estacionário*. A evolução temporal desse estado pode ser calculada facilmente usando o operador de evolução (3.11). A probabilidade de que o spin seja observado na direção *anti-paralela* ao eixo z no tempo t será dada então por

$$p_{S \rightarrow -S}(t) = \sin^2 \left(\frac{E_+ - E_-}{2\hbar} t \right) = \sin^2 \frac{|\gamma|}{\hbar} t$$

e tem portanto um comportamento oscilatório de amplitude 1 e período $\pi\hbar/|\gamma|$, que aumenta à medida que a separação em energia dos dois níveis diminui. Essa mudança de orientação é usualmente chamada “tunelamento” do spin S , devido ao fato de que as orientações intermediárias do spin (que na realidade não intervêm explicitamente nesta descrição) envolvem energias maiores que a energia disponível nos dois estados considerados e são portanto “classicamente proibidas”.

Uma interpretação mais simples, que aliás adere estritamente à maneira de funcionar do procedimento de cálculo utilizado, consiste em considerar a reorientação do spin como um efeito quântico de interferência entre os estados estacionários do sistema de dois níveis. Esses estados estacionários têm, como visto, a propriedade crucial de serem “deslocalizados” quanto à orientação do spin. O estado “tunelante” $|S\hbar\rangle$ constrói a localização direcional do spin através de uma superposição particular dos estados estacionários na qual as componentes correspondentes à orientação segundo a direção oposta à desejada se cancelam por interferência destrutiva. Para que isso aconteça, é preciso ajustar adequadamente as amplitudes e a fase relativa dos estados estacionários ao construir a superposição apropriada. Devido à diferença de energia entre os dois estados estacionários, a *fase relativa* entre as duas componentes do estado tunelante varia com o tempo, com o resultado de que a localização inicial por interferência é progressivamente destruída durante um semi-período $\pi\hbar/2\gamma$ e subseqüentemente *reconstruída* à medida que a variação acumulada da fase relativa se aproxima de 2π . Deste ponto de vista, o tunelamento pode ser visto como resultante de uma perda da localização direcional do spin devida à evolução temporal da fase relativa entre estados estacionários não degenerados e deslocalizados.

O tunelamento do spin S pode ser estudado também na situação mais geral em que existe um campo magnético externo, uniforme e paralelo à direção z , que pode ser descrita acrescentando ao hamiltoniano um terceiro termo

$$h_{\text{ext}} = BS_z.$$

Para que o problema resultante possa ainda ser tratado em termos de um sistema de dois níveis, é preciso que a quebra da degenerescência de h_0 por h_{ext} seja muito menor que o espaçamento em energia dos diferentes dubletos de h_0 . Mais especificamente, no caso do dubleto correspondente ao estado fundamental de h_0 é preciso que

$$\Delta \equiv |2B\hbar S| \ll |A\hbar^2[S^2 - (S-1)^2] = |A\hbar^2(2S-1)|.$$

Nesse caso o modelo de dois níveis a ser considerado envolve o hamiltoniano (3.8) ainda com $\delta_1 = \delta_2 = 0$ mas com $\Delta \neq 0$. Os estados estacionários são dados por (3.9), o “estado tunelante” não estacionário é novamente $|S\hbar\rangle$ e a probabilidade de encontrar o spin no estado $|-S\hbar\rangle$ no tempo t tem novamente um comportamento oscilante dado pela equação (3.12). Neste caso a frequência angular das oscilações da orientação de S é

$$\frac{E_+ - E_-}{\hbar} = \frac{\sqrt{\Delta^2 + 4|\gamma|^2}}{\hbar}$$

e portanto *maior* que no caso em que $\Delta = 0$. A amplitude das oscilações é no entanto agora menor que 1, o que significa, em particular, que a probabilidade de encontrar o spin ainda com a orientação inicial é diferente de zero em qualquer tempo. A situação de amplitude máxima e igual a 1, correspondente a $\Delta = 0$, é usualmente chamada *tunelamento ressonante*.

3.3 A representação de Schrödinger.

Enquanto a versão da mecânica quântica desenvolvida originalmente por Heisenberg confere um papel especial à energia (através da condição de que a matriz que representa a energia seja diagonal), a versão devida a Schrödinger privilegia, de uma forma até certo ponto análoga, as variáveis de posição q . O papel desempenhado na formulação de Heisenberg pela base constituída pelos autovetores do hamiltoniano é, “mutatis mutandis” (o que inclui, em particular, a forma de tratar a dinâmica) substituído por uma “base” de autovetores de q

$$q|x\rangle = x|x\rangle, \quad x \text{ real}$$

os quais, tratando-se de um operador com espectro contínuo (o que significa que qualquer número real pode resultar de uma medida de posição em uma dimensão), são na realidade vetores “impróprios”, cuja ortogonalidade e normalização é definida por $\langle x|x'\rangle = \delta(x - x')$,

onde a última função é a função delta de Dirac. Desse modo, um vetor de estado qualquer $|\varphi\rangle$ é representado³ por uma função complexa da variável real

$$|\varphi\rangle \rightarrow \varphi(x) \equiv \langle x|\varphi\rangle.$$

O produto escalar de dois vetores pode ser traduzido para essa representação fazendo uso da resolução da unidade em termos dos autovetores impróprios $\{|x\rangle\}$

$$\hat{1} = \int_{-\infty}^{+\infty} dx |x\rangle\langle x|$$

que dá

$$\langle\psi|\varphi\rangle = \int_{-\infty}^{+\infty} dx \langle\psi|x\rangle\langle x|\varphi\rangle = \int_{-\infty}^{+\infty} dx \psi^*(x)\varphi(x).$$

O caso particular em que $\psi(x) = \varphi(x)$ dá o quadrado da norma do vetor $|\varphi\rangle$, donde segue que os vetores do espaço de Hilbert são representados por funções de (módulo) quadrado integrável. De acordo com a interpretação padrão da mecânica quântica, a quantidade

$$p_{\Delta} = \int_{x \in \Delta} dx |\langle x|\varphi\rangle|^2 = \int_{x \in \Delta} dx |\varphi(x)|^2$$

corresponde à probabilidade de obter um valor da posição no intervalo Δ numa medição feita sobre o estado $|\varphi\rangle$, de modo que o integrando $|\varphi(x)|^2$ deve ser interpretado como uma *densidade de probabilidade* para os valores da posição no estado $|\varphi\rangle$. As funções $\varphi(x) = \langle x|\varphi\rangle$ que representam vetores do espaço de Hilbert são usualmente chamadas *funções de onda*.

A ação do operador de posição se torna particularmente simples nessa representação. De fato, a representação de Schrödinger do vetor resultante da ação de q sobre $|\varphi\rangle$ pode ser obtida calculando

$$\langle x|q|\varphi\rangle = x\langle x|\varphi\rangle = x\varphi(x)$$

que mostra que a função de onda que representa o vetor resultante é a função de onda que representa o vetor de partida *multiplicada por x* . Em outras palavras, no espaço de funções de onda o operador q é realizado como o operador “multiplicação por x ”.

³É importante não confundir a expressão *representação de Schrödinger* com a *descrição de Schrödinger* da dinâmica quântica, em que a dependência temporal é realizada fazendo o vetor de estado depender do tempo. A expressão *representação de Schrödinger* não se refere a uma particular forma de realizar a dinâmica quântica, mas a uma forma de *representar* os vetores de estado e as variáveis dinâmicas *independentemente* de como venha a ser realizada a dinâmica.

É preciso ainda obter uma representação para o operador associado ao momento, p . Isso significa obter um operador p_S , agindo sobre funções de onda, que seja tal que

$$\langle x|p|\varphi\rangle = p_S\varphi(x)$$

onde p satisfaz a relação de comutação de Heisenberg, Born e Jordan com o operador de posição q

$$[q, p] = i\hbar.$$

Para identificar uma escolha possível para p_S basta notar que, para qualquer vetor $|\varphi\rangle$ (e portanto para qualquer função de onda $\varphi(x)$), deve-se ter

$$\langle x|[q, p]|\varphi\rangle = xp_S\varphi(x) - p_Sx\varphi(x) = i\hbar\varphi(x),$$

donde segue que p_S pode ser tomado como o operador de diferenciação

$$p_S = -i\hbar\frac{\partial}{\partial x}.$$

O caráter hermiteano desse operador no espaço de funções de onda pode ser verificado por meio de uma integração por partes. De fato,

$$\begin{aligned} \int_{-\infty}^{+\infty} dx \psi^*(x) \left[-i\hbar\frac{\partial\varphi(x)}{\partial x} \right] &= -i\hbar\psi^*(x)\varphi(x) \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} dx \left[i\hbar\frac{\partial\psi^*(x)}{\partial x} \right] \varphi(x) \\ &= \left\{ \int_{-\infty}^{+\infty} dx \varphi^*(x) \left[-i\hbar\frac{\partial\psi(x)}{\partial x} \right] \right\}^* \end{aligned}$$

onde o anulamento do termo integrado resulta de que as funções de onda são funções de quadrado integrável.

Dessa forma, a equação de movimento para os vetores de estado, correspondente portanto à descrição de Schrödinger, para o caso de um hamiltoniano que descreve uma partícula de massa M num potencial $V(q)$

$$i\hbar\frac{d|\varphi(t)\rangle}{dt} = \left[\frac{p^2}{2M} + V(q) \right] |\varphi(t)\rangle$$

adquire, no espaço de funções de onda, a forma de uma equação linear de derivadas parciais

$$i\hbar \frac{\partial \varphi(x, t)}{\partial t} = -\frac{\hbar^2}{2M} \frac{\partial^2 \varphi(x, t)}{\partial x^2} + V(x)\varphi(x, t)$$

onde $\varphi(x, t) = \langle x | \varphi(t) \rangle$. Essa equação é na realidade uma equação de movimento para a função de onda que descreve o estado da partícula, a “equação de Schrödinger dependente do tempo”. Para escrever o termo envolvendo o potencial $V(q)$ foi usada a relação $V(q)|x\rangle = V(x)|x\rangle$, que permite escrever o operador $V(q)$ no espaço de funções de onda como o operador “multiplicação por $V(x)$ ”.

A extensão desse esquema para o caso de um espaço de posições tridimensional é simples. Nesse caso, a posição está associada a um conjunto de tres operadores que podem ser convenientemente tomados como as tres projeções cartesianas q_x , q_y e q_z , que será mais conveniente designar por q_1 , q_2 e q_3 , de um *operador vetorial* \vec{q} . É importante observar que o caráter vetorial de \vec{q} se refere ao fato de que as suas componentes cartesianas se transformam como as componentes de um vetor *no espaço tridimensional de posições* por uma mudança do sistema cartesiano adotado. Da mesma forma, o momento está associado a um operador vetorial \vec{p} que pode ser decomposto em tres componentes p_1 , p_2 e p_3 também com propriedades de transformação de componentes de um vetor por uma mudança do sistema cartesiano de referência. Esse conjunto de operadores deve satisfazer as relações de comutação

$$[q_i, q_j] = 0, \quad [p_i, p_j] = 0 \quad \text{e} \quad [q_i, p_j] = i\hbar \delta_{ij}$$

que caracterizam as três componentes cartesianas como *graus de liberdade independentes* (cf. seção 2.2.5). Vale a pena notar que $\{\delta_{ij}\}$, $i, j = 1, 2, 3$ pode ser visto como um tensor cartesiano de segunda ordem, consistentemente com o caráter vetorial de $\{q_i\}$, $i = 1, 2, 3$ e $\{p_j\}$, $j = 1, 2, 3$. Desse modo, toda a discussão anterior do caso unidimensional se aplica a cada uma das componentes cartesianas considerada separadamente. Os autovetores simultâneos das tres componentes de \vec{q} (impróprios) são designados pela notação $|\vec{r}\rangle$. Dessa forma, escrevendo as componentes de \vec{r} como x_1 , x_2 e x_3 ,

$$q_i |\vec{r}\rangle = x_i |\vec{r}\rangle;$$

a função de onda que representa o vetor de estado $|\varphi\rangle$ é agora a função complexa de tres variáveis reais

$$\varphi(\vec{r}) = \langle \vec{r} | \varphi \rangle$$

e o vetor resultante da ação de q_i sobre $|\varphi\rangle$ é representado por $\langle \vec{r} | q_i | \varphi \rangle = x_i \varphi(\vec{r})$. Da mesma forma a ação das componentes do momento é representada em termos das funções de onda pelos operadores diferenciais

$$\langle \vec{r} | p_i | \varphi \rangle = -i\hbar \frac{\partial \varphi(\vec{r})}{\partial x_i}.$$

O caráter de vetores no espaço tridimensional de posições dos operadores \vec{q} e \vec{p} permite escrever essas relações de uma forma que não faz referência explícita a um sistema de coordenadas determinado, que é

$$\langle \vec{r} | \vec{q} | \varphi \rangle = \vec{r} \varphi(\vec{r}) \quad \text{e} \quad \langle \vec{r} | \vec{p} | \varphi \rangle = -i\hbar \nabla \varphi(\vec{r})$$

onde ∇ é o operador diferencial associado ao gradiente. Nessas expressões, $\vec{r} \varphi(\vec{r})$ e $-i\hbar \nabla \varphi(\vec{r})$ devem ser entendidos como *vetores* no espaço tridimensional, que podem portanto ser escritos em termos de tres componentes (complexas), associadas às componentes respectivas dos operadores vetoriais envolvidos.

Finalmente, a equação de Schrödinger dependente do tempo que descreve a dinâmica de uma partícula de massa M em um potencial $V(\vec{q})$ tem a forma

$$i\hbar \frac{\partial \varphi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2M} \nabla^2 \varphi(\vec{r}, t) + V(\vec{r}) \varphi(\vec{r}, t) \quad (3.14)$$

em que é importante notar o caráter *escalar* do operador diferencial associado à energia cinética, representado pelo operador laplaciano ∇^2 . Essa equação de movimento é uma equação de derivadas parciais linear para quatro variáveis. Sendo uma equação de primeira ordem no tempo, o problema de condições iniciais em que $\varphi(\vec{r}, t = 0)$ é dada determina univocamente a solução para outros valores de t .

3.3.1 Propriedades gerais da dinâmica de uma partícula em um potencial externo.

A propriedade mais geral da dinâmica quântica é a conservação de norma do vetor de estado. Ela decorre, como foi visto, da unitariedade do operador de evolução $U(t, t_0)$, por sua vez ligada à hermiticidade do hamiltoniano. Na descrição de Schrödinger (vetor de estado dependente do tempo, variáveis dinâmicas independentes do tempo) e na representação de Schrödinger (vetores de estado tratados em termos de funções de onda) a equação de movimento é uma equação diferencial também nas variáveis espaciais. A conservação de norma pode então ser expressa, em forma diferencial, por uma equação de continuidade para a densidade de probabilidade dependente do tempo

$$\rho(\vec{r}, t) \equiv |\varphi(\vec{r}, t)|^2.$$

De fato, diferenciando $\rho(\vec{r}, t)$ com relação ao tempo e usando a equação de movimento (3.14) resulta

$$\begin{aligned}\frac{\partial \rho(\vec{r}, t)}{\partial t} &= \varphi^*(\vec{r}, t) \frac{\partial \varphi(\vec{r}, t)}{\partial t} + \frac{\partial \varphi^*(\vec{r}, t)}{\partial t} \varphi(\vec{r}, t) \\ &= \frac{1}{i\hbar} [\varphi^*(\vec{r}, t) H \varphi(\vec{r}, t) - \varphi(\vec{r}, t) (H \varphi^*(\vec{r}, t))] \\ &= -\frac{1}{i\hbar} \frac{\hbar^2}{2M} [\varphi^*(\vec{r}, t) \nabla^2 \varphi(\vec{r}, t) - \varphi(\vec{r}, t) (\nabla^2 \varphi^*(\vec{r}, t))]\end{aligned}$$

onde foi usada a hermiticidade de H e, na última passagem, a realidade de $V(\vec{r})$. A partir da última expressão acima é fácil ver finalmente que

$$\frac{\partial \rho(\vec{r}, t)}{\partial t} + \nabla \cdot \vec{j}(\vec{r}, t) = 0$$

com a definição da *densidade de corrente de probabilidade*

$$\vec{j}(\vec{r}, t) = \frac{\hbar}{2Mi} [\varphi^*(\vec{r}, t) \nabla \varphi(\vec{r}, t) - \varphi(\vec{r}, t) (\nabla \varphi^*(\vec{r}, t))]. \quad (3.15)$$

A densidade de probabilidade $\rho(\vec{r}, t)$ é claramente independente da fase da função de onda, mas é interessante analisar o papel essencial que essa fase desempenha para a densidade de corrente $\vec{j}(\vec{r}, t)$. Para isso é conveniente separar explicitamente a amplitude e a fase da função de onda $\varphi(\vec{r}, t)$ escrevendo-a sob a forma

$$\varphi(\vec{r}, t) = A(\vec{r}, t) e^{\frac{i}{\hbar} S(\vec{r}, t)}$$

onde $A(\vec{r}, t)$ e $S(\vec{r}, t)$ são funções *reais* e a constante \hbar foi introduzida na definição de S por conveniência, dando a essa função dimensões de ação. Substituindo essa forma de $\varphi(\vec{r}, t)$ na equação de movimento (3.14) e igualando a zero separadamente as partes real e imaginária da equação resultante se obtém, respectivamente

$$A(\vec{r}, t) \left[\frac{\partial S(\vec{r}, t)}{\partial t} + \frac{(\nabla S(\vec{r}, t))^2}{2M} + V(\vec{r}) \right] = \frac{\hbar^2}{2M} \nabla^2 A(\vec{r}, t) \quad (3.16)$$

e

$$\frac{\partial A(\vec{r}, t)}{\partial t} = -\frac{1}{2M} A(\vec{r}, t) \nabla^2 S(\vec{r}, t) - \frac{1}{M} \nabla A(\vec{r}, t) \cdot \nabla S(\vec{r}, t). \quad (3.17)$$

A equação (3.17), correspondente à parte imaginária, pode ser facilmente reduzida à equação de continuidade. Multiplicando-a, de fato, por $2A(\vec{r}, t)$ ela pode ser reescrita sob a forma

$$\frac{\partial A^2(\vec{r}, t)}{\partial t} + \nabla \cdot \left[A^2(\vec{r}, t) \frac{\nabla S(\vec{r}, t)}{M} \right] = 0.$$

Nessa equação $A^2(\vec{r}, t)$ é evidentemente a densidade de probabilidade. Daí decorre que a densidade de corrente é dada por

$$\vec{j}(\vec{r}, t) = A^2(\vec{r}, t) \frac{\nabla S(\vec{r}, t)}{M} = \rho(\vec{r}, t) \frac{\nabla S(\vec{r}, t)}{M} \quad (3.18)$$

que identifica $-S(\vec{r}, t)/M$ como um *potencial de velocidades* para o fluxo da densidade de probabilidade prescrito pela equação de movimento, no sentido que $\vec{j}(\vec{r}, t) = \rho(\vec{r}, t) \vec{v}(\vec{r}, t)$ onde o campo de velocidades $\vec{v}(\vec{r}, t)$ é dado como

$$\vec{v}(\vec{r}, t) = -\nabla \left(\frac{-S(\vec{r}, t)}{M} \right).$$

Por outro lado, as duas funções reais $A(\vec{r}, t)$ e $S(\vec{r}, t)$ estão ligadas em sua evolução temporal pela equação (3.16), proveniente dos termos reais da equação de movimento. Essa equação pode ser reescrita sob a forma

$$\frac{\partial S(\vec{r}, t)}{\partial t} + \frac{(\nabla S(\vec{r}, t))^2}{2M} + V(\vec{r}) - \frac{\hbar^2}{2M} \frac{\nabla^2 A(\vec{r}, t)}{A(\vec{r}, t)} = 0 \quad (3.19)$$

que, a menos do último termo, corresponde à equação clássica de Hamilton-Jacobi, em que $S(\vec{r}, t)$ é a função principal de Hamilton, para o problema de uma partícula de massa M sujeita ao potencial $V(\vec{r})$. Essa equação pode portanto ser usada para identificar um *limite clássico* do tratamento quântico do problema de uma partícula em um potencial externo. De fato, sempre que os efeitos do último termo puderem ser ignorados quantitativamente na determinação de $S(\vec{r}, t)$ pela equação (3.19), essa função será tal que estarão satisfeitas as relações clássicas

$$\nabla S(\vec{r}, t) = \vec{p}_c(\vec{r}, t) = M \vec{v}_c(\vec{r}, t)$$

onde p_c e \vec{v}_c são respectivamente um momento e uma velocidade clássica. Se a energia total tiver um valor bem definido vale ainda a relação

$$\frac{\partial S(\vec{r}, t)}{\partial t} = -E.$$

Nesse limite, as linhas de velocidade associadas à densidade de corrente de probabilidade se reduzem às trajetórias clássicas. É importante notar que a descrição do sistema permanece

como sendo feita em termos de uma distribuição contínua de probabilidades, de modo que a situação clássica a que este limite se refere deve ser entendida como sendo realizada em termos de propriedades médias em uma coleção de sistemas idênticos com uma distribuição de condições iniciais dada. As condições em que um tal limite pode ser usado como aproximação para o problema quântico podem ser caracterizadas da seguinte forma. Supondo que a energia tenha uma dispersão pequena em torno de E , a equação (3.19) pode ser aproximada por

$$(\nabla S(\vec{r}, t))^2 \simeq 2M(E - V(\vec{r})) + \frac{\hbar^2}{2M} \frac{\nabla^2 A(\vec{r}, t)}{A(\vec{r}, t)}.$$

Uma “aproximação clássica”, consistindo em ignorar o último termo, será válida quando

$$\frac{\nabla^2 A(\vec{r}, t)}{A(\vec{r}, t)} \ll \frac{2M}{\hbar^2} (E - V(\vec{r})) \equiv \left(\frac{2\pi}{\lambda(\vec{r})} \right)^2$$

onde $\lambda(\vec{r})$ é o chamado *comprimento de onda de de Broglie*. Essa condição mostra, em termos qualitativos, que efeitos de curvatura na dependência espacial da amplitude $A(\vec{r}, t)$ devem ser pequenos na escala do comprimento de onda de de Broglie.

Esse tratamento das condições de validade para o limite clássico da dinâmica quântica na representação de Schrödinger contém na realidade uma analogia óptica freqüentemente usada. De fato, considerando agora estados de energia total bem definida E , a equação para $S(\vec{r})$ pode ser escrita, no limite clássico, como

$$\left[\nabla \left(\frac{\lambda_0}{2\pi\hbar} S(\vec{r}) \right) \right]^2 = \frac{\lambda_0^2}{\lambda^2(\vec{r})} \equiv n^2(\vec{r})$$

onde λ_0 é um “valor de referência” escolhido para o comprimento de onda de de Broglie (por exemplo, $\lambda_0^2 = (2\pi\hbar)^2/2ME$). Essa equação diferencial é conhecida como “equação do Eikonal” (que, no caso, é a função $\lambda_0 S(\vec{r})/2\pi\hbar$). A quantidade adimensional $n(\vec{r})$ tem o sentido de um índice de refração relativo no ponto \vec{r} . As superfícies $S = \text{constante}$ são superfícies de fase constante, e as normais a essas superfícies (que correspondem às linhas de velocidade da mecânica clássica) são os “raios” da óptica geométrica. Dessa forma, variações do potencial $V(\vec{r})$ podem ser interpretadas em termos de variações de um índice de refração.

Outro resultado geral interessante para o movimento de uma partícula em um potencial é o chamado Teorema de Ehrenfest, que dá uma expressão para a variação no tempo do valor esperado do momento

$$\langle \vec{p} \rangle(t) = \int \varphi^*(\vec{r}, t) \vec{p} \varphi(\vec{r}, t) d^3r$$

onde as funções de onda são supostamente normalizadas. A derivada temporal desse valor médio pode ser obtida usando a equação de movimento de Schrödinger

$$\begin{aligned}\frac{d}{dt}\langle\vec{p}\rangle(t) &= \frac{\hbar}{i} \int \left[\frac{\partial\varphi^*}{\partial t} \nabla\varphi + \varphi^* \nabla \frac{\partial\varphi}{\partial t} \right] d^3r \\ &= -\frac{\hbar^2}{2M} \int \left[\nabla^2\varphi^* \nabla\varphi - \varphi^* \nabla \nabla^2\varphi \right] d^3r - \int \varphi^* (\nabla V) \varphi d^3r.\end{aligned}$$

A primeira integral se anula, como é fácil verificar fazendo uma dupla integração por partes, e a segunda é o valor médio do gradiente do potencial. Portanto

$$\frac{d}{dt}\langle\vec{p}\rangle(t) = -\langle\nabla V\rangle(t).$$

A dependência temporal nos dois termos provém da dependência temporal da função de onda, determinada pela equação de Schrödinger dependente do tempo. Esse resultado é *exato*, e não está ligado a aproximações clássicas. É importante notar, porém, que a “força” que age sobre o valor médio do momento é a média da força clássica no estado quântico considerado, que em geral difere substancialmente da força clássica calculada, por exemplo, na posição média $\langle\vec{r}\rangle(t)$. Apenas nos casos em que a partícula sujeita a um potencial linear (força constante, possivelmente nula, $\nabla V = \text{constante}$) e de oscilador harmônico ($\nabla V = k\vec{r}$) é possível escrever

$$-\langle\nabla V\rangle(t) = -\nabla V(\langle\vec{r}\rangle(t)),$$

isto é, que a força que age sobre o momento médio da partícula é a força clássica correspondente à posição média da partícula.

3.3.2 Funções de onda estacionárias para o oscilador harmônico.

As funções de onda $\phi_\nu(x) \equiv \langle x|\nu\rangle$ correspondentes aos autovetores do hamiltoniano de um oscilador harmônico unidimensional podem ser obtidas como as soluções de quadrado integrável da equação diferencial de autovalores

$$-\frac{\hbar^2}{2M} \frac{d^2\phi_\nu(x)}{dx^2} + \frac{M\omega_0^2}{2} x^2 \phi_\nu(x) = E_\nu \phi_\nu(x).$$

Alternativamente, essas funções de onda estacionárias podem ser calculadas diretamente a partir dos autovetores obtidos de forma “abstrata” na seção **3.1.1** através da introdução da representação de Schrödinger. Para simplificar as expressões envolvidas, é conveniente introduzir inicialmente as versões adimensionais dos operadores associados à posição e ao momento

$$Q = \sqrt{\frac{M\omega_0}{\hbar}} q = \frac{q}{b} \quad \text{e} \quad P = \frac{1}{\sqrt{\hbar M\omega_0}} p = \frac{bp}{\hbar}$$

obtidas introduzindo convenientemente a escala de comprimento dada pelo chamado “parâmetro do oscilador” $b = \sqrt{\hbar/M\omega_0}$. Como é imediato verificar a partir das relações de comutação satisfeitas por q e p , satisfazem à relação de comutação $[Q, P] = i$, e numa representação de Schrödinger em que o operador Q é realizado como operador de multiplicação pela posição adimensional $X = x/b$, o operador P pode ser realizado em termos do operador diferencial $-i\partial/\partial X$.

Em termos desses operadores, o operador a , que aniquila o estado fundamental do oscilador harmônico, isto é, $a|\nu=0\rangle=0$ (v. seção 3.1.1), pode ser escrito como

$$a = \frac{1}{\sqrt{2}}(Q + iP) \rightarrow \frac{1}{\sqrt{2}}\left(X + \frac{\partial}{\partial X}\right).$$

Dessa forma, se $\phi_0(X) = \langle X|\phi_0\rangle$ é a função de onda estacionária correspondente ao estado fundamental $|\nu=0\rangle$, segue que, nessa representação de Schrödinger ela deve satisfazer a equação diferencial

$$\frac{1}{\sqrt{2}}\left(X + \frac{\partial}{\partial X}\right)\phi_0(X) = 0 \quad \text{ou} \quad \frac{\partial\phi_0(X)}{\partial X} = -X\phi_0(X),$$

cujas soluções são facilmente reconhecíveis como

$$\phi_0(X) = Ne^{-\frac{X^2}{2}}$$

onde N é uma constante de normalização a ser escolhida de forma conveniente. Para isso, o que é feito usualmente envolve uma volta à variável de posição $x = bX$ e o uso da condição de normalização de que a integral *em* x do quadrado de ϕ_0 seja igual a 1. Disso resulta

$$\phi_0(x) = \left[\frac{1}{b\sqrt{\pi}}\right]^{\frac{1}{2}} e^{-\frac{x^2}{2b^2}}$$

que corresponde a uma densidade de probabilidade gaussiana, de variância b , em torno da origem $x=0$.

A função de onda $\phi_\nu(x)$ correspondente ao ν -ésimo estado excitado $|\nu\rangle$ do oscilador pode ser obtida facilmente da função de onda que representa o estado fundamental por aplicação do operador $a^{\dagger\nu}/\sqrt{\nu!}$. Esse operador é representado por

$$\frac{1}{\sqrt{\nu!}} a^{\dagger \nu} \rightarrow \frac{1}{\sqrt{2^\nu \nu!}} \left(X - \frac{\partial}{\partial X} \right)^\nu$$

de modo que

$$\phi_\nu(X) = N \frac{1}{\sqrt{2^\nu \nu!}} \left(X - \frac{\partial}{\partial X} \right)^\nu e^{-\frac{X^2}{2}}. \quad (3.20)$$

A menos dos coeficientes de normalização apropriados, essas funções tem portanto a forma de produtos da gaussiana que aparece no estado fundamental por polinômios em X que podem ser escritos como

$$H_\nu(X) \equiv e^{\frac{X^2}{2}} \left(X - \frac{\partial}{\partial X} \right)^\nu e^{-\frac{X^2}{2}} \quad (3.21)$$

que são os chamados *polinômios de Hermite*. As funções normalizadas em termos da variável $x = bX$ são dadas pela expressão (3.20) com a constante de normalização N obtida para o estado fundamental, e são portanto expressas em termos dessa variável como

$$\phi_\nu(x) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} \frac{1}{\sqrt{2^\nu \nu!}} H_\nu \left(\frac{x}{b} \right) e^{-\frac{x^2}{2b^2}}.$$

O uso dessas funções de onda para o cálculo de elementos de matriz na base dos estados estacionários do oscilador harmônico envolve integrais contendo produtos de polinômios de Hermite no integrando. Integrais desse tipo podem ser convenientemente calculadas utilizando a *função geratriz dos polinômios de Hermite*, definida como a soma da série

$$G(z, X) = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} H_\nu(X).$$

Essa função geratriz pode ser obtida em forma fechada notando que ela é a solução da equação diferencial

$$\frac{\partial G(z, X)}{\partial z} = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} H_{\nu+1}(X) = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \left(2XH_\nu(X) - \frac{dH_\nu(X)}{dX} \right) = 2XG(z, X) - \frac{\partial G(z, X)}{\partial X}$$

com a condição de contorno $G(0, X) = H_0(X) = 1$, que é dada simplesmente como

$$G(z, X) = e^{-z^2 + 2zX}.$$

Na obtenção da equação diferencial para a função geratriz foi feito uso de uma *formula de derivação* para os polinômios de Hermite que é facilmente obtida da definição (3.21). De fato,

$$\begin{aligned} H_{\nu+1}(X) &= e^{\frac{X^2}{2}} \left(X - \frac{\partial}{\partial X} \right) e^{-\frac{X^2}{2}} e^{\frac{X^2}{2}} \left(X - \frac{\partial}{\partial X} \right)^{\nu} e^{-\frac{X^2}{2}} \\ &= e^{\frac{X^2}{2}} \left(X - \frac{\partial}{\partial X} \right) e^{-\frac{X^2}{2}} H_{\nu}(X) = 2XH_{\nu}(X) - \frac{dH_{\nu}(X)}{dX}, \end{aligned}$$

que pode ser vista como fornecendo uma expressão para a derivada de $H_{\nu}(X)$ em termos de $H_{\nu}(X)$ e de $H_{\nu+1}(X)$.

Como exemplo do uso da função geratriz para o cálculo de integrais envolvendo polinômios de Hermite, seja o problema de determinar a solução da equação de Schrödinger dependente do tempo para o oscilador harmônico

$$-\frac{\hbar^2}{2M} \frac{\partial^2 \varphi(x, t)}{\partial x^2} + \frac{M\omega_0^2}{2} x^2 \varphi(x, t) = i\hbar \frac{\partial \varphi(x, t)}{\partial t}$$

com a condição inicial

$$\varphi(x, 0) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} e^{-\frac{(x-x_0)^2}{2b^2}} \quad (3.22)$$

que é uma função de onda semelhante à que corresponde ao estado fundamental, mas deslocada da origem de uma distância x_0 . O procedimento padrão para obter a solução buscada através das autofunções do hamiltoniano é resumido na expressão

$$\varphi(x, t) = \sum_{\nu=0}^{\infty} \phi_{\nu}(x) e^{-i(n+\frac{1}{2})\omega_0 t} \int dx' \phi_{\nu}^*(x') \varphi(x', 0)$$

que pode ser descrita como o resultado da aplicação do operador de evolução $\exp(-iHt/\hbar)$ sobre a expansão da condição inicial em termos das autofunções de H . Os coeficientes dessa expansão são obtidos através das integrais sobre x' , que podem ser escritas explicitamente no caso em questão como

$$c_{\nu} \equiv \frac{1}{b\sqrt{\pi}} \frac{1}{\sqrt{2^{\nu}\nu!}} \int dx' H_{\nu}\left(\frac{x'}{b}\right) e^{-\frac{x'^2}{2b^2} - \frac{(x'-x_0)^2}{2b^2}}.$$

Essa integral envolve um polinômio de Hermite e pode ser calculada substituindo-o pela função geratriz. Com isso o resultado aparece como uma função de z cuja expansão em série de

potências permite identificar a integral de partida como o coeficiente de $z^\nu/\nu!$ nessa expansão. Utilizando, por simplicidade de notação, a variável adimensional $X = x/b$, a nova integral a ser calculada é então

$$\frac{1}{\sqrt{\pi}2^\nu\nu!} \int dX e^{-z^2+2zX} e^{-\frac{X^2}{2}-\frac{(X-X_0)^2}{2}} = \frac{1}{\sqrt{2^\nu\nu!}} e^{zX_0-\frac{X_0^2}{4}}$$

onde $X_0 = x_0/b$. Da expansão da exponencial envolvendo z em série de potências resulta finalmente

$$c_\nu = \frac{X_0^\nu e^{-\frac{X_0^2}{4}}}{\sqrt{2^\nu\nu!}}$$

de modo que a solução desse problema de condições iniciais é dada pela série

$$\varphi(X, t) = \sum_{\nu=0}^{\infty} \phi_\nu(bX) e^{-i(n+\frac{1}{2})\omega_0 t} \frac{X_0^\nu e^{-\frac{X_0^2}{4}}}{\sqrt{2^\nu\nu!}} = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} \sum_{\nu=0}^{\infty} e^{-i(n+\frac{1}{2})\omega_0 t} H_\nu(X) e^{-\frac{X^2}{2}} \frac{X_0^\nu e^{-\frac{X_0^2}{4}}}{2^\nu\nu!}.$$

É fácil ver ainda que a soma dessa série pode ser obtida em forma fechada utilizando mais uma vez a função geratriz dos polinômios de Hermite. De fato, colecionando os ingredientes independentes de ν , ela pode ser reescrita como

$$\varphi(X, t) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} e^{-i\frac{\omega_0 t}{2}-\frac{X_0^2}{2}-\frac{X^2}{2}} \sum_{\nu=0}^{\infty} \left(\frac{X_0 e^{-i\omega_0 t}}{2} \right)^\nu \frac{H_\nu(X)}{\nu!}$$

onde a soma pode ser reconhecida como a função geratriz calculada para $z = X_0 e^{-i\omega_0 t}/2$

$$G\left(\frac{X_0 e^{-i\omega_0 t}}{2}, X\right) = \exp\left(-\frac{X_0^2}{4} e^{-2i\omega_0 t} + X_0 X e^{-i\omega_0 t}\right).$$

Usando esse resultado, a solução $\varphi(x, t)$ pode ser facilmente reduzida à forma fechada

$$\varphi(x, t) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} e^{-i\left(\frac{\omega_0 t}{2} + \frac{xx_0}{b^2} \sin \omega_0 t - \frac{x_0^2}{b^2} \sin 2\omega_0 t\right)} e^{-\frac{1}{2b^2}(x-x_0 \cos \omega_0 t)^2}.$$

A densidade de probabilidade no tempo t correspondente a essa solução é dada simplesmente por

$$\rho(x, t) = |\varphi(x, t)|^2 = \frac{1}{b\sqrt{\pi}} e^{-\frac{1}{b^2}(x-x_0 \cos \omega_0 t)^2}$$

que tem a mesma forma que $|\varphi(x, 0)|^2$ mas oscila harmonicamente, com a frequência do oscilador, em torno de $x = 0$ com amplitude x_0 . A parte imaginária do expoente identifica o potencial do campo de velocidades associado ao fluxo de probabilidade, como discutido na seção anterior, equação (3.18). Neste caso o campo de velocidades é espacialmente uniforme e dado por

$$\vec{v} = \frac{\hbar x_0}{b^2} \text{sen } \omega_0 t.$$

O fato de ele ser espacialmente uniforme é consistente com a preservação da forma da distribuição de probabilidades ao longo do tempo, e desempenha um papel essencial na determinação do valor médio do momento como função do tempo, que também oscila em torno de zero com a frequência ω_0 .

Essas características são as que correspondem à evolução temporal de um *estado coerente* do oscilador (v. Problema 3.6), e é fácil verificar que a condição inicial (3.22) é de fato um estado coerente. Para isso basta calcular

$$\frac{1}{\sqrt{2}}(Q - iP)\varphi(x, 0) = \frac{1}{\sqrt{2}} \left(X + \frac{d}{dX} \right) \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} e^{-\frac{(X-X_0)^2}{2}} = \frac{X_0}{\sqrt{2}} \varphi(x, 0)$$

que mostra que $\varphi(x, 0)$ é uma autofunção do operador que corresponde, na representação de Schrödinger, ao operador a , com autovalor $x_0/\sqrt{2}$. Da mesma forma se pode verificar que a solução $\varphi(x, t)$ é também um estado coerente (com autovalor em geral complexo) para qualquer t .

3.3.3 Comportamento quântico de uma partícula livre.

A dinâmica de uma partícula livre é caracterizada pelo hamiltoniano

$$H = \frac{\vec{p}^2}{2M}$$

que depende apenas do momento \vec{p} , e portanto pode ser diagonalizada simultaneamente com \vec{p} . A equação de autovalores para esta variável dinâmica é

$$\vec{p}|\vec{k}\rangle = \hbar\vec{k}|\vec{k}\rangle$$

que, na representação de Schrödinger, escrevendo a função de onda $\langle\vec{r}|\vec{k}\rangle$ como $\phi_{\vec{k}}(\vec{r})$, aparece sob a forma

$$-i\hbar\nabla\phi_{\vec{k}}(\vec{r}) = \hbar\vec{k}\phi_{\vec{k}}(\vec{r}).$$

O autovalor foi escrito, por conveniência, como $\hbar\vec{k}$, de modo que as componentes cartesianas do vetor \vec{k} têm dimensão de inverso de comprimento. Essa equação admite como soluções as "ondas planas"

$$\phi_{\vec{k}}(\vec{r}) = \langle \vec{r} | \vec{k} \rangle = N_{\vec{k}} e^{i\vec{k} \cdot \vec{r}} \quad (3.23)$$

que devem ser ainda convenientemente normalizadas através da escolha da constante $N_{\vec{k}}$.

Essas funções não têm norma finita, e de fato correspondem, na representação de Schrödinger, aos autovetores impróprios $|\vec{k}\rangle$ do operador com espectro puramente contínuo \vec{p} , isto é, $\phi_{\vec{k}}(\vec{r}) = \langle \vec{r} | \vec{k} \rangle$. As ondas planas $\phi_{\vec{k}}(\vec{r})$ podem portanto ser vistas, a menos da questão de normalização, também como elementos de uma transformação unitária que leva da representação de Schrödinger, na qual o operador \vec{q} é realizado como um operador de multiplicação (sendo portanto diagonal), para uma *representação de momento*, na qual o operador \vec{p} é aquele que desempenha papel análogo ao do operador \vec{q} na representação de Schrödinger. De fato, a "função de onda no espaço de momento" associada a um vetor de estado qualquer $|\varphi\rangle$ corresponde à amplitude

$$\tilde{\varphi}(\vec{k}) \equiv \langle \vec{k} | \varphi \rangle = \int d^3r \langle \vec{k} | \vec{r} \rangle \langle \vec{r} | \varphi \rangle = N_{\vec{k}} \int d^3r e^{-i\vec{k} \cdot \vec{r}} \varphi(\vec{r}), \quad (3.24)$$

o que mostra que essa transformação consiste de fato numa *transformação de Fourier* da função de onda de Schrödinger. Para assegurar a sua unitariedade formal, no sentido que os elementos da transformação inversa sejam os complexo conjugados dos elementos da transformação direta transposta, a escolha apropriada da constante de normalização é, então,

$$N_{\vec{k}} = \frac{1}{(2\pi)^{3/2}}. \quad (3.25)$$

Essa escolha permite escrever a transformação inversa (do espaço de momento para a representação de Schrödinger) como

$$\varphi(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int d^3k e^{i\vec{k} \cdot \vec{r}} \tilde{\varphi}(\vec{k}),$$

que, combinada com a (3.24), corresponde a uma resolução formal da unidade escrita como

$$\delta(\vec{r} - \vec{r}') = \frac{1}{(2\pi)^3} \int d^3k e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}. \quad (3.26)$$

A validade dessa expressão formal está ligada à validade da inversão *ponto a ponto* da transformação de Fourier.

As funções (3.23) (com a normalização (3.25)) são também soluções da equação de autovalores para o hamiltoniano da partícula livre

$$-\frac{\hbar^2}{2M}\nabla^2\phi_{\vec{k}}(\vec{r}) = E_k\phi_{\vec{k}}(\vec{r})$$

em que os autovalores são dados por

$$E_k = \frac{\hbar^2 k^2}{2M}.$$

Ao contrário do que ocorre com o espectro do momento, os autovalores do espectro contínuo de H são infinitamente degenerados. De fato, eles dependem apenas do módulo do vetor \vec{k} , o que implica (para cada $k \neq 0$) numa degenerescência também *contínua*, correspondente às diferentes orientações possíveis de \vec{k} .

O problema da dinâmica da partícula livre, utilizando a descrição de Schrödinger formulada em termos de funções de onda dependentes do tempo, consiste em obter soluções da equação de movimento

$$-\frac{\hbar^2}{2M}\nabla^2\varphi(\vec{r}, t) = i\hbar\frac{\partial\varphi(\vec{r}, t)}{\partial t}$$

satisfazendo condições iniciais $\varphi(\vec{r}, t = 0)$ dadas. Tais soluções podem ser obtidas através da representação de Schrödinger do operador de evolução, escrita em termos das autofunções simultâneas de \vec{p} e H ,

$$U(\vec{r}, t; \vec{r}', 0) \equiv \langle \vec{r} | e^{-\frac{i}{\hbar} H t} | \vec{r}' \rangle = \int d^3k \langle \vec{r} | \vec{k} \rangle e^{-i\frac{\hbar k^2}{2M}t} \langle \vec{k} | \vec{r}' \rangle = \frac{1}{(2\pi)^3} \int d^3k e^{i\vec{k} \cdot (\vec{r} - \vec{r}') - i\frac{\hbar k^2}{2M}t} \quad (3.27)$$

De fato, a solução do problema de condições iniciais para a partícula livre é dada em termos desse operador simplesmente como

$$\varphi(\vec{r}, t) = \langle \vec{r} | e^{-\frac{i}{\hbar} H t} | \varphi(t = 0) \rangle = \int d^3r' U(\vec{r}, t; \vec{r}', 0) \varphi(\vec{r}', 0). \quad (3.28)$$

Na realidade, existem duas formas de implementar esse esquema em situações concretas. A primeira segue passo a passo o procedimento usado na seção anterior para resolver o problema análogo para o oscilador harmônico. De fato, a integração sobre \vec{r}' em (3.28) corresponde a aplicar uma transformação de Fourier à condição inicial, o que efetivamente a reduz à representação em que o hamiltoniano é diagonal. Desse modo a solução desejada é obtida calculando

$$\varphi(\vec{r}, t) = \frac{1}{(2\pi)^{3/2}} \int d^3k e^{i\vec{k} \cdot \vec{r} - i\frac{\hbar k^2}{2M}t} \tilde{\varphi}(\vec{k}, 0) \quad (3.29)$$

onde $\tilde{\varphi}(\vec{k}, 0)$ é a transformada de Fourier da condição inicial. Alternativamente, é possível fazer a integração sobre \vec{k} na equação (3.27) *independentemente da condição inicial*, e depois usar o resultado como indicado em (3.28). Esta última estratégia tem a vantagem de que sua primeira etapa pode ser implementada de uma vez por todas, independentemente da condição inicial, sendo esta introduzida posteriormente, através da integração sobre \vec{r}' em (3.28). O que se obtém desse modo para $U(\vec{r}, t; \vec{r}', 0)$ é (v. problema 3.13)

$$U(\vec{r}, t; \vec{r}', 0) = \left(\frac{M}{2\pi i \hbar t} \right)^{\frac{3}{2}} e^{\frac{iM(\vec{r}-\vec{r}')^2}{2\hbar t}}. \quad (3.30)$$

Esse objeto é a representação de Schrödinger do operador de evolução (também eventualmente chamado *propagador de Schrödinger* para a partícula livre), e consiste na realidade de um "elemento de matriz" desse operador, tomado com os autovetores impróprios do operador de posição, que pode ser interpretado como a amplitude da probabilidade de que uma partícula livre localizada na posição \vec{r}' em $t = 0$ seja encontrada na posição \vec{r} no tempo t . A relação (3.28) pode então ser vista como uma consequência imediata do caráter linear da equação de movimento quântica. As duas estratégias de cálculo para $\varphi(\vec{r}, t)$ são evidentemente equivalentes, e a preferência por uma ou outra deve ser decidida por eventuais razões de conveniência.

Um exemplo que envolve cálculos relativamente simples, mas que é bastante útil para ilustrar várias propriedades da dinâmica quântica de uma partícula livre, é o de uma condição inicial que consiste de um estado descrito pela função de onda gaussiana normalizada (usualmente chamada "pacote de ondas gaussiano")

$$\varphi(\vec{r}, 0) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{3}{2}} e^{i\vec{k}_0 \cdot \vec{r}} e^{-\frac{(\vec{r}-\vec{r}_0)^2}{2b^2}}.$$

Os valores médios da posição e do momento nesse estado são facilmente obtidos como

$$\langle \vec{q} \rangle = \vec{r}_0 \quad \text{e} \quad \langle \vec{p} \rangle = \hbar \vec{k}_0$$

e a densidade de corrente (3.18) envolve um campo de velocidades uniforme $\vec{v}_0 = \hbar \vec{k}_0 / M$. A transformada de Fourier, que corresponde à função de onda no espaço de momento $\langle \vec{k} | \varphi(t=0) \rangle$ do estado tomado como condição inicial, é

$$\tilde{\varphi}(\vec{k}, 0) = \left[\frac{2b}{\sqrt{\pi}} \right]^{\frac{3}{2}} e^{-i(\vec{k}-\vec{k}_0) \cdot \vec{r}_0} e^{-2b^2(\vec{k}-\vec{k}_0)^2}$$

que mostra que a distribuição de probabilidade desse estado no espaço de momento é também gaussiana, com largura $1/2b$ em termos dos vetores de onda \vec{k} . A condição inicial é portanto um estado de incerteza mínima.

Uma primeira forma de chegar à função de onda no tempo t é dada pela equação (3.29), na qual aparece a transformada de Fourier da condição inicial. Essa forma leva, no caso que está sendo considerado, a uma expressão envolvendo uma integração sobre o espaço de momento

$$\varphi(\vec{r}, t) = \frac{1}{(2\pi)^{3/2}} \left[\frac{2b}{\sqrt{\pi}} \right]^{\frac{3}{2}} \int d^3k \, e^{i\vec{k} \cdot \vec{r} - i \frac{\hbar k^2}{2M} t} e^{-i(\vec{k} - \vec{k}_0) \cdot \vec{r}_0} e^{-2b^2(\vec{k} - \vec{k}_0)^2}$$

que pode ser calculada em termos de tres integrais gaussianas envolvendo as componentes cartesianas de \vec{k} . Alternativamente, $\varphi(\vec{r}, t)$ pode ser obtida usando as equações (3.28) e (3.30), o que leva a uma expressão que envolve uma integral em \vec{r}'

$$\varphi(\vec{r}, t) = \left(\frac{M}{2\pi i \hbar t} \right)^{\frac{3}{2}} \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{3}{2}} \int d^3r' \, e^{\frac{iM(\vec{r} - \vec{r}')^2}{2\hbar t}} e^{i\vec{k}_0 \cdot \vec{r}'} e^{-\frac{(\vec{r}' - \vec{r}_0)^2}{2b^2}}$$

que pode ser simplificada introduzindo variáveis de posição deslocadas $\vec{s} \equiv \vec{r} - \vec{r}_0$,

$$\varphi(\vec{r}, t) = \left(\frac{M}{2\pi i \hbar t} \right)^{\frac{3}{2}} \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{3}{2}} e^{i\vec{k}_0 \cdot \vec{r}_0} \int d^3s' \, e^{\frac{iM(\vec{s} - \vec{s}')^2}{2\hbar t}} e^{i\vec{k}_0 \cdot \vec{s}'} e^{-\frac{s'^2}{2b^2}}.$$

Essa integral pode também ser calculada em termos de integrais gaussianas nas componentes cartesianas de \vec{s}' . O resultado que se obtém nos dois casos é

$$\varphi(\vec{r}, t) = \left[\frac{1}{B(t)\sqrt{\pi}} \right]^{\frac{3}{2}} \exp \left[-\frac{\left(\vec{r} - \vec{r}_0 - \frac{\hbar \vec{k}_0 t}{M} \right)^2}{2B^2(t)} \left(1 - \frac{i\hbar t}{Mb^2} \right) - \frac{i\hbar k_0^2 t}{2M} + i\vec{k}_0 \cdot \vec{r} \right] \quad (3.31)$$

onde $B(t)$ é dado por

$$B(t) = b\sqrt{1 + \frac{\hbar^2 t^2}{M^2 b^4}} = b\sqrt{1 + \frac{t^2}{\tau^2}}$$

onde foi definida a escala de tempo, dependente da massa e da largura inicial do pacote gaussiano, $\tau = Mb^2/\hbar$.

A densidade de probabilidade associada a $\varphi(\vec{r}, t)$

$$\rho(\vec{r}, t) = \left[\frac{1}{B(t)\sqrt{\pi}} \right]^3 \exp \left[-\frac{\left(\vec{r} - \vec{r}_0 - \frac{\hbar \vec{k}_0 t}{M} \right)^2}{B^2(t)} \right]$$

corresponde a uma distribuição espacial gaussiana, com largura dependente do tempo $B(t)$, cujo centro se desloca a partir da posição inicial \vec{r}_0 com velocidade constante $\hbar\vec{k}_0/M = \vec{v}_0$. A largura aumenta com o tempo de uma forma *não linear* tal que dB/dt cresce monotonicamente de zero, para $t = 0$, até o valor assintótico $\hbar/Mb = b/\tau$, para $t \rightarrow \infty$. É importante notar também que, sendo o momento \vec{p} uma constante do movimento para a partícula livre, a distribuição de probabilidade no espaço de momento $|\tilde{\varphi}(\vec{k}, t)|^2$ é independente do tempo. De fato, a aplicação do operador de evolução sobre $\tilde{\varphi}(\vec{k}, 0)$ introduz apenas fatores de fase dependentes de \vec{k} . Desse modo, a evolução temporal destrói a condição de incerteza mínima que se aplica ao estado inicial.

Uma imagem mais clara desse processo de alargamento espacial do pacote gaussiano pode ser obtida estudando também a distribuição de corrente de probabilidade ou, equivalentemente, o campo determinado pelo gradiente da fase de $\varphi(\vec{r}, t)$

$$\frac{S}{\hbar} = \vec{k}_0 \cdot \vec{r} - \frac{\hbar k_0^2 t}{2M} + \frac{\left(\vec{r} - \vec{r}_0 - \frac{\hbar \vec{k}_0 t}{M}\right)^2}{2b^2 \left(1 + \frac{t^2}{\tau^2}\right)} \frac{t}{\tau}.$$

O campo de velocidades associado à distribuição de probabilidade $\rho(\vec{r}, t)$ é, nesse caso, dado por

$$\vec{v}(\vec{r}, t) \equiv \frac{\nabla S}{M} = \frac{\hbar \vec{k}_0}{M} + \left(\vec{r} - \vec{r}_0 - \frac{\hbar \vec{k}_0 t}{M}\right) \frac{t}{\tau^2 + t^2} = \vec{v}_0 + (\vec{r} - \vec{r}_0 - \vec{v}_0 t) \frac{t}{\tau^2 + t^2}.$$

Essa expressão mostra que, para $t > 0$, o campo uniforme inicial \vec{v}_0 é modificado por uma componente adicional que tem a propriedade de ser *isotrópica* e *linearmente crescente* com relação ao centro do pacote $\vec{r}_0 + \vec{v}_0 t$, com uma taxa de variação com a distância que depende do tempo. Introduzindo a distância ao centro do pacote $\Delta\vec{r} \equiv \vec{r} - \vec{r}_0 - \vec{v}_0 t$ essa componente "intrínseca" do campo de velocidades, determinada apenas pela massa M da partícula e pela largura inicial b do pacote, é dada por

$$\Delta\vec{v}(\Delta\vec{r}, t) = \frac{t}{\tau^2 + t^2} \Delta\vec{r}$$

e dá lugar a uma componente correspondente da corrente de probabilidade

$$\Delta\vec{j}(\Delta\vec{r}, t) = \left[\frac{1}{B(t)\sqrt{\pi}} \right]^3 e^{-\frac{\Delta r^2}{B^2(t)}} \frac{t}{\tau^2 + t^2} \Delta\vec{r}$$

que mostra que os efeitos do crescimento com a distância Δr da componente intrínseca da velocidade são na realidade limitados pelo decréscimo gaussiano da densidade de probabilidade.

Dessa forma, a ordem de grandeza dos seus efeitos pode ser estimada calculando $\Delta\vec{v}(\Delta\vec{r}, t)$ para distâncias ao centro do pacote da ordem do seu tamanho no tempo t , $\Delta\vec{r} = B(t)\hat{r}$, onde \hat{r} é o vetor unitário na direção de $\Delta\vec{r}$. Essa quantidade

$$\Delta\vec{v}(B(t)\hat{r}, t) = \frac{b}{\tau} \frac{t}{\sqrt{\tau^2 + t^2}} \hat{r}$$

de fato reproduz a taxa de variação da largura com o tempo, $dB(t)/dt$. Um ponto importante a ser discutido é o mecanismo responsável pela dependência dessa quantidade com o tempo. Em outras palavras, é preciso elucidar as razões que fazem com que a tendência à expansão espacial do pacote de ondas se instale progressivamente, sendo nula em $t = 0$ e se aproximando assintoticamente do limite b/τ para $t \rightarrow \infty$.

Para tratar essa questão é preciso observar inicialmente que a escala de tempo τ é determinada fundamentalmente pela relação de incerteza de Heisenberg. De fato, a dispersão quadrática média de qualquer das componentes cartesianas do operador posição na condição inicial é $\sigma_q(t = 0) = b/\sqrt{2}$. Dado que se trata de um estado de incerteza mínima, isso implica em uma dispersão correspondente em qualquer das componentes cartesianas do momento que é dada por $\sigma_p = \hbar/b\sqrt{2}$. Como o momento é uma constante do movimento, esta última dispersão é preservada ao longo do tempo. Essas dispersões são *propriedades intrínsecas* do estado inicial considerado, e a escala de tempo τ pode ser escrita em termos delas como

$$\tau = \frac{(\sigma_q(t = 0))}{\left(\frac{\sigma_p}{M}\right)}.$$

O numerador nessa razão representa a escala das dimensões espaciais do pacote inicial, ao passo que o denominador representa a escala das *diferenças de velocidade* impostas pelas relações de incerteza. Dessa forma τ corresponde essencialmente ao tempo durante o qual uma distância da ordem da extensão espacial do pacote inicial é percorrida com a velocidade correspondente à *dispersão* de velocidades, e deve ser visto portanto como um "tempo de maturação" do estado inicial, durante o qual as componentes com velocidade relativa ao centro do pacote com maior módulo tendem a se concentrar na região superficial do pacote. Isso implica num processo de correlacionamento de posição e momento, e leva finalmente à expansão assintoticamente linear em t (sempre com referência ao centro do pacote), e com a taxa determinada pela dispersão em momento σ_p .

Para quantificar essas idéias é possível (embora não seja usual) introduzir um "parâmetro de maturidade" $\mu(t)$ para o estado gaussiano dependente do tempo, definido como

$$\sin \mu(t) = \frac{M\Delta\vec{v}(B(t)\hat{r}, t) \cdot \hat{r}}{\sqrt{2}\sigma_p} = \frac{t}{\sqrt{\tau^2 + t^2}}.$$

Esse parâmetro mede a dispersão do campo intrínseco de velocidades em extremidades opostas do pacote em unidades da dispersão de velocidades correspondente à relação de incerteza. Essa quantidade exprime portanto em que medida essa dispersão de velocidades, de origem cinemática, se realiza *em termos de correlações entre posição e momento*. É fácil ver que, com essa definição, $\tan \mu = t/\tau$ corresponde à "idade" do pacote expressa em unidades de τ .

A correlação entre posição e momento obtida acima, no pacote com "idade" diferente de zero, é capturada também pela covariância das componentes de \vec{p} e \vec{q} , incluída na relação geral de incerteza devida a Schrödinger (v. seção **2.2.4**). De fato, considerando por simplicidade o caso de um pacote gaussiano unidimensional centrado em $x = 0$ e com momento médio nulo⁴, a covariância de p e q é dada por

$$\begin{aligned}\sigma_{pq}^{(\varphi)}(t) &= \frac{1}{2} \langle \varphi(t) | (pq + qp) | \varphi(t) \rangle = \frac{1}{2} \langle \varphi(t) | (2pq + i\hbar) | \varphi(t) \rangle = \\ &= \frac{i\hbar}{2} + \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi^*(x, t) \frac{d}{dx} x \varphi(x, t) dx = \frac{\hbar^2 t}{2Mb^2}.\end{aligned}$$

A expressão para $\varphi(x, t)$ a ser usada para o cálculo da integral é facilmente deduzida de (3.31) anulando \vec{r}_0 e \vec{k}_0 e reduzindo a uma única dimensão, o que leva ao resultado

$$\varphi(x, t) = \left[\frac{1}{B(t)\sqrt{\pi}} \right]^{\frac{1}{2}} \exp \left[-\frac{x^2}{2B^2(t)} \left(1 - \frac{i\hbar t}{Mb^2} \right) \right].$$

As dispersões $\sigma_q^{(\varphi)2}(t)$ e $\sigma_p^{(\varphi)2}(t)$ calculadas com essa mesma função de onda dão, por outro lado,

$$\sigma_q^{(\varphi)2}(t) = \frac{B^2(t)}{2} \quad \text{e} \quad \sigma_p^{(\varphi)2}(t) = \frac{\hbar^2}{2b^2}$$

de modo que o determinante da matriz de covariância é

$$\det \begin{pmatrix} \frac{B^2(t)}{2} & \frac{\hbar^2 t}{2Mb^2} \\ \frac{\hbar^2 t}{2Mb^2} & \frac{\hbar^2}{2b^2} \end{pmatrix} = \frac{\hbar^2}{4} \left(\frac{B^2(t)}{b^2} - \frac{\hbar^2 t^2}{M^2 b^4} \right) = \frac{\hbar^2}{4}$$

⁴Usando argumentos de paridade é fácil ver que, no caso de um pacote gaussiano tridimensional, covariâncias do tipo $\sigma_{r_i r_j}^{(\varphi)}$, $\sigma_{p_i p_j}^{(\varphi)}$ e $\sigma_{r_i p_j}^{(\varphi)}$ com $i \neq j$ se anulam, de modo que esse caso pode ser tratado componente por componente com o mesmo resultado obtido no caso unidimensional. A restrição a momento médio e posição média nulos também não implica em perda de generalidade, pois apenas serve para dispensar as subtrações que de outra forma seriam necessárias para o cálculo dos elementos da matriz de covariância.

mostrando que o pacote gaussiano *satura* a relação de incerteza de Schrödinger *em qualquer tempo*. Isso permite ainda a vinculação do pacote, mesmo em $t \neq 0$, a um tipo generalizado de estado de incerteza mínima, e pode ser interpretado como indicativo de que o crescimento com t do produto das incertezas (segundo Heisenberg) é na realidade um reflexo do processo de correlação posição-momento que se instala através da dinâmica quântica.

Uma aplicação adicional da relação geral de incerteza de Schrödinger consiste no estudo do tempo característico associado à variação da dispersão quadrática do pacote gaussiano livre. Considerando ainda o caso do pacote unidimensional centrado na origem e com momento médio nulo, ela consiste em calcular, para esse caso, os vários termos da relação (2.35) tomando $g = x^2$. Para isso é preciso primeiramente construir os operadores cujo valor médio no pacote de ondas é nulo no tempo t

$$g_{(\varphi)} \equiv x^2 - \langle \varphi(t) | x^2 | \varphi \rangle = x^2 - \frac{B^2(t)}{2} \quad \text{e} \quad H_{(\varphi)} \equiv H - \langle \varphi(t) | H | \varphi \rangle = \frac{p^2}{2M} - \frac{\hbar^2}{4Mb^2}.$$

Esses operadores permitem calcular os elementos da matriz de covariância como

$$\begin{aligned} \sigma_g^{(\varphi)2} &= \langle \varphi | g_{(\varphi)}^2 | \varphi \rangle = \frac{B^4(t)}{2} = \frac{b^2}{2} \left(1 + \frac{\hbar^2 t^2}{M^2 b^4} \right)^2 \\ \sigma_H^{(\varphi)2} &= \langle \varphi | H_{(\varphi)}^2 | \varphi \rangle = \frac{\hbar^4}{8M^2 b^4} \quad \text{e} \\ \sigma_{gH}^{(\varphi)} &= \frac{\langle \varphi | H_{(\varphi)} g_{(\varphi)} - g_{(\varphi)} H_{(\varphi)} | \varphi \rangle}{2} = \frac{\langle \varphi | p^2 x^2 + x^2 p^2 | \varphi \rangle}{4M} - \frac{\hbar^2 B^2(t)}{8Mb^2} \\ &= -\frac{\hbar^2}{4M} \left(1 - \frac{\hbar^2 t^2}{M^2 b^4} \right) \end{aligned}$$

Levando esses ingredientes à relação (2.35) resulta, após alguns passos algébricos simples,

$$\sqrt{\frac{2\hbar^4 t^2}{(\hbar^2 t^2 + M^2 b^4)^2}} \tau_{(x^2)}^{(\varphi)} \geq 1$$

ou seja

$$\tau_{(x^2)}^{(\varphi)} \geq \frac{1}{\sqrt{2}} \left(t + \frac{M^2 b^4}{\hbar^2 t} \right). \quad (3.32)$$

Para a obtenção da relação (2.35), o tempo característico $\tau_g^{(\varphi)}$ foi definido como

$$\tau_g^{(\varphi)} \equiv \left| \frac{\hbar \sigma_g^{(\varphi)}}{\langle \varphi | [g, H] | \varphi \rangle} \right|,$$

e essa quantidade pode ser calculada diretamente neste caso para $g = x^2$ com o resultado

$$\tau_{(x^2)}^{(\varphi)} = \frac{1}{\sqrt{2}} \left(t + \frac{M^2 b^4}{\hbar^2 t} \right) \quad (3.33)$$

o que mostra que a desigualdade (3.32) é também saturada neste caso (isto é, ela vale como igualdade *em qualquer tempo*). Para entender o comportamento singular da expressão que aparece em (3.32) e (3.33) para $t \rightarrow 0$ e $t \rightarrow \infty$ basta notar que, para $t \rightarrow 0$, a velocidade $\langle d(x^2)/dt \rangle = \hbar^2 t / M^2 b^2$ se anula enquanto a dispersão $\sigma_{x^2}^{(\varphi)}$ é diferente de zero; e que para $t \rightarrow \infty$ o crescimento linear com t da velocidade é dominado pelo crescimento quadrático da dispersão. As duas singularidades de $\tau_{(x^2)}^{(\varphi)}$ têm portanto significados de certa forma diferentes, pois enquanto a singularidade em $t = 0$ se deve à estacionaridade inicial da dispersão, a singularidade em $t \rightarrow \infty$ reflete apenas as taxas diferentes de variação com o tempo de $\sigma_{x^2}^{(\varphi)}$ e de $\langle d(x^2)/dt \rangle = d\sigma_{x^2}^{(\varphi)}/dt$, e é portanto apenas um artefato da definição de $\tau_{(x^2)}^{(\varphi)}$.

Finalmente, é interessante comparar esses resultados com o que decorre do uso da relação (2.34), proveniente da relação de incerteza que ignora a contribuição da covariância, que dá simplesmente

$$\sigma_H^{(\varphi)} \tau_{(x^2)}^{(\varphi)} \geq \frac{\hbar}{2} \quad \text{ou seja} \quad \tau_{(x^2)}^{(\varphi)} \geq \frac{\sqrt{2} M b^2}{\hbar}.$$

O valor mínimo de $\tau_{(x^2)}^{(\varphi)}$ pode ser obtido de (3.33) como sendo

$$\tau_{(x^2)}^{(\varphi)} \Big|_{\min} = \frac{\sqrt{2} M b^2}{\hbar}, \quad \text{para} \quad t \rightarrow t_0 = \frac{M b^2}{\hbar}$$

mostrando que essa desigualdade é saturada apenas no instante t_0 , em que a covariância $\sigma_{gH}^{(\varphi)}$ se anula. O fato de que nesse instante a covariância $\sigma_{pq}^{(\varphi)}$ é diferente de zero (pois se anula apenas para $t = 0$) ilustra o fato de que as correlações expressas pela covariância dependem crucialmente não apenas do estado mas também das variáveis dinâmicas consideradas.

Com alguns passos adicionais é possível também identificar pacotes gaussianos com maturidade e idade *negativas*. De fato, considerando um novo problema de condições iniciais em que o estado inicial $\psi(\vec{r}, 0)$ é definido em termos do conjugado complexo de um pacote gaussiano “maduro”, isto é

$$\psi(\vec{r}, t = 0) \equiv \varphi^*(\vec{r}, T)$$

onde T é um tempo grande na escala de τ , é fácil ver que a solução da equação de Schrödinger dependente do tempo satisfazendo essa condição inicial é

$$\psi(\vec{r}, t) = \varphi^*(\vec{r}, T - t).$$

De fato, a derivada com relação ao tempo dessa função é dada por

$$\frac{\partial \psi(\vec{r}, t)}{\partial t} = - \frac{\partial \varphi^*(\vec{r}, t')}{\partial t'} \Big|_{t'=T-t} ;$$

como, por outro lado, $H\varphi^*(\vec{r}, t') = -i\hbar\partial\varphi^*/\partial t'$, resulta que $\psi(\vec{r}, t)$ satisfaz a equação

$$i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial t} = H\psi(\vec{r}, t).$$

A densidade de probabilidade associada ao pacote $\psi(\vec{r}, t)$ é dada por

$$\rho_\psi(\vec{r}, t) = \left[\frac{1}{B(T-t)\sqrt{\pi}} \right]^3 \exp \left[- \frac{(\vec{r} - \vec{r}_0' + \vec{v}_0 t)^2}{B^2(T-t)} \right]$$

onde $\vec{r}_0' = \vec{r}_0 + \vec{v}_0 T$ é a posição inicial do centro desse pacote. Para tempos no intervalo $0 \leq t \leq T$, portanto, ocorre uma *contração* das dimensões espaciais do pacote até que, para $t = T$, ele volta a ser representado por $\varphi(\vec{r}, t = 0)$. Para $t > T$ a evolução temporal deste último estado é reproduzida com as propriedades discutidas anteriormente.

O campo de velocidades obtido a partir da análise da fase de $\psi(\vec{r}, t)$ é dado por

$$\vec{v}_\psi(\vec{r}, t) = -\vec{v}_0 - (\vec{r} - \vec{r}_0' + \vec{v}_0 t) \frac{T-t}{\tau^2 + (T-t)^2}$$

que corresponde ao campo de velocidades “intrínseco”, relativo ao centro do pacote,

$$\Delta\vec{v}_\psi(\Delta\vec{r}, t) = - \frac{T-t}{\tau^2 + (T-t)^2} \Delta\vec{r}$$

onde agora $\Delta\vec{r} = \vec{r} - \vec{r}_0' + \vec{v}_0 t$. Para $t < T$ esse campo radial é dirigido para *dentro* do pacote, dando lugar portanto a valores *negativos* do parâmetro de maturidade μ e da “idade” $\tan\mu$. Ambas essas quantidades são porém *crescentes*, como no caso tratado anteriormente.

O processo que leva da solução $\varphi(\vec{r}, t)$ à solução $\psi(\vec{r}, t)$ é claramente um processo de *inversão temporal*. De fato, a conjugação complexa aplicada a $\varphi(\vec{r}, t)$ mantém a densidade de probabilidade inalterada e troca o sinal do potencial do campo de velocidades, de modo que a evolução de $\psi(\vec{r}, t)$ repete, em seqüência inversa, a evolução de $\varphi(\vec{r}, t)$. Como a inversão temporal implica na inversão de todas as velocidades do sistema, a sua aplicação a um pacote

gaussiano com correlações posição-momento inverte os momentos envolvidos nessas correlações e dessa forma inverte o sinal de μ . Com isso a expansão quase linear do pacote “maduro” é substituída por uma contração que o leva de volta à situação de “idade zero” (campo intrínseco de velocidades nulo). Ao contrário do que ocorre com a expansão, a contração de um pacote gaussiano só pode ocorrer durante um tempo *finito*.

Ainda outra forma de expor a correlação entre posição e momento no pacote de ondas gaussiano consiste em utilizar uma representação do estado quântico num espaço de fases *clássico*, introduzida por Wigner (*Phys. Rev.* **40**, 749 (1932)). Considerando, por simplicidade, o caso de um pacote unidimensional (na coordenada x), com $k_0 = 0$ e $x_0 = 0$, (a discussão precedente mostra que esse caso contém na realidade uma descrição completa do movimento intrínseco de um pacote tridimensional com momento médio diferente de zero) essa representação tem como ponto de partida a forma “não diagonal” da densidade de probabilidade

$$\rho(x, x', t) \equiv \varphi(x, t) \varphi^*(x', t) = \langle x | \varphi(t) \rangle \langle \varphi(t) | x' \rangle$$

onde esta última notação foi introduzida para justificar a denominação “matriz densidade” usualmente dada a esse objeto. No caso específico do pacote gaussiano

$$\rho(x, x', t) = \frac{1}{B(t)\sqrt{\pi}} \exp - \left[\frac{x^2}{2B^2(t)} \left(1 - \frac{i\hbar t}{Mb^2} \right) + \frac{x'^2}{2B^2(t)} \left(1 + \frac{i\hbar t}{Mb^2} \right) \right].$$

Os “elementos diagonais” dessa matriz densidade correspondem claramente à densidade de probabilidade na representação utilizada. A *transformada de Wigner* de $\rho(x, x', t)$ é definida, em primeiro lugar, reescrevendo essa densidade em termos das coordenadas

$$X = \frac{x + x'}{2} \quad \text{e} \quad \xi = x' - x$$

ou seja

$$x = X - \frac{\xi}{2} \quad \text{e} \quad x' = X + \frac{\xi}{2}$$

e tomando em seguida a transformada de Fourier com relação a ξ com um coeficiente apropriado

$$\rho_W(X, k, t) = \frac{1}{2\pi} \int d\xi e^{-ik\xi} \rho(X, \xi, t).$$

A distribuição de Wigner $\rho_W(X, k, t)$ está definida no “espaço de fases clássico” $X \times k$ e é uma função real, mas não necessariamente positiva, o que prejudica a sua interpretação

simplesmente como uma distribuição de probabilidades no espaço de fases clássico. No entanto, não é difícil verificar que as integrais

$$\rho(X, t) = \int dk \rho_W(X, k, t) \quad \text{e} \quad \bar{\rho}(k, t) = \int dX \rho_W(X, k, t) \quad (3.34)$$

correspondem efetivamente às distribuições de probabilidade respectivamente no espaço de posição e de momento, segundo a prescrição padrão da teoria quântica. A escolha do coeficiente $1/2\pi$ na definição de $\rho_W(X, k, t)$ tem na realidade o sentido de uma normalização que garante essas propriedades.

No caso do pacote gaussiano, a transformada de Wigner da matriz densidade pode ser obtida facilmente e tem a forma simples

$$\rho_W(X, k, t) = \frac{1}{\pi} e^{-b^2 k^2 - \frac{(X + \hbar k t / M)^2}{b^2}}$$

na qual é possível verificar explicitamente as propriedades anunciadas acima. Neste caso a distribuição de Wigner é no entanto uma função positiva de X e k . Para a condição inicial ($t = 0$) ela assume valores iguais sobre curvas do plano $X \times k$ que, calibrando o eixo X em unidades de b e o eixo k em unidades de $1/b$, são círculos centrados na origem. Os pontos dos semiplanos superior e inferior se deslocam respectivamente para a direita e para a esquerda, paralelamente ao eixo X e com velocidades proporcionais à sua distância a esse eixo. Desse modo os círculos da condição inicial se deformam quando $t > 0$ para elipses cujos eixos maiores estão inicialmente ao longo da bissetriz do primeiro quadrante e se alongam com o tempo, tendendo assintoticamente a se alinhar com o eixo X . A correlação momento-posição corresponde ao fato de que, de acordo com as relações (3.34) e para $t > 0$, as contribuições para a distribuição de probabilidade em posições $X > 0$ e $X < 0$ provém predominantemente de regiões com $k > 0$ e $k < 0$ respectivamente.

O espectro do momento \vec{p} foi tratado até aqui como um espectro contínuo e portanto associado a autofunções impróprias (de norma infinita) sujeitas à normalização especial em termos de funções delta de Dirac. Uma forma freqüentemente utilizada de evitar a consideração de "autofunções" de norma infinita é impor *condições de contorno periódicas* em um volume $V = L^3$ sobre as soluções da equação de autovalores para o momento. Isso significa que as soluções (3.23) dessa equação devem satisfazer à condição de periodicidade

$$\phi_{\vec{k}}(\vec{r} + \vec{n}L) = \phi_{\vec{k}}(\vec{r})$$

onde $\vec{n} = \{n_1, n_2, n_3\}$ é um vetor de componentes *inteiras*, negativas, positivas ou nulas. Essa condição restringe os valores de $\vec{k} = \{k_1, k_2, k_3\}$ de modo que

$$(\vec{k} \cdot \vec{n})L = \pm 2\pi N,$$

com N inteiro, o que corresponde a vetores \vec{k} do tipo

$$\vec{k} = \frac{2\pi}{L} \vec{\nu}$$

onde $\vec{\nu} \equiv \{\nu_1, \nu_2, \nu_3\}$ é também um vetor de componentes inteiras, que também podem ser negativas, nulas ou positivas. A condição de periodicidade implica portanto na *discretização* do espectro do momento.

O conjunto de autofunções periódicas $\phi_{\vec{k}}(\vec{r})$ permite representar funções igualmente periódicas de \vec{r} no sentido em que tais funções podem ser representadas por *séries* de Fourier. Isso implica, por outro lado, em redefinir o produto escalar de funções periódicas do tipo considerado como

$$\langle \psi | \varphi \rangle \rightarrow \int_{\vec{r} \in V} d^3r \, \psi^*(\vec{r}) \varphi(\vec{r})$$

onde a integração é restrita ao volume de periodicidade $V = L^3$, de modo que as autofunções periódicas do momento têm de fato norma finita e a constante de normalização $N_{\vec{k}}$ pode ser escolhida de forma a satisfazer a condição

$$\int_{\vec{r} \in V} d^3r \, \phi_{\vec{k}}^*(\vec{r}) \phi_{\vec{k}'}(\vec{r}) = \delta_{\vec{k}, \vec{k}'}$$

onde $\delta_{\vec{k}, \vec{k}'}$ deve ser entendido como um delta de Kronecker, consistentemente com o caráter discreto do espectro. Com essa normalização, as autofunções periódicas do momento têm a forma

$$\phi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{r}}. \quad (3.35)$$

As situações típicas em que o uso de condições de contorno periódicas para as autofunções do momento é conveniente *não* exigem qualquer propriedade de periodicidade correspondente para o sistema considerado. De fato, a idéia de um volume de periodicidade é usualmente entendida de forma que esse volume seja suficientemente grande para conter essencialmente *todo* o sistema de interesse. Dessa forma o trabalho pode ser desenvolvido sem os eventuais inconvenientes decorrentes da consideração de autovetores impróprios. O resultado final obtido para quantidades físicas deve evidentemente ser independente de V . Um recurso de cálculo bastante útil para expressões que envolvem somas sobre o espectro discreto do momento, com condições de contorno periódicas, consiste em “passar ao limite do contínuo”, o que significa fazer $L \rightarrow \infty$, com o que o espectro discreto do momento se torna progressivamente mais denso e somas sobre o espectro discreto podem ser substituídas por integrais sobre \vec{k} . Essa substituição envolve um “jacobiano” $dn/d\vec{k}$ que tem o sentido de uma densidade de estados (discretos) por

unidade de volume do espaço de momento. Como o “volume por estado discreto” no espaço \vec{k} é $(\Delta k_1)^3 = (2\pi/L)^3$, essa passagem de somas a integrais deve ser feita de acordo com

$$\sum_{\vec{k}} F(\vec{k}) \xrightarrow{V \rightarrow \infty} \int d^3 k \left(\frac{dn}{d\vec{k}} \right) F(\vec{k}) = \frac{V}{(2\pi)^3} \int d^3 k F(\vec{k}).$$

Em particular, é imediato verificar que essa prescrição transforma elementos de matriz calculados com as funções de onda (3.35) nos elementos de matriz correspondentes, calculados com a normalização do contínuo (3.25).

A dedução das relações de incerteza feita no capítulo anterior envolve implicitamente algumas hipóteses que podem não ser satisfeitas quando se usa autofunções do momento com condições de contorno periódicas, conduzindo eventualmente a situações paradoxais. Por exemplo, no caso mais simples da relação de incerteza posição-momento

$$\sigma_q \sigma_p \geq \frac{\hbar}{2}$$

uma dispersão nula do momento acarreta $\sigma_q \rightarrow \infty$. No entanto, as autofunções do momento com condições de contorno periódicas e normalização (3.35) dão, em uma dimensão, $\sigma_q \sim L$.

A origem desta dificuldade pode ser encontrada reexaminando, neste caso, a dedução feita ali da relação de incerteza. Como antes, é conveniente definir os operadores deslocados

$$q_{(\varphi)} = q - \langle \varphi | q | \varphi \rangle \quad \text{e} \quad p_{(\varphi)} = p - \langle \varphi | p | \varphi \rangle$$

onde $\langle x | \varphi \rangle = \varphi(x)$ é uma função de onda derivável satisfazendo agora às condições de periodicidade

$$\varphi(x) = \varphi(x + nL) \quad \text{e} \quad \frac{d\varphi(x)}{dx} = \frac{d\varphi(x + nL)}{dx}.$$

Desse modo,

$$\sigma_q^2 \sigma_p^2 = \langle \varphi | q_{(\varphi)}^2 | \varphi \rangle \langle \varphi | p_{(\varphi)}^2 | \varphi \rangle.$$

Introduzindo as funções

$$\bar{\varphi}_q(x) = q_{(\varphi)} \varphi(x) \quad \text{e} \quad \bar{\varphi}_p(x) = p_{(\varphi)} \varphi(x)$$

e usando as condições de periodicidade impostas sobre $\varphi(x)$ (que são necessárias para garantir o anulamento de termos integrados em integrações por partes), é fácil ver que

$$\langle \varphi | q_{(\varphi)}^2 | \varphi \rangle = \int_{x \in L} dx |\bar{\varphi}_q(x)|^2$$

e uma relação análoga para $\langle \varphi | p_{(\varphi)}^2 | \varphi \rangle$ envolvendo $\bar{\varphi}_p(x)$. A desigualdade de Schwartz dá, então

$$\langle \varphi | q_{(\varphi)}^2 | \varphi \rangle \langle \varphi | p_{(\varphi)}^2 | \varphi \rangle \geq \left| \int dx \bar{\varphi}_p^*(x) \bar{\varphi}_q(x) \right|^2 \geq \frac{1}{4} \left| \int dx \bar{\varphi}_p^*(x) \bar{\varphi}_q(x) - \int dx \bar{\varphi}_q^*(x) \bar{\varphi}_p(x) \right|^2.$$

No capítulo anterior, o último termo foi reduzido a uma expressão envolvendo o comutador de $q_{(\varphi)}$ e $p_{(\varphi)}$ usando

$$\int dx \bar{\varphi}_q^*(x) \bar{\varphi}_p(x) = \int dx \left(q_{(\varphi)} \varphi(x) \right)^* p_{(\varphi)} \varphi(x) = \int dx \varphi(x)^* \left(q_{(\varphi)} p_{(\varphi)} \right) \varphi(x)$$

e

$$\int dx \bar{\varphi}_p^*(x) \bar{\varphi}_q(x) = \int dx \left(p_{(\varphi)} \varphi(x) \right)^* q_{(\varphi)} \varphi(x) = \int dx \varphi(x)^* \left(p_{(\varphi)} q_{(\varphi)} \right) \varphi(x) \quad (?)$$

A última relação, porém, *não* é correta neste caso. O produto escalar $(\bar{\varphi}_p, \bar{\varphi}_q)$ pode, de fato, se calculado explicitamente com o resultado

$$(\bar{\varphi}_p, \bar{\varphi}_q) = -i\hbar (x - \langle \varphi | q | \varphi \rangle) |\varphi(x)|^2 \Big|_0^L + \int dx \varphi(x)^* \left(p_{(\varphi)} q_{(\varphi)} \right) \varphi(x)$$

onde o primeiro termo, proveniente da integração por partes envolvida no remanejamento de $p_{(\varphi)}$, é $-i\hbar L |\varphi(L)|^2$. Esse termo se anula apenas se $\varphi(L) = \varphi(0) = 0$, o que *não* foi exigido. Em particular, esta última condição não é satisfeita pelas autofunções periódicas do momento, para as quais $\phi_{\vec{k}}(L) = \phi_{\vec{k}}(0) = 1/\sqrt{L}$. A contribuição do termo integrado é então $-i\hbar$ e *cancela* a contribuição vinda do comutador de $q_{(\varphi)}$ e $p_{(\varphi)}$.

O problema encontrado no início, portanto, pode ser relacionado com uma quebra de hermiticidade de $p_{(\varphi)}$. A hermiticidade desse operador depende, pelo visto, da periodicidade das funções de onda, e a função $\bar{\varphi}_q(x) = q_{(\varphi)} \varphi(x)$ *não é periódica*. Por outro lado, a relação

$$\langle \varphi | q_{(\varphi)}^2 | \varphi \rangle \langle \varphi | p_{(\varphi)}^2 | \varphi \rangle \geq \frac{1}{4} \left| \int dx \bar{\varphi}_p^*(x) \bar{\varphi}_q(x) - \int dx \bar{\varphi}_q^*(x) \bar{\varphi}_p(x) \right|^2$$

tem validade geral. Para que uma relação de incerteza com a forma escrita no início seja válida é preciso que valha a relação

$$\int dx \left(q_{(\varphi)} \varphi(x) \right)^* p_{(\varphi)} \varphi(x) - \int dx \varphi(x)^* \left(p_{(\varphi)} q_{(\varphi)} \right) \varphi(x) = i\hbar \int dx |\varphi(x)|^2$$

que é a “forma fraca” do comutador de $q_{(\varphi)}$ e $p_{(\varphi)}$. Essa relação não é verificada no caso de condições de contorno periódicas quando $\varphi(x)$ não se anula nos extremos do período de integração.

3.3.4 Partícula sujeita a potencial externo em uma dimensão.

O tratamento quântico de uma partícula sujeita a um potencial externo $V(\vec{r})$ geral, em três dimensões, leva, na representação de Schrödinger, a uma equação diferencial parcial linear envolvendo três variáveis espaciais. Estados não estacionários envolvem também o tempo como quarta variável. O tratamento, mesmo numérico, desse tipo geral de equação é bastante trabalhoso. As dificuldades podem no entanto ser bastante aliviadas caso existam propriedades de simetria que possam ser exploradas para desacoplar algumas das variáveis envolvidas. O caso especial de potenciais centrais $V(\vec{r}) \rightarrow V(|\vec{r}|)$ será tratado dessa forma no capítulo 5. Existem, no entanto várias características das soluções da equação de Schrödinger geral que já aparecem no tratamento de problemas solúveis envolvendo o caso especial de uma única dimensão espacial. O objetivo desta seção é apresentar e discutir algumas dessas características.

A equação de Schrödinger que descreve uma partícula quântica de massa M em uma única dimensão espacial, sujeita a um potencial $V(x)$, é

$$-\frac{\hbar^2}{2M} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t) = i\hbar \frac{\partial \psi(x, t)}{\partial t}.$$

As soluções estacionárias dessa equação têm a forma

$$\psi_E(x, t) = \phi_E(x) e^{-\frac{iEt}{\hbar}}$$

que, substituída na equação de Schrödinger, mostra que as funções $\phi_E(x)$ devem satisfazer a equação diferencial de autovalores

$$-\frac{\hbar^2}{2M} \frac{d^2 \phi_E(x)}{dx^2} + V(x)\phi_E(x) = E\phi_E(x)$$

que é claramente a equação de autovalores para o operador hamiltoniano escrita na representação de Schrödinger. Essa equação adquire uma forma mais conveniente com as definições

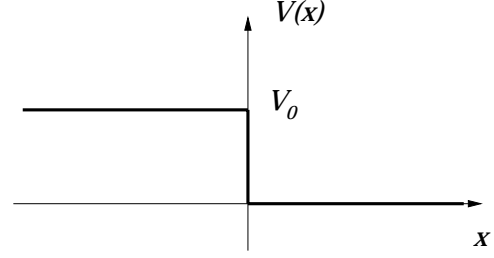
$$U(x) \equiv \frac{2M}{\hbar^2} V(x) \quad \text{e} \quad k^2 \equiv \frac{2ME}{\hbar^2}$$

que permitem reescreve-la como

$$\frac{d^2 \phi_E(x)}{dx^2} + [k^2 - U(x)] \phi_E(x) = 0. \quad (3.36)$$

As soluções buscadas dessa equação devem satisfazer condições de contorno apropriadas. Auto-funções $\phi_E(x)$ correspondentes ao espectro discreto do hamiltoniano devem ser normalizáveis, o que exige que se aproximem de zero de forma suficientemente rápida quando $x \rightarrow \pm\infty$.

Figura 3.1: Degrau de potencial de altura V_0 em $x = 0$.



Autofunções correspondentes ao espectro contínuo não são normalizáveis, mas são também especificadas prescrevendo o seu comportamento para valores muito grandes de x . Essas questões são ilustradas a seguir em algumas situações típicas.

1. Degrau de potencial. Neste caso a dependência de x do potencial se reduz a uma descontinuidade finita em um determinado ponto, que pode ser tomado como correspondendo a $x = 0$ (v. figura 3.1):

$$V(x) = \begin{cases} V_0 > 0 & \text{para } x \leq 0 \\ 0 & \text{para } x > 0. \end{cases}$$

Introduzindo a quantidade

$$k_0^2 \equiv k^2 - \frac{2MV_0}{\hbar^2}$$

a equação de autovalores (3.36) pode ser escrita como um par de equações, válidas para $x \leq 0$ e para $x > 0$ respectivamente

$$\begin{aligned} \frac{d^2 \phi_E}{dx^2} + k_0^2 \phi_E &= 0, & x \leq 0 \\ \frac{d^2 \phi_E}{dx^2} + k^2 \phi_E &= 0, & x > 0 \end{aligned}$$

cujas soluções gerais são⁵

⁵No caso em que k_0^2 e k^2 forem ambos positivos é conveniente definir k_0 e k que aparecem nas soluções gerais como a raiz quadrada *positiva* de k_0^2 e k^2 ; caso uma destas quantidades (ou ambas) seja (ou sejam) negativas, a forma das soluções gerais permanece válida usando valores imaginários puros para k_0 e/ou k . Para manter a associação de A_+ e B_+ com as exponenciais positivas basta tomar $k_0 = -i\kappa_0$ e $k = -i\kappa$ com κ_0 e κ positivos.

$$\phi_E(x \leq 0) = A_+ e^{ik_0 x} + A_- e^{-ik_0 x}$$

$$\phi_E(x > 0) = B_+ e^{ikx} + B_- e^{-ikx}.$$

Essa solução satisfaz a equação original (3.36) em qualquer ponto $x \neq 0$, mas a presença da segunda derivada e a descontinuidade do potencial em $x = 0$ exigem cuidados adicionais para que essa equação esteja satisfeita *também* em $x = 0$. Como o potencial $V(x)$ é *finito* (embora descontínuo) nesse ponto, é preciso que a segunda derivada esteja definida e seja também finita em $x = 0$. Isso só será possível se estiverem satisfeitas as condições

$$\lim_{\delta x \rightarrow 0} \phi_E(x + \delta x) - \phi_E(x - \delta x) = 0 \quad (3.37)$$

$$\lim_{\delta x \rightarrow 0} \frac{d\phi_E(x + \delta x)}{dx} - \frac{d\phi_E(x - \delta x)}{dx} = 0, \quad \delta x > 0$$

isto é, $\phi_E(x)$ e sua primeira derivada devem ser contínuas em $x = 0$. Essas condições obrigam as constantes A_{\pm} e B_{\pm} a satisfazer as relações

$$B_+ + B_- = A_+ + A_- \quad \text{e} \quad k(B_+ - B_-) = k_0(A_+ - A_-).$$

A determinação completa das quatro constantes A_{\pm} e B_{\pm} exige porém duas condições adicionais, que são obtidas impondo sobre a solução condições de contorno apropriadas.

Para discutir essas condições de contorno para $E > 0$ (e portanto $k^2 > 0$), é conveniente tratar separadamente casos em que $k_0^2 > 0$ (isto é, $E > V_0$) e casos em que $k_0^2 < 0$ (isto é, $0 < E < V_0$). No primeiro caso, tanto k como k_0 são reais, e as soluções têm um comportamento oscilatório com x dos dois lados do degrau de potencial. Como a equação diferencial é de segunda ordem, existem duas soluções independentes ϕ_E^i , $i = 1, 2$, para cada valor de E , que podem ser especificadas impondo duas condições adicionais independentes sobre as constantes A_{\pm} e B_{\pm} . Uma escolha usual para essas condições adicionais é

$$\phi_E^{(1)}(x) \rightarrow A_+^{(1)} = 1, \quad B_-^{(1)} = 0; \quad \phi_E^{(2)}(x) \rightarrow A_+^{(2)} = 0, \quad B_-^{(2)} = 1.$$

A escolha do valor 1 para $A_+^{(1)}$ e $B_-^{(2)}$ pode ser vista como um critério de normalização. De fato, como a equação de autovalores é linear e homogênea, as soluções são definidas apenas a menos de uma constante multiplicativa. Combinadas com as condições exigidas pela equação no ponto $x = 0$, essas condições de contorno levam às soluções

$$\phi_E^{(1)}(x) = \begin{cases} e^{ik_0x} + \frac{k-k_0}{k+k_0}e^{ik_0x}, & x < 0 \\ \frac{2k_0}{k+k_0}e^{ikx}, & x > 0 \end{cases}, \quad \phi_E^{(2)}(x) = \begin{cases} \frac{2k}{k+k_0}e^{-ik_0x}, & x < 0 \\ \frac{k-k_0}{k+k_0}e^{ikx} + e^{-ikx}, & x > 0 \end{cases}.$$

No limite em que $V_0 \rightarrow 0$ essas duas soluções se reduzem simplesmente às soluções independentes $e^{\pm ikx}$ para uma partícula livre. Elas não são normalizáveis (a não ser no sentido impróprio, com funções delta de Dirac) e o espectro é contínuo. Os autovalores na faixa de energia considerada, $E > V_0$, são duplamente degenerados.

No segundo caso, em que $0 < E < V_0$, $k_0 < 0$ e a solução geral para $x < 0$ é uma combinação de exponenciais reais, uma crescentes e outra decrescente. Para $x > 0$ a solução geral é ainda oscilatória. As soluções com comportamento exponencialmente crescente para $x \rightarrow -\infty$ correspondem a um crescimento exponencial da densidade de probabilidade nessa região classicamente proibida. Fisicamente, essa situação pode ser associada à presença de fontes arbitrariamente intensas de partículas em abscissas arbitrariamente negativas. Situações desse tipo são usualmente excluídas, o que leva à exclusão concomitante da solução exponencialmente crescente. A solução que resta é portanto a que satisfaz à condição de contorno $A_- = 0$ com o critério de normalização $B_- = 1$ que, juntamente com as condições que devem ser satisfeitas na origem, leva à solução

$$\phi_E^{(0)}(x) = \begin{cases} \frac{2k}{k+i\kappa_0}e^{\kappa_0x}, & x < 0, \quad \kappa_0 = +\sqrt{\frac{2M}{\hbar^2}(V_0 - E)} \\ \frac{k-i\kappa_0}{k+i\kappa_0}e^{ikx} + e^{-ikx}, & x > 0 \end{cases}.$$

O espectro é também contínuo nessa região de energias, mas não mais degenerado, devido à exclusão das soluções não limitadas em $x \rightarrow -\infty$.

Finalmente, no caso em que $E < 0$ as soluções gerais envolvem apenas exponenciais reais, e as condições exigidas em $x = 0$ impõe sempre comportamentos exponencialmente crescentes, seja para $x \rightarrow -\infty$ seja para $x \rightarrow +\infty$. A exemplo do que ocorre com a solução desse tipo no caso em que $0 < E < V_0$, essas soluções são descartadas.

As condições de contorno escolhidas para obter as diferentes soluções acima da equação de Schrödinger unidimensional têm interpretações físicas bastante simples. Elas prescrevem uma componente de amplitude normalizada à unidade e com vetor de onda dirigido para o degrau de potencial (o que pode ser realizado de dois modos quando $E > V_0$, correspondendo respectivamente a incidências a partir da esquerda e da direita e às soluções $\phi_E^{(1)}$ e $\phi_E^{(2)}$) além de outras componentes necessárias para regularizar o comportamento da solução na origem. Para as soluções $\phi_E^{(1)}$ e $\phi_E^{(2)}$ essas outras componentes se anulam no limite $V_0 \rightarrow 0$, e correspondem portanto ao espalhamento da componente incidente pelo degrau de potencial. A solução única

$\phi_E^{(0)}(x)$ no caso em que $0 < E < V_0$ deixa de existir no limite $V_0 \rightarrow 0$. A componente espalhada cujo vetor de onda é paralelo ao da onda incidente é chamada onda transmitida, enquanto a componente cujo vetor de onda é anti-paralelo ao da onda incidente é a onda refletida. Por exemplo, a solução $\phi_E^{(1)}$ tem uma onda transmitida de amplitude $2k_0/(k + k_0)$ e uma onda transmitida de amplitude $-(k - k_0)/(k + k_0)$. A solução única no caso $0 < E < V_0$, por outro lado, tem apenas uma onda refletida de amplitude $(k - i\kappa_0)/(k + i\kappa_0)$.

É instrutivo calcular a densidade de corrente de probabilidade associada a essas soluções. Tomando novamente como exemplo a solução $\phi_E^{(1)}$, o que se obtém para $x < 0$ é

$$\begin{aligned} j^{(1)}(x < 0) &= \frac{\hbar}{2iM} \left[\phi_E^{(1)*}(x) \frac{d}{dx} \phi_E^{(1)}(x) - \left(\frac{d}{dx} \phi_E^{(1)*}(x) \right) \phi_E^{(1)}(x) \right] = \\ &= \frac{\hbar k_0}{M} \left[1 - \frac{(k - k_0)^2}{(k + k_0)^2} \right] \equiv \frac{\hbar k_0}{M} (1 - R). \end{aligned}$$

Para $x > 0$ o que se obtém de forma análoga é

$$j^{(1)}(x > 0) = \frac{\hbar k_0}{M} \left[1 - \frac{4k_0^2}{(k + k_0)^2} \right] \equiv \frac{\hbar k}{M} T = \frac{\hbar k_0}{M} (1 - R).$$

Essas relações incluem as definições dos coeficientes de reflexão R e de transmissão T , como o quadrado dos módulos das amplitudes das ondas refletida e transmitida, quando a amplitude da onda incidente é igual a 1. A corrente em $x < 0$ é proporcional à velocidade associada ao vetor de onda da onda incidente e tem uma contribuição da onda incidente e outra (negativa) da onda refletida. Deve ser notado que esta última, nas condições deste caso, é *nula* na mecânica clássica. Em $x > 0$ a corrente é dada pelo produto da coeficiente de transmissão pela velocidade associada ao vetor de onda da onda transmitida. Essa quantidade é na realidade *igual* à densidade de corrente calculada para $x < 0$ (isto é, há *conservação do fluxo de probabilidade* através do degrau de potencial em $x = 0$) e implica na relação

$$R + \frac{k}{k_0} T = 1$$

entre o coeficiente de reflexão e o coeficiente de transmissão. O escalonamento do coeficiente de transmissão (definido em termos da amplitude da onda transmitida) pelo fator k/k_0 é uma consequência da conservação do fluxo de probabilidade, pois tem em conta a diferença das velocidades associadas aos vetores de onda em cada uma das duas regiões.

Para a solução única correspondente ao caso $0 < E < V_0$ é fácil verificar que o coeficiente de transmissão é zero e o coeficiente de reflexão é igual a 1, enquanto a densidade de corrente

é nula tanto para $x < 0$ como para $x > 0$. Como $R=1$, a amplitude da onda refletida difere da amplitude da onda incidente por uma fase apenas, que é usualmente explicitada através das definições

$$\frac{k - i\kappa_0}{k + i\kappa_0} \equiv -e^{2i\delta(k)} \equiv -S(k), \quad \tan \left[\frac{\pi}{2} - \delta(k) \right] = -\frac{\kappa_0}{k}$$

com as quais a solução pode ser escrita, na região $x > 0$, como

$$\phi_E^{(0)}(x > 0) = e^{-ikx} - e^{2i\delta(k)} e^{ikx}.$$

O fato de que a defasagem da onda refletida com relação à onda incidente é uma função de k (ou, equivalentemente, da energia E) tem conseqüências físicas importantes que podem ser identificadas construindo em $t = 0$ um pacote de ondas da forma

$$\psi(x, t = 0) = \frac{1}{2\pi} \int dk \tilde{u}(k - \bar{k}) \phi_E^{(0)}(x)$$

onde $\tilde{u}(k - \bar{k})$ é uma função real e positiva que difere de zero apenas em uma faixa de vetores de onda suficientemente estreita, em torno de $k = \bar{k}$, de forma que apenas soluções do tipo $\phi_E^{(0)}(x)$ estejam envolvidas na integral. A evolução temporal livre desse pacote dá, no tempo t e na região $x > 0$,

$$\begin{aligned} \psi(x > 0, t) &= \frac{1}{2\pi} \int dk \tilde{u}(k - \bar{k}) e^{-\frac{i\hbar}{2M} k^2 t} \left[e^{-ikx} - e^{2i\delta(k)} e^{ikx} \right] \\ &= \frac{1}{2\pi} \int dk \tilde{u}(k - \bar{k}) \left[e^{-i\varphi_-(k, x, t)} - e^{i\varphi_+(k, x, t)} \right] \end{aligned}$$

com as fases $\varphi_{\pm}(k, x, t)$ dadas por

$$\varphi_-(k, x, t) = \frac{\hbar k^2 t}{2M} + kx \quad \text{e} \quad \varphi_+(k, x, t) = -\frac{\hbar k^2 t}{2M} + 2\delta(k) + kx.$$

Num dado tempo t , a região do espaço x em que a amplitude $\psi(x, t)$ do pacote é maior é aquela em que as componentes com diferentes momentos tiverem a mesma fase. Isso acontece, para $\varphi_-(k, x, t)$, onde estiver satisfeita a relação

$$0 \simeq \left. \frac{d\varphi_-}{dk} \right|_{k \simeq \bar{k}} = -\frac{\hbar \bar{k}}{M} t - x \quad \text{ou} \quad x \simeq -\bar{v}t.$$

enquanto que, para $\varphi_+(k, x, t)$,

$$0 \simeq \left. \frac{d\varphi_+}{dk} \right|_{k \simeq \bar{k}} = -\frac{\hbar \bar{k}}{M} t + 2 \left. \frac{d\delta}{dk} \right|_{k \simeq \bar{k}} + x \quad \text{ou} \quad x \simeq \bar{v} t - 2 \left. \frac{d\delta}{dk} \right|_{k \simeq \bar{k}}.$$

Como $x > 0$, a primeira dessas condições só pode ser satisfeita para $t < 0$, e representa um pacote de ondas incidente, que se aproxima da origem com velocidade $\bar{v} = \hbar \bar{k}/M$, e em condições de atingi-la em $t = 0$. Não há, portanto, contribuição apreciável do termo envolvendo $\varphi_-(k, x, t)$ para $t > 0$. A condição envolvendo $\varphi_+(k, x, t)$, por outro lado, representa um pacote de ondas que se afasta da origem também com velocidade \bar{v} . Para identificar a sua posição como função do tempo é conveniente reescrever o termo que contém a defasagem como

$$2 \left. \frac{d\delta}{dk} \right|_{k \simeq \bar{k}} = 2 \left. \frac{d\delta}{dE} \frac{dE}{dk} \right|_{k \simeq \bar{k}} = 2\bar{v} \hbar \left. \frac{d\delta}{dE} \right|_{k \simeq \bar{k}}$$

de modo que

$$x \simeq \bar{v} \left(t - 2\hbar \left. \frac{d\delta}{dE} \right|_{k \simeq \bar{k}} \right).$$

Essa posição corresponde a um pacote que se desloca na direção positiva x e que parte da origem em $t = 2\hbar d\delta/dE$. Este pacote pode ser visto como resultante da reflexão do pacote incidente pelo degrau de potencial, que sofre, durante o processo de reflexão, um *atraso* dado por

$$\tau \simeq 2\hbar \left. \frac{d\delta}{dE} \right|_{k \simeq \bar{k}}.$$

Quando a defasagem $\delta(k)$ cresce com a energia (como acontece no presente caso), $\tau > 0$. Existem no entanto casos envolvendo outros potenciais que levam a $\tau < 0$. Isso representa um *adiantamento* do pacote refletido, com relação ao tempo livre de trânsito de ida e volta à origem.

O problema de um degrau de potencial em uma dimensão é talvez a versão mais simples de um problema de espalhamento em mecânica quântica. Ele envolve dois *canais*, representados pelas regiões $x > 0$ e $x < 0$ respectivamente, e a altura V_0 corresponde ao “valor Q ” que caracteriza a inelasticidade dos processos de transição entre os canais. As soluções $\phi_E^{(0)}$ e $\phi_E^{(2)}$ contém uma onda incidente no canal $x > 0$. Na primeira delas, o canal $x < 0$ é um *canal fechado*, isto é, energeticamente proibido, devido ao valor de Q . Em $\phi_E^{(2)}$, porém os dois canais são canais abertos (energeticamente permitidos) e a onda

incidente dá lugar às ondas transmitida e refletida, que correspondem a ondas emergentes do potencial espalhador nos canais inelástico e elástico, respectivamente. Quando há um único canal aberto ($0 < E < V_0$) o único processo possível é o de espalhamento elástico, e a amplitude emergente só pode diferir da incidente por uma fase. Esse fato exprime a conservação do fluxo de probabilidade, nesse caso. No caso em que há dois canais abertos, o fluxo de probabilidade incidente é distribuído entre eles. A conservação do fluxo de probabilidade se exprime então em termos das amplitudes emergentes nos dois canais através da relação $R + (k/k_0)T = 1$.

A questão da completeza das autofunções do hamiltoniano, com as condições de contorno de espalhamento usadas acima, é na realidade um caso particular do problema geral de completeza encontrado em problemas de espalhamento. Neste caso, ele pode também ser abordado do ponto de vista das propriedades da equação diferencial de autovalores. Esse problema não será tratado aqui.

2. Poço de potencial. O problema de uma partícula de massa M em uma dimensão espacial, sujeita a um potencial dado por (v. figura 3.2)

$$V(x) = \begin{cases} +\infty & \text{para } x \leq 0 \\ -V_0 \ (V_0 > 0) & \text{para } 0 < x \leq a \\ 0 & \text{para } x > a \end{cases}$$

é útil tanto para ilustrar a aplicabilidade das idéias gerais introduzidas no tratamento do problema mais simples do degrau de potencial como para exemplificar o aparecimento de uma parte discreta no espectro do hamiltoniano. Como no caso do item anterior, o procedimento consistirá em obter as autofunções (eventualmente impróprias) e autovalores do hamiltoniano, em termos do que é possível obter a solução de problemas dependentes do tempo, através de argumentos de completeza das autofunções.

A equação de autovalores é, neste caso,

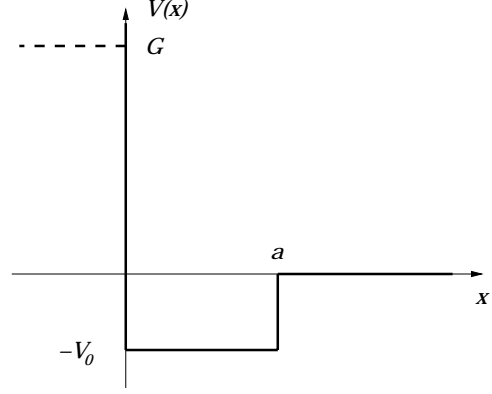
$$\begin{aligned} \frac{d^2 \phi_E}{dx^2} + K^2 \phi_E &= 0, & 0 < x \leq a; & & K^2 &\equiv \frac{2M}{\hbar^2}(E + V_0) \\ \frac{d^2 \phi_E}{dx^2} + k^2 \phi_E &= 0, & x > a; & & k^2 &\equiv \frac{2M}{\hbar^2}E \end{aligned}$$

cujas soluções gerais são dadas por combinações lineares do tipo⁶

$$\phi_E(0 < x \leq a) = A_+ e^{iKx} + A_- e^{-iKx} \quad \text{e} \quad \phi_E(x > a) = a_+ e^{ikx} + a_- e^{-ikx}.$$

⁶V. nota à página 153.

Figura 3.2: Poço de potencial de largura a e profundidade V_0 . O comportamento da função de onda na parede infinita em $x = 0$ pode ser obtida a partir do limite da função de onda para um degrau de altura G , fazendo $G \rightarrow \infty$.



As combinações admissíveis das quatro constantes são limitadas por condições que garantam a adequação da solução nos pontos especiais $x = 0$ e $x = a$. Essas condições podem ser obtidas a partir do que já foi feito com relação ao tratamento do degrau de potencial, considerando o potencial $V(x)$ como o limite, para $G \rightarrow +\infty$, do potencial $V_G(x)$ dado por

$$V_G(x) = \begin{cases} G & \text{para } x \leq 0 \\ -V_0 & \text{para } 0 < x \leq a \\ 0 & \text{para } x > a \end{cases}.$$

Este potencial, também representado na figura 3.2, pode ser visto como a combinação de dois degraus de potencial, nos quais deve ser imposta a continuidade da função de onda e de sua derivada. Além disso, eliminando as soluções exponencialmente crescentes para $x \rightarrow -\infty$, vem, no limite em que $G \rightarrow +\infty$, $\phi_E(x = 0) \rightarrow 0$, como pode ser facilmente verificado a partir da solução $\phi_E^{(0)}(x)$ do degrau de potencial. As condições que devem ser impostas em $x = 0$ e $x = a$ são, então,

$$\phi_E(0) = 0 \quad \text{logo} \quad A_+ = -A_- \equiv \frac{A}{2i}$$

e

$$\begin{aligned} A \sen Ka &= a_+ e^{ika} + a_- e^{-ika} \\ AK \cos Ka &= ik (a_+ e^{ika} - a_- e^{-ika}). \end{aligned}$$

Existem portanto duas condições para as três constantes remanescentes, A , a_+ e a_- , após a implementação da condição em $x = 0$. Da mesma forma que no caso do degrau de potencial, porém, uma delas é fixada pelo critério de normalização que seja adotado, dado o caráter

homogêneo da equação diferencial, de modo que as duas últimas condições na realidade fixam univocamente a solução (a menos de uma constante multiplicativa global).

Para obter efetivamente as soluções é preciso distinguir dois casos. No primeiro, $E > 0$, de modo que as soluções têm caráter oscilatório em todo o semi-eixo x positivo. Fixando a normalização com a condição $a_- = 1$ (isto é, uma onda incidente com amplitude igual a 1), existe uma única solução para cada valor de $E > 0$, dada pelas relações

$$A = \frac{e^{-ika}}{\sin Ka} (1 - e^{2i\varphi(k)}) \quad \text{com} \quad \tan \varphi(k) = \frac{k}{K} \tan Ka$$

e

$$a_+ = -e^{2i\delta(k)} \quad \text{com} \quad \delta(k) = \varphi(k) - ka.$$

Essas soluções podem ser interpretadas nas mesmas linhas que as soluções $\phi_E^{(0)}(x)$ obtidas no caso do degrau de potencial. Elas descrevem situações de espalhamento com conservação do fluxo de probabilidade, e a dependência com a energia da defasagem $\delta(k)$ afeta a evolução temporal de pacotes de onda após a sua colisão com o poço de potencial. Deve ser notado que, neste caso, a defasagem $\delta(k)$ inclui efeitos devidos à região *interna*, $x < a$, do poço de potencial através de $\varphi(k)$.

A constante A representa a amplitude da função de onda na região interna do poço, $x < a$, e é interessante estudar o seu comportamento como função de E (ou K). Pode ser verificado facilmente que

$$|A|^2 = \frac{4E}{E + V_0 \cos^2 Ka}.$$

Essa quantidade, vista como função de Ka , tem um comportamento oscilatório, com máximos (que correspondem a máximos na probabilidade de encontrar a partícula na região $x < a$) para valores de Ka próximos de múltiplos ímpares de $\pi/2$ e mínimos para valores de Ka próximos de múltiplos pares de $\pi/2$.

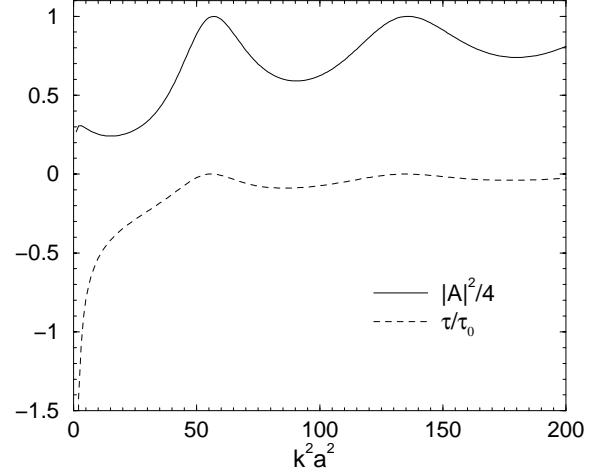
O comportamento de $|A|^2$ com Ka pode ser correlacionado com o comportamento da defasagem $\delta(k)$ como função de k . De fato,

$$\frac{d\delta}{dE} = \frac{M}{\hbar^2 k} \frac{d\delta}{dk} = \frac{M}{\hbar^2 k} \left(\frac{d\varphi}{dk} - a \right)$$

e

$$\frac{d\varphi}{dk} = \frac{aE}{E + V_0 \cos^2 Ka} \left(\frac{V_0}{E} \frac{\sin 2Ka}{2Ka} + 1 \right) = \frac{a|A|^2}{4} \left(\frac{V_0}{E} \frac{\sin 2Ka}{2Ka} + 1 \right).$$

Figura 3.3: Quadrado da amplitude da parte interna da função de onda $\phi_E(x)$, $|A|^2$, e tempo de atraso, τ , em unidades de $\tau_0 = Ma^2/\hbar$, como função de $k^2a^2 = 2MEa^2/\hbar^2$, para um poço quadrado com $2MV_0a^2/\hbar^2 = 64$.



Essas expressões mostram que o atraso $\tau = 2\hbar d\delta/dE$ contém uma contribuição negativa proveniente do termo $-ka$ da defasagem, além da contribuição proveniente de $\varphi(k)$. A primeira dessas duas contribuições tende a produzir um adiantamento do pacote refletido, devido à reflexão que ocorre na borda externa do potencial, em $x = a$. A contribuição que envolve $d\varphi/dk$ é positiva e tem um comportamento oscilatório como função de k (ou de E), e pode ser interpretado em termos da variação com k do “tempo médio de permanência” no interior do poço de potencial. Essa interpretação é corroborada pela proporcionalidade de $d\varphi/dk$ ao quadrado da amplitude da função de onda no interior do poço de potencial, na medida que isso mostra que um aumento na probabilidade de encontrar a partícula na região $x < a$ tende a aumentar a contribuição positiva para o atraso do pacote refletido. Um exemplo dos resultados para o quadrado do módulo da amplitude interna $|A|^2$ e para o tempo de atraso τ no caso de um poço de potencial com $2MV_0a^2/\hbar^2 = 64$ é mostrado na figura 3.3.

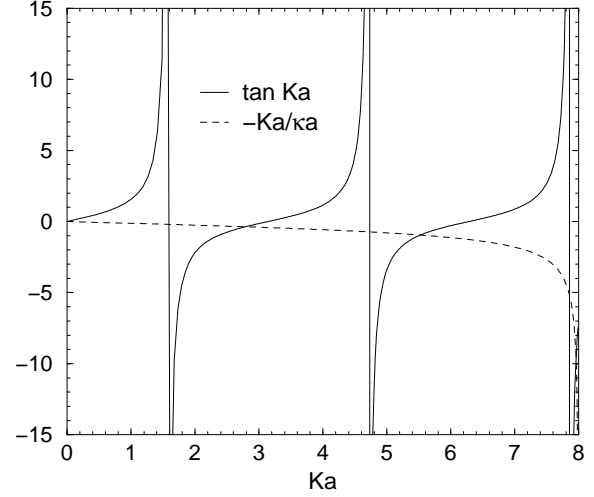
Como não existem soluções limitadas para $E < -V_0$, resta apenas considerar o caso $-V_0 < E < 0$. Aqui, para $x < a$ as soluções são ainda oscilatórias, enquanto que para $x > a$ elas são exponenciais reais, crescentes e decrescentes. Argumentos análogos aos usados no caso do degrau de potencial levam à eliminação das soluções exponencialmente crescentes, de modo que as soluções limitadas são do tipo

$$\phi_E(x) = \begin{cases} A \sin Kx, & x < a, & K^2 = \frac{2M}{\hbar^2}(V_0 + E) \\ a_- e^{-\kappa x}, & x > a, & \kappa^2 = -\frac{2ME}{\hbar^2} \end{cases}$$

sendo que as condições de continuidade em $x = a$ levam às equações

$$A \sin Ka = a_- e^{-\kappa a}$$

Figura 3.4: Solução gráfica da equação (3.38), que determina os estados ligados do poço de potencial, para o mesmo caso usado na figura 3.3, em que $2MV_0a^2/\hbar^2 = 64$. O gráfico inclui as assíntotas de $\tan Ka$, e $\kappa a \rightarrow 0$ para $Ka \rightarrow 8$, que corresponde a $E \rightarrow 0$. Há três estados ligados, com o terceiro na parte direita e inferior da figura.



$$KA \cos Ka = -\kappa a_- e^{-\kappa a}$$

Lembrando que uma das constantes A , a_- é fixada por uma condição de normalização, é claro que essas *duas* equações só podem ser satisfeitas para valores particulares de E , ou, equivalentemente, dos vetores de onda K e κ . O espectro é portanto *discreto* nesse intervalo de energias, e os autovalores do hamiltoniano são dados pelas soluções da equação que resulta da eliminação de a_-/A entre as duas equações resultantes das condições de continuidade em $x = a$

$$\tan Ka = -\frac{Ka}{\kappa a}, \quad -V_0 < E < 0. \quad (3.38)$$

A solução trivial $Ka = 0$ (que corresponde a $E = -V_0$) é excluída por levar a uma função de onda identicamente nula. Para cada uma das soluções positivas $K_n a$, $n = 1, 2, \dots$ dessa equação o valor da razão a_-/A é dado por

$$\left(\frac{a_-}{A}\right)_n = e^{\kappa_n a} \sin K_n a$$

e a normalização é fixada escolhendo o valor de A_n de modo a satisfazer a condição

$$1 = \int_0^\infty |\phi_{E_n}(x)|^2 dx = |A_n|^2 \left(\int_0^a \sin^2 K_n x dx + e^{2\kappa_n a} \sin^2 K_n a \int_a^\infty e^{-2\kappa_n x} dx \right).$$

O lado direito da equação (3.38) é sempre negativo, e se aproxima assintoticamente de $-\infty$ quando κ (ou E) tende a zero (v. exemplo na figura 3.4). Nesse limite, por outro lado, $K \rightarrow K_0 \equiv (2MV_0/\hbar^2)^{1/2}$. Tendo em conta o comportamento de $\tan Ka$, resulta que o número

de estados ligados admitidos pelo poço de potencial é dado pelo número de ramos negativos da tangente quando o seu argumento varia de zero a K_0a . Portanto, para que haja pelo menos um estado ligado é necessário que $K_0a > \pi/2$, ou seja

$$(K_0a)^2 = \frac{2M}{\hbar^2}V_0a^2 > \frac{\pi^2}{4} \quad \text{ou ainda} \quad V_0a^2 > \frac{\pi^2\hbar^2}{8M}.$$

A condição para que haja pelo menos n estados ligados é, analogamente, $K_0a > (2n - 1)\pi/2$, ou

$$V_0a^2 > (2n - 1)^2 \frac{\pi^2\hbar^2}{8M}.$$

Essas relações mostram também que o número de estados ligados não se altera quando a largura a e a profundidade V_0 do poço de potencial são variados de modo que V_0a^2 permaneça constante.

Problemas.

3.1 - Escreva a energia cinética e a energia potencial de um oscilador harmônico de massa M e frequência ω_0 em termos de a e a^\dagger , na descrição de Schrödinger. Use esses resultados para calcular os valores médios da energia cinética e da energia potencial em um autovetor $|\nu\rangle$ do hamiltoniano, e verifique que eles satisfazem o teorema do virial (cf. problema 2.7).

3.2 - Resolva o problema de condições iniciais para um oscilador harmônico livre, de massa M e frequência ω_0 , que se encontra inicialmente no estado $|t = 0\rangle = (|\nu_1\rangle + |\nu_2\rangle)/\sqrt{2}$ onde $|\nu_1\rangle$ e $|\nu_2\rangle$ são autovetores do hamiltoniano com $\nu_1 \neq \nu_2$. Calcule o valor médio do observável $P_1(t = 0) \equiv |\phi_-\rangle\langle\phi_-|$, sendo o estado $|\phi_-\rangle$ definido como $|\phi_-\rangle = (|\nu_1\rangle - |\nu_2\rangle)/\sqrt{2}$, em função de t para $t > 0$. Relacione a escala de tempo em que esse valor médio varia (obtida simplesmente por inspeção do valor médio de P_1 como função do tempo) com a incerteza em energia do estado considerado, e confronte o resultado com a relação de incerteza energia-tempo. **Sugestão:** a título de exercício, trate o problema duas vezes, usando as descrições de Schrödinger e de Heisenberg, respectivamente.

3.3 - Obtenha a forma explícita da solução do problema de condições iniciais para o oscilador forçado que se encontra inicialmente no seu estado fundamental para o caso de uma força externa impulsiva no instante $t = t_0 > 0$ dada por $f(t) = f_0\delta(t - t_0)$, onde f_0 é uma constante real e a função $\delta(t - t_0)$ é a função delta de Dirac. Trate o problema tanto na descrição de Heisenberg como também na de Schrödinger.

3.4 - Mostre que a probabilidade de obter o valor $(\nu + 1/2)\hbar\omega_0$ numa medição da energia de um oscilador harmônico que se encontra no estado coerente $|z\rangle$ tal que $a|z\rangle = z|z\rangle$ tem uma distribuição *poissoniana* em ν (lembre-se que uma distribuição poissoniana é uma distribuição do tipo $e^{-a}a^\nu/\nu!$).

3.5 - a) Mostre que dois estados coerentes $|z_1\rangle$ e $|z_2\rangle$, com $z_1 \neq z_2$, não são ortogonais, calculando explicitamente o produto escalar $\langle z_2|z_1\rangle$.

b) Verifique a resolução da unidade em termos de estados coerentes

$$\hat{1} = \int |z\rangle \frac{d^2z}{\pi} \langle z|$$

onde a integração é estendida a todo o plano complexo z . **Sugestão:** escreva os estados coerentes do integrando em termos dos autovetores $|\nu\rangle$ do hamiltoniano e faça a integração usando coordenadas polares no plano complexo. Verifique desse modo que a integral se reduz à soma $\sum_\nu |\nu\rangle\langle\nu|$.

c) Mostre que a resolução da unidade estabelecida em **b** permite escrever um estado coerente $|z\rangle$ como uma superposição de estados coerentes $|z'\rangle$ com valores de z' que podem ser diferentes de z . Isso mostra que o conjunto de todos os estados coerentes não é linearmente independente, e por essa razão a resolução da unidade mencionada em **b)** é chamada uma relação de *supercompleteza*.

3.6 - a) Resolva o problema de condições iniciais para um oscilador harmônico com a condição inicial $|t=0\rangle = |z_0\rangle$, onde $|z_0\rangle$ é um estado coerente. Use a descrição de Schrödinger e mostre que $|t>0\rangle = |z(t)\rangle$, é um estado também coerente, dependente do tempo.

b) Calcule os valores médios do momento p e da posição q em um estado coerente genérico $|z\rangle$. Use os resultados para obter os valores médios de p e q como função do tempo a partir do estado $|z(t)\rangle$ obtido em **a)**. Compare $\langle p\rangle(t)$ e $\langle q\rangle(t)$ assim obtidos com $p(t)$ e $q(t)$ de um oscilador *clássico*.

3.7 - Mostre que um estado coerente $|z\rangle$ é um estado de incerteza mínima para qualquer valor de z . **Sugestão:** Calcule os valores médios de $(p - \langle p\rangle)^2$ e de $(q - \langle q\rangle)^2$ no estado coerente, com $\langle p\rangle = \langle z|p|z\rangle$ e uma definição análoga para $\langle q\rangle$, mostrando em particular que o produto das incertezas não depende de z .

3.8 - Como os estados coerentes $|z\rangle$ são estados de incerteza mínima, nos quais os valores médios dos operadores de posição e momento são dados por $\langle q\rangle = bQ$ e $\langle p\rangle = \hbar P/b$, com $b = \sqrt{\hbar/M\omega_0}$ e $z \equiv (Q + iP)/\sqrt{2}$, a quantidade $f(Q, P) \equiv |\langle z|\varphi\rangle|^2$ é útil como representação de como a

densidade de probabilidade associada ao estado $|\varphi\rangle$ se distribui num *espaço de fase* $Q \times P$ *clássico*. A função $f(Q, P)$ que representa vetores de estado $|\varphi\rangle$ num tal espaço de fase clássico é conhecido pelo nome de “função-Q”, ou também “função de Husimi”.

a) Calcule a função de Husimi $f_\nu(Q, P) = |\langle z|\nu\rangle|^2$ correspondente ao estado estacionário de energia $\hbar\omega_0(\nu + 1/2)$ de um oscilador harmônico unidimensional. Estude o perfil dessa função, comparando-o com a órbita de mesma energia (no espaço de fase) de um oscilador clássico.

b) Calcule a função de Husimi $f_{z_0}(Q, P) = |\langle z|z_0\rangle|^2$ correspondente *ao estado coerente* $|z_0\rangle$, e compare-a com a que corresponde ao estado fundamental $|\nu = 0\rangle$, obtida em **a)**.

3.9 - Considere a probabilidade de transição

$$p_{1\rightarrow 2}(t) = |\langle \phi_2|U(t, 0)|\phi_1\rangle|^2$$

onde $U(t, 0)$ é o operador de evolução associado ao hamiltoniano de um sistema de dois níveis que tem a forma

$$H = \begin{pmatrix} 0 & \gamma \\ \gamma^* & \Delta \end{pmatrix}$$

e $\{|\phi_i\rangle\}$, $i = 1, 2$ são os vetores da base ortonormal usada para representar H . Mostre que $p_{1\rightarrow 2}(t)$ *não depende de* Δ para tempos *muito curtos*, no sentido que $(E_+ - E_-)t/\hbar \ll 1$, sendo E_\pm os autovalores de H . **Observação:** Note que $p_{1\rightarrow 2}(t)$ tem um comportamento oscilatório no tempo com uma amplitude que *diminui* e com uma frequência que *aumenta* quando Δ aumenta. O fato de que, para tempos curtos, $p_{1\rightarrow 2}(t)$ não depende de Δ significa então que, nessa escala de tempo, esses dois efeitos na realidade se compensam.

Haveria um argumento capaz de dar conta desse comportamento com base na relação de incerteza energia-tempo?

3.10 - Deduza o Teorema de Ehrenfest usando a descrição de Heisenberg da dinâmica quântica. Nessa descrição os operadores dependentes do tempo $\vec{q}(t)$ e $\vec{p}(t)$ satisfazem equações de movimento

$$i\hbar \frac{d\vec{q}(t)}{dt} = [\vec{q}(t), H]$$

e uma equação análoga para $\vec{p}(t)$. Suponha que o hamiltoniano não dependa explicitamente do tempo e tenha a forma

$$H = \frac{\vec{p}^2}{2M} + V(\vec{q}),$$

podendo ser portanto expresso com essa mesma forma em termos dos operadores da descrição de Heisenberg $\vec{q}(t)$ e $\vec{p}(t)$. **Sugestão:** Use o fato de que a unitariedade da transformação que leva os operadores da descrição de Schrödinger aos da descrição de Heisenberg garante que as relações de comutação entre eles é preservada. Dessa forma, não é necessário resolver as equações de movimento de Heisenberg.

3.11 - O operador q pode ser representado em termos de seus “autovetores” impróprios pela “matriz” diagonal $\langle x|q|x'\rangle = x\delta(x-x')$, onde esta última função é a função delta de Dirac. Isso é consistente com a representação de q pelo operador de multiplicação por x no espaço de funções de onda, pois

$$\langle x|q|\varphi\rangle = \int_{-\infty}^{+\infty} dx' \langle x|q|x'\rangle \langle x'|\varphi\rangle = \int_{-\infty}^{+\infty} dx' x\delta(x-x')\varphi(x') = x\varphi(x).$$

a) Mostre que a representação “matricial” correspondente para o operador p pode ser escrita como

$$\langle x|p|x'\rangle = -i\hbar \delta'(x-x')$$

onde $\delta'(x-x')$ é a derivada da função delta de Dirac com relação ao seu argumento, $d\delta(x-x')/dx$.

Sugestão: proceda com o operador p de forma análoga à usada acima para o operador q , e faça a integração envolvendo $\delta'(x-x')$ por meio de uma integração por partes. Note que a função delta de Dirac é uma “função par” no sentido que $\delta(x) = \delta(-x)$ e que a sua derivada é uma “função ímpar”, $\delta'(x) = -\delta'(-x)$.

b) Tomando a representação “matricial” nos autovetores impróprios de q da relação de comutação $i\hbar\hat{1} = [q, p]$ o que se obtém é

$$i\hbar \delta(x-x') = \langle x|[q, p]|x'\rangle = (x-x')\langle x|p|x'\rangle = -i\hbar (x-x')\delta'(x-x').$$

Mostre que essa relação é consistente no sentido que, para qualquer função de onda $\varphi(x)$ contínua no ponto x ,

$$\int_{-\infty}^{+\infty} dx' i\hbar \delta(x-x')\varphi(x') = - \int_{-\infty}^{+\infty} dx' i\hbar (x-x')\delta'(x-x')\varphi(x').$$

3.12 - Para estudar a variação com o tempo da largura de um pacote de ondas livre *qualquer* em uma dimensão, que evolui portanto de acordo com a equação de Schrödinger dependente do tempo

$$i\hbar \frac{d|\varphi(t)\rangle}{dt} = \frac{p^2}{2M}|\varphi(t)\rangle$$

a partir de uma condição inicial dada $|\varphi(t=0)\rangle$, mostre, primeiramente, que valem as relações

$$\begin{aligned} \text{i)} \quad & \frac{d\langle q \rangle}{dt} = \frac{\langle p \rangle}{M} \\ \text{ii)} \quad & \frac{d\langle p \rangle}{dt} = 0 \\ \text{iii)} \quad & \frac{d\langle q^2 \rangle}{dt} = \frac{1}{M} \langle (qp + pq) \rangle \\ \text{iv)} \quad & \frac{d\langle (qp + pq) \rangle}{dt} = \frac{2}{M} \langle p^2 \rangle \quad \text{e} \\ \text{v)} \quad & \frac{d\langle p^2 \rangle}{dt} = 0. \end{aligned}$$

onde a notação $\langle \rangle$ indica o valor médio tomado no estado dependente do tempo $|\varphi(t)\rangle$ que descreve o pacote.

Mostre em seguida que, das relações **iii)** a **v)**, decorre que a dependência com o tempo de $\langle q^2 \rangle$ é dada por

$$\langle q^2 \rangle = A + Bt + \frac{C}{2}t^2$$

com $A = \langle q^2 \rangle_{t=0}$, $B = (1/M)\langle (qp + pq) \rangle_{t=0}$ e $C = (2/M^2)\langle p^2 \rangle$, enquanto das relações **i)** e **ii)** decorre que $\langle q \rangle = \langle q \rangle_{t=0} + (1/M)\langle p \rangle t$.

Use então a dependência temporal assim determinada de $\langle q^2 \rangle$ e de $\langle q \rangle$ para analisar a dependência temporal da dispersão $\sigma_q^2(t) \equiv \langle q^2 \rangle - \langle q \rangle^2$. Considere especialmente a dependência da dispersão em momento do pacote, $\sigma_p^2 \equiv \langle p^2 \rangle - \langle p \rangle^2$, e da covariância $\langle (qp + pq) \rangle$. Note que esta última quantidade depende do tempo!

3.13 - Mostre que a função de onda de uma partícula livre no espaço infinito pode ser expressa em termos da função de onda num ponto diferente do espaço-tempo como

$$\varphi(\vec{r}, t) = \int d^3r' U(\vec{r}, t; \vec{r}', t_0) \varphi(\vec{r}', t_0)$$

onde

$$U(\vec{r}, t; \vec{r}', t_0) = \frac{1}{(2\pi)^3} \int d^3k e^{i\vec{k} \cdot (\vec{r} - \vec{r}') - i\omega_k(t - t_0)} = \left(\frac{M}{2\pi i \hbar t} \right)^{\frac{3}{2}} e^{\frac{iM(\vec{r} - \vec{r}')^2}{2\hbar t}}$$

com $\omega_k = \hbar k^2 / 2M$.

- Sugestões:** i) $\varphi(\vec{r}, t) = e^{-iHt/\hbar} \varphi(\vec{r}, 0)$;
 ii) $\varphi(\vec{r}, t) = \int d^3r' \delta(\vec{r} - \vec{r}') \varphi(\vec{r}', t)$;
 iii) $\delta(\vec{r} - \vec{r}') = \frac{1}{(2\pi)^3} \int d^3k e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}.$

3.14 - Reduza a relação do problema anterior ao caso de uma partícula livre em uma dimensão. Obtenha a função de onda que descreve no tempo t uma partícula que, em $t = 0$, está no estado descrito pela função de onda gaussiana

$$\varphi(x, t = 0) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{1}{2}} e^{ikx} e^{-\frac{x^2}{2b^2}}.$$

3.15 - Considere o problema unidimensional de uma partícula de massa M sujeita ao potencial dado por

$$V(x) = \begin{cases} +\infty & \text{para } x \leq 0 \\ V_0 > 0 & \text{para } 0 < x < a \\ 0 & \text{para } x \geq a. \end{cases}$$

Obtenha as autofunções do hamiltoniano e estude o atraso $\tau = 2\hbar d\delta/dE$ tanto na região de energia $0 < E < V_0$ quanto na região $E > V_0$.

Referências gerais.

Os tópicos tratados neste capítulo são discutidos também nos textos já citados nos capítulos anteriores. Um capítulo muito útil sobre propriedades e usos da representação de Wigner se encontra num livro cujo foco principal não é a mecânica quântica:

★ **S. R. de Groot e L. G. Suttorp**, *Foundations of Electrodynamics*, North-Holland Publishing Company (Amsterdam) 1972, Capítulo VI.

Capítulo 4

Métodos de aproximação.

A possibilidade de obter a solução exata para um problema de mecânica quântica é, em termos práticos, uma situação bastante excepcional. Isso confere importância considerável ao desenvolvimento, sistematização e uso de métodos de aproximação adaptados a diversos tipos de problemas. Este capítulo trata, da forma mais simples e direta possível, de algumas técnicas de aproximação de uso freqüente, e da sua aplicação típica a exemplos que ilustrem tanto as suas virtudes quanto as suas limitações. A existência, geralmente concomitante, de ambas, faz com que a frase de Griffin e Wheeler segundo a qual essas técnicas “não constituem uma maquinaria mágica que pode ser operada sem o exercício de julgamento” (*Phys. Rev.* **108**, 311 (1957)) seja sempre pertinente.

Os métodos de aproximação utilizados na mecânica quântica podem ser divididos em dois grandes grupos, geralmente chamados *métodos perturbativos* e *métodos não perturbativos*. A base comum aos métodos chamados perturbativos é uma expansão em série de potências de algum parâmetro que possa ser considerado como “pequeno”. A solução exata para o sistema de dois níveis obtida na seção **3.2.1**, por exemplo, pode ser expandida em potências de $\lambda = |\gamma|/(\delta_2 + \Delta - \delta_1)$ quando $|\gamma| \ll (\delta_2 + \Delta - \delta_1)$, dando para as energias E_- e E_+ , do estado fundamental e do estado excitado

$$E_- \simeq \delta_1 - \frac{|\gamma|^2}{\delta_2 + \Delta - \delta_1} \quad \text{e} \quad E_+ \simeq \delta_2 + \Delta + \frac{|\gamma|^2}{\delta_2 + \Delta - \delta_1}$$

mais termos de ordem superior em λ . Uma expansão correspondente pode também ser escrita para os autovetores. O propósito dos métodos perturbativos de aproximação é estabelecer um conjunto de procedimentos que permita escrever expansões desse tipo independentemente do uso da solução exata, e que portanto possam ser aplicados quando ela não é disponível. No exemplo acima, a convergência da expansão pode ser inferida diretamente da forma (em particular das propriedades de analiticidade) da solução exata. No uso típico desses métodos,

em situações em que não se dispõe de uma solução exata, essa questão tende a ser bastante problemática.

O sub-grupo mais importante de métodos não perturbativos é a dos *métodos variacionais*. Eles são aplicáveis tipicamente quando a solução *exata* do problema considerado pode ser caracterizada através de uma condição de extremo, aplicada a um funcional apropriado de variáveis dinâmicas e vetores de estado, com relação a variações definidas dos ingredientes do funcional. Soluções aproximadas podem então ser obtidas restringindo o domínio das variações de acordo com condições efetivas de cálculo, independentemente da identificação de parâmetros “pequenos” para orientar a busca. As limitações comumente associadas ao uso de métodos de aproximação variacionais se relacionam com a dificuldade de avaliar quantitativamente a confiabilidade dos resultados obtidos. Em particular, eles tipicamente não fornecem procedimentos gerais ou sistemáticos para o cálculo de correções “mais finas”. Em outras palavras, os métodos variacionais são caracteristicamente poderosos mas de difícil controle.

Um exemplo tão simples quanto importante de um procedimento variacional é o que permite obter uma aproximação para o menor autovalor γ_0 de uma variável dinâmica qualquer g . Ele pode ser formulado considerando um vetor de estado *qualquer* $|\varphi\rangle$ e utilizando a resolução da unidade em termos dos autovetores $|\gamma\rangle$ de g para escreve-lo como

$$|\varphi\rangle = \sum_{\gamma} |\gamma\rangle \langle \gamma | \varphi \rangle \equiv \sum_{\gamma} |\gamma\rangle c_{\gamma}$$

onde o espectro de g foi tomado como sendo puramente discreto. Essa suposição não é no entanto necessária, e pode ser eliminada simplesmente usando formas mais gerais da resolução da unidade em que estejam incluídos eventuais autovetores impróprios associados ao espectro contínuo. O valor médio de g no vetor de estado $|\varphi\rangle$ (não necessariamente de norma 1) pode então ser escrito como

$$\frac{\langle \varphi | g | \varphi \rangle}{\langle \varphi | \varphi \rangle} = \frac{1}{\sum_{\gamma} |c_{\gamma}|^2} \sum_{\gamma} \gamma |c_{\gamma}|^2 \geq \gamma_0 \quad (4.1)$$

onde a última desigualdade resulta simplesmente do fato de que $\gamma \geq \gamma_0$, dado que γ_0 é o menor autovalor de g . É fácil ver também que a igualdade valerá apenas no caso em que $|c_{\gamma}|^2 = 0$ para $\gamma \neq \gamma_0$, ou seja, apenas no caso em que $|\varphi\rangle$ seja autovetor de g com autovalor γ_0 . Desse modo, a condição de autovetor de g com o menor autovalor γ_0 corresponde ao mínimo (absoluto) do funcional

$$G[|\varphi\rangle] \equiv \frac{\langle \varphi | g | \varphi \rangle}{\langle \varphi | \varphi \rangle}.$$

Uma aproximação variacional para γ_0 é obtida calculando efetivamente o valor desse funcional para alguma família de vetores de estado φ e determinando o menor valor que ele assume para os

vetores dessa família. Devido à relação (4.1), esse valor será sempre um *limite superior* para γ_0 , que será de fato igual a esse autovalor se a família escolhida incluir o autovetor correspondente. Quando a natureza deste autovetor e o espectro de g são mal conhecidos, é claro que não é possível avaliar quantitativamente a eventual discrepância entre o valor mínimo obtido para o funcional e o valor de γ_0 .

4.1 Métodos perturbativos.

Os métodos perturbativos de uso corrente na mecânica quântica se agrupam também em duas grandes famílias que correspondem respectivamente a problemas envolvendo equações de autovalores e a problemas envolvendo equações de movimento. A família correspondente ao tratamento de equações de movimento tem, por razões óbvias, a designação comum de *teoria de perturbações dependente do tempo*. Como os métodos perturbativos de aproximação para equações de autovalores são usualmente desenvolvidos com referência à equação de autovalores para o hamiltoniano (que identifica os estados estacionários do sistema considerado), eles são usualmente designados como *teoria de perturbações independente do tempo*.

Em termos um pouco mais específicos, a situação típica pode ser caracterizada da seguinte forma. O hamiltoniano do problema a ser tratado (que será tomado como a variável a ser considerada, no caso de equações de autovalores) é escrito como a soma de duas partes

$$H = H_0 + H_1.$$

No caso da teoria de perturbações independente do tempo, supõe-se que o problema de autovalores tenha sido resolvido para o “hamiltoniano não perturbado” H_0 , isto é, os autovalores $E_n^{(0)}$ e autovetores $|\phi_n^{(0)}\rangle$ que são soluções de

$$H_0|\phi_n^{(0)}\rangle = E_n^{(0)}|\phi_n^{(0)}\rangle$$

são supostamente conhecidos. A teoria desenvolve então procedimentos sistemáticos de expansão em série para aproximar as soluções de

$$(H_0 + H_1)|\phi_n\rangle = E_n|\phi_n\rangle$$

em termos dos vetores de estado $|\phi_n^{(0)}\rangle$, dos autovalores $E_n^{(0)}$ e da “perturbação” H_1 . A teoria de perturbações dependente do tempo, por outro lado, calcula sistematicamente correções devidas a H_1 para o problema de condições iniciais

$$(H_0 + H_1)|\varphi(t)\rangle = i\hbar \frac{\partial |\varphi(t)\rangle}{\partial t}, \quad |\varphi(t=0)\rangle = |\varphi_0\rangle$$

fazendo uso também dos autovalores e autovetores de H_0 . Em sua forma mais simples, as duas teorias são desenvolvidas especificamente para *espectros discretos*. Com uso suficiente de “julgamento”, e com interpretação cuidadosa de resultados, é possível estender o seu uso a situações envolvendo espectros contínuos, para as quais, no entanto, métodos mais adequados são desenvolvidos no contexto da teoria dos processos de espalhamento.

4.1.1 Perturbação de estados não degenerados em espectros discretos.

O tipo mais simples de expansão perturbativa é o que aproxima os efeitos de uma perturbação H_1 sobre um autovetor $|\phi_{n_0}^{(0)}\rangle$ correspondente ao autovalor *não degenerado* $E_{n_0}^{(0)}$ do hamiltoniano não perturbado H_0 . A quantidade λ que funciona neste caso como “parâmetro pequeno” para a expansão pode ser caracterizada, como se verá em seguida, em termos da razão entre o valor absoluto de elementos de matriz de H_1 que ligam esse autovetor a outros autovetores de H_0 e a diferença entre E_{n_0} e os autovalores associados a esses outros autovetores, isto é

$$\lambda \sim \mathcal{O} \left(\frac{|\langle \phi_{n_0}^{(0)} | H_1 | \phi_n^{(0)} \rangle|}{|E_{n_0}^{(0)} - E_n^{(0)}|} \right)$$

com os autovetores $|\phi_n^{(0)}\rangle$ devidamente normalizados. O caráter não degenerado de $E_{n_0}^{(0)}$ garante que o denominador dessa razão não seja nulo. Além disso, para que a expansão possa ser útil é necessário que a razão seja pequena, $\lambda \ll 1$. Quando essa condição não é satisfeita o autovalor $E_{n_0}^{(0)}$ é dito “quase-degenerado”, para efeitos da perturbação, exigindo então tratamento especial, a ser discutido adiante. Em vista da relação linear que existe entre λ e H_1 , para facilitar a contabilização das diferentes ordens da expansão perturbativa é conveniente escrever o hamiltoniano completo como

$$H = H_0 + \xi H_1$$

onde o *parâmetro de expansão* ξ é finalmente tomado como igual a 1.

Nessas condições, o efeito da perturbação pode ser visto como o de transformar a solução de

$$H_0 |\phi_{n_0}^{(0)}\rangle = E_{n_0}^{(0)} |\phi_{n_0}^{(0)}\rangle, \quad \langle \phi_{n_0}^{(0)} | \phi_{n_0}^{(0)} \rangle = 1 \quad (4.2)$$

na solução de

$$(H_0 + \xi H_1) |\phi_{n_0}\rangle = E_{n_0} |\phi_{n_0}\rangle, \quad \langle \phi_{n_0} | \phi_{n_0} \rangle = 1 \quad (4.3)$$

tal que, se $\xi \rightarrow 0$, o vetor $|\phi_{n_0}\rangle$ se reduz a $|\phi_{n_0}^{(0)}\rangle$ e o autovalor E_{n_0} se reduz a $E_{n_0}^{(0)}$. É conveniente então escrever o autovetor perturbado sob a forma

$$|\phi_{n_0}\rangle = N_{n_0}(|\phi_{n_0}^{(0)}\rangle + |\chi_{n_0}\rangle) \quad (4.4)$$

onde N_{n_0} é um coeficiente de normalização e a componente adicional $|\chi_{n_0}\rangle$ introduzida pela perturbação pode ser tomada, sem perda de generalidade, como sendo ortogonal ao estado não perturbado, $\langle\chi_{n_0}|\phi_{n_0}^{(0)}\rangle = 0$.

Substituindo a decomposição (4.4) na equação de autovalores (4.3) resulta

$$(H_0 + \xi H_1)(|\phi_{n_0}^{(0)}\rangle + |\chi_{n_0}\rangle) = E_{n_0}(|\phi_{n_0}^{(0)}\rangle + |\chi_{n_0}\rangle)$$

ou ainda, usando (4.2),

$$(E_{n_0} - E_{n_0}^{(0)})|\phi_{n_0}^{(0)}\rangle = \xi H_1|\phi_{n_0}^{(0)}\rangle + (H_0 + \xi H_1 - E_{n_0})|\chi_{n_0}\rangle.$$

O produto escalar dessa relação com $|\phi_{n_0}^{(0)}\rangle$, tendo em conta a ortogonalidade desse vetor a $|\chi_{n_0}\rangle$, dá

$$E_{n_0} - E_{n_0}^{(0)} = \xi(\langle\phi_{n_0}^{(0)}|H_1|\phi_{n_0}^{(0)}\rangle + \langle\phi_{n_0}^{(0)}|H_1|\chi_{n_0}\rangle). \quad (4.5)$$

Por outro lado, o produto escalar da mesma relação com $|\phi_m^{(0)}\rangle$, $m \neq n_0$, tendo em conta a ortogonalidade desse vetor a $|\phi_{n_0}^{(0)}\rangle$, dá

$$\begin{aligned} \langle\phi_m^{(0)}|\chi_{n_0}\rangle &= \xi \left(\frac{\langle\phi_m^{(0)}|H_1|\phi_{n_0}^{(0)}\rangle + \langle\phi_m^{(0)}|H_1|\chi_{n_0}\rangle}{E_{n_0} - E_m^{(0)}} \right) \\ &= \xi \frac{\langle\phi_m^{(0)}|H_1|\phi_{n_0}^{(0)}\rangle}{E_{n_0} - E_m^{(0)}} + \xi \sum_{m' \neq n_0} \frac{\langle\phi_m^{(0)}|H_1|\phi_{m'}^{(0)}\rangle}{E_{n_0} - E_m^{(0)}} \langle\phi_{m'}^{(0)}|\chi_{n_0}\rangle, \quad m \neq n_0. \end{aligned} \quad (4.6)$$

Uma expansão para a componente adicional do autovetor perturbado pode agora ser obtida iterando a relação (4.6)

$$\begin{aligned} |\chi_{n_0}\rangle &= \sum_{m \neq n_0} |\phi_m^{(0)}\rangle \langle\phi_m^{(0)}|\chi_{n_0}\rangle = \\ &= \xi \sum_{m \neq n_0} |\phi_m^{(0)}\rangle \frac{\langle\phi_m^{(0)}|H_1|\phi_{n_0}^{(0)}\rangle}{E_{n_0} - E_m^{(0)}} + \\ &+ \xi^2 \sum_{m \neq n_0} \sum_{m' \neq n_0} |\phi_m^{(0)}\rangle \frac{\langle\phi_m^{(0)}|H_1|\phi_{m'}^{(0)}\rangle \langle\phi_{m'}^{(0)}|H_1|\phi_{n_0}^{(0)}\rangle}{(E_{n_0} - E_m^{(0)})(E_{n_0} - E_{m'}^{(0)})} + \\ &+ \dots \end{aligned} \quad (4.7)$$

e, substituindo essa série em (4.5), resulta

$$\begin{aligned}
E_{n_0} &= E_{n_0}^{(0)} + \xi \langle \phi_{n_0}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle + \\
&+ \xi^2 \sum_{m \neq n_0} \frac{\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{E_{n_0} - E_m^{(0)}} + \\
&+ \xi^3 \sum_{m \neq n_0} \sum_{m' \neq n_0} \frac{\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{m'}^{(0)} \rangle \langle \phi_{m'}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{(E_{n_0} - E_m^{(0)})(E_{n_0} - E_{m'}^{(0)})} + \\
&+ \dots
\end{aligned} \tag{4.8}$$

As duas séries (4.7) e (4.8) constituem a chamada *expansão perturbativa de Brillouin-Wigner*. Os elementos de matriz que aparecem nos numeradores envolvem apenas os autovetores de H_0 , e podem portanto ser calculados uma vez que o problema de autovalores para esse Hamiltoniano tenha sido resolvido, mas os denominadores envolvem o autovalor *perturbado* E_n . O procedimento para obter uma aproximação para esse autovalor utilizando esta expansão consiste em truncar a série (4.8) na ordem desejada e resolver a equação não linear resultante para E_n . O valor assim encontrado é então utilizado na série (4.7) para obter a aproximação correspondente para o autovetor perturbado. É importante notar que, sendo soluções de uma equação algébrica cujos termos contém diferentes potências do parâmetro de expansão ξ , as aproximações obtidas para o autovalor e para o autovetor perturbado não são homogêneas em ξ , isto é, não correspondem a “ordens” definidas de uma expansão em potências de ξ .

Uma alternativa à expansão de Brillouin-Wigner, de uso mais freqüente e cujos termos são homogêneos no parâmetro de expansão ξ , é a chamada expansão de Rayleigh-Schrödinger. Esta expansão pode ser obtida da expansão de Brillouin-Wigner escrevendo o autovalor perturbado como uma série de potências em ξ ,

$$E_{n_0} = E_{n_0}^{(0)} + \sum_{p=1}^{\infty} \xi^p e_{n_0}^{(p)},$$

substituindo essa expansão nas séries (4.7) e (4.8). e finalmente homogeneizando os termos no parâmetro de expansão ξ . Para isso os denominadores envolvendo o autovalor perturbado são reescritos usando

$$\begin{aligned}
\frac{1}{E_{n_0} - E_m^{(0)}} &= \frac{1}{E_{n_0}^{(0)} + \sum_{p=1}^{\infty} \xi^p e_{n_0}^{(p)} - E_m^{(0)}} = \\
&= \frac{1}{E_{n_0}^{(0)} - E_m^{(0)}} - \xi \frac{e_{n_0}^{(1)}}{(E_{n_0}^{(0)} - E_m^{(0)})^2} - \xi^2 \left[\frac{e_{n_0}^{(2)}}{(E_{n_0}^{(0)} - E_m^{(0)})^2} - \frac{e_{n_0}^{(1)2}}{(E_{n_0}^{(0)} - E_m^{(0)})^3} \right] + \\
&+ \dots
\end{aligned}$$

com o que a equação (4.8) se transforma em

$$\begin{aligned}
E_{n_0}^{(0)} &+ \sum_{p=1}^{\infty} \xi^p e_{n_0}^{(p)} = E_{n_0}^{(0)} + \xi \langle \phi_{n_0}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle + \\
&+ \xi^2 \sum_{m \neq n_0} \frac{|\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle|^2}{E_{n_0}^{(0)} - E_m^{(0)}} + \\
&+ \xi^3 \left[\sum_{m, m' \neq n_0} \frac{\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{m'}^{(0)} \rangle \langle \phi_{m'}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{(E_{n_0}^{(0)} - E_m^{(0)})(E_{n_0}^{(0)} - E_{m'}^{(0)})} - e_{n_0}^{(1)} \sum_{n \neq n_0} \frac{|\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle|^2}{(E_{n_0}^{(0)} - E_m^{(0)})^2} \right] + \\
&+ \mathcal{O}(\xi^4)
\end{aligned}$$

que identifica as correções para o autovalor, ordem por ordem, como

$$\begin{aligned}
e_{n_0}^{(1)} &= \langle \phi_{n_0}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle \\
e_{n_0}^{(2)} &= \sum_{m \neq n_0} \frac{|\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle|^2}{E_{n_0}^{(0)} - E_m^{(0)}} \\
e_{n_0}^{(3)} &= \left[\sum_{m, m' \neq n_0} \frac{\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{m'}^{(0)} \rangle \langle \phi_{m'}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{(E_{n_0}^{(0)} - E_m^{(0)})(E_{n_0}^{(0)} - E_{m'}^{(0)})} - e_{n_0}^{(1)} \sum_{n \neq n_0} \frac{|\langle \phi_{n_0}^{(0)} | H_1 | \phi_m^{(0)} \rangle|^2}{(E_{n_0}^{(0)} - E_m^{(0)})^2} \right] \\
&\dots
\end{aligned}$$

As correções de ordens sucessivas para o vetor de estado podem ser obtidas de modo semelhante, usando a equação (4.7). De fato, fazendo uso da expansão em potências de ξ do autovalor perturbado, ela pode ser reescrita sob a forma

$$\begin{aligned}
|\chi_{n_0}\rangle &= \xi \sum_{m \neq n_0} \frac{|\phi_m^{(0)}\rangle \langle \phi_m^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{E_{n_0}^{(0)} - E_m^{(0)}} + \\
&+ \xi^2 \left[\sum_{m, m' \neq n_0} \frac{|\phi_m^{(0)}\rangle \langle \phi_m^{(0)} | H_1 | \phi_{m'}^{(0)} \rangle \langle \phi_{m'}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{(E_{n_0}^{(0)} - E_m^{(0)})(E_{n_0}^{(0)} - E_{m'}^{(0)})} - e_{n_0}^{(1)} \sum_{m \neq n_0} \frac{|\phi_m^{(0)}\rangle \langle \phi_m^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle}{(E_{n_0}^{(0)} - E_m^{(0)})^2} \right] + \\
&+ \mathcal{O}(\xi^3)
\end{aligned}$$

que identifica completamente essas correções. Como pode ser notado, a correção de uma determinada ordem para o autovetor tem uma estrutura semelhante à da correção *de ordem*

imediatamente maior para o autovalor. A partir da segunda ordem para o vetor de estado (e da terceira para o autovalor), além dos termos semelhantes aos da expansão de Brillouin-Wigner, aparecem termos adicionais, provenientes da expansão dos denominadores de energia, que contém como fatores correções de ordens mais baixas para o autovalor. Esses termos adicionais diferem dos primeiros (nos quais o estado $|\phi_{n_0}^{(0)}\rangle$, cuja perturbação está sendo estudada, aparece uma única vez em cada termo da série correspondente ao autovetor, e apenas no primeiro e último elementos de matriz em cada termo da série correspondente ao autovalor) por conterem sempre um ou mais retornos a esse estado após uma sequência de interações H_1 . É preciso lembrar ainda que o vetor de estado deve ser normalizado pelo fator

$$N_{n_0} = \frac{1}{\sqrt{1 + \langle \chi_{n_0} | \chi_{n_0} \rangle}}$$

que pode ser igualmente expandido em potências de ξ . É fácil ver que a correção de normalização de ordem mais baixa é de ordem ξ^2 .

Os termos da expansão de Rayleigh-Schrödinger podem ser calculados ordem por ordem a partir dos autovalores não perturbados e dos elementos de matriz da perturbação H_1 calculados na base dos autovetores não perturbados. A expansão será tanto mais útil quanto mais rapidamente as contribuições de ordens mais altas se tornem desprezíveis. Para que isso aconteça é necessário que os elementos de matriz de H_1 sejam sempre menores que os espaçamentos entre autovalores que os compensam dimensionalmente nos denominadores de energia, como mencionado no início desta seção, no contexto da identificação de um “parâmetro pequeno” para a expansão perturbativa. Essa condição não é, no entanto, suficiente para assegurar a convergência da série perturbativa. Uma indicação de possíveis problemas nesse sentido pode de fato ser vista já na expansão de Brillouin-Wigner, na qual os denominadores de energia envolvem o autovalor *perturbado*. Sempre que este se aproximar de autovalores não perturbados haverá oportunidade para a manifestação de eventuais problemas de convergência.

As correções perturbativas usadas com maior frequência são as de Rayleigh-Schrödinger de primeira ordem para o vetor de estado e de primeira e segunda ordens para o autovalor. Correções de normalização não intervêm nessas ordens. Uma forma conveniente para lembrar o sinal dos denominadores de energia que aparecem nessas expressões de uso mais freqüente é notar que a correção de segunda ordem *para a energia do estado fundamental* é sempre *negativa*. Como, nesse caso, $E_{n_0}^{(0)}$ é o menor autovalor e as quantidades que aparecem no numerador são não negativas, essa observação fixa univocamente os denominadores correspondentes às correções de primeira ordem para o autovetor, e de segunda ordem para o autovalor, como $E_{n_0}^{(0)} - E_m^{(0)}$, que são sempre negativos para o estado fundamental. O uso da expansão perturbativa de Rayleigh-Schrödinger é ilustrado nos ítems abaixo em algumas situações simples, mas típicas.

1. Oscilador carregado em um campo elétrico externo. Um problema simples que pode ser usado para ilustrar o uso da expansão acima consiste em avaliar a polarizabilidade do estado fundamental de um oscilador harmônico carregado quando submetido a um campo elétrico externo, uniforme. Mais especificamente, uma partícula de massa M e carga elétrica e , presa a um potencial harmônico no qual ela oscila com frequência ω_0 , e submetida ainda a um campo elétrico uniforme \vec{E}_0 , é descrita pelo Hamiltoniano

$$H = \frac{\vec{p}^2}{2M} + \frac{M\omega_0^2}{2}\vec{q}^2 - e\vec{E}_0 \cdot \vec{q}.$$

O problema de autovalores para esse Hamiltoniano pode ser facilmente separado em tres problemas independentes dos quais apenas um é “interessante”, no sentido de conter efeitos do campo aplicado externamente. De fato, escolhendo um sistema de coordenadas cartesianas tal que as componentes do campo elétrico sejam dadas por $\vec{E}_0 \equiv \{0, 0, E_0\}$, o Hamiltoniano pode ser escrito como

$$H = \sum_{j=1}^3 \left(\frac{p_j^2}{2M} + \frac{M\omega_0^2}{2}q_j^2 - eE_0q_3\delta_{j3} \right) \equiv H_1 + H_2 + H_3$$

que corresponde a tres osciladores independentes (pois $[q_i, p_j] = 0$ para $i \neq j$), apenas um dos quais (com Hamiltoniano H_3) depende de \vec{E}_0 . Os autovetores de H podem então ser vistos como produtos de autovetores de cada um desses tres osciladores, sendo que os autovetores $|\nu_1\rangle$ e $|\nu_2\rangle$ de H_1 e H_2 respectivamente são independentes do campo externo. O “estado não perturbado” sobre o qual se quer avaliar o efeito desse campo é um estado não degenerado que pode ser visto simplesmente como o produto dos estados fundamentais dos tres osciladores com $\vec{E}_0 = 0$ e, para uma avaliação perturbativa, é suficiente portanto estudar os efeitos da “perturbação” \vec{E}_0 sobre o estado fundamental de H_3 .

Considerando então a expansão de Rayleigh-Schrödinger para $H_3 \equiv H_0 + \xi H_1$ com

$$H_0 = \frac{p_3^2}{2M} + \frac{M\omega_0^2}{2}q_3^2 \quad \text{e} \quad \xi H_1 = -\xi eE_0q_3$$

resulta que as correções de primeira e segunda ordens para a energia do estado fundamental são dadas por

$$\begin{aligned} e_0^{(1)} &= -\xi eE_0 \langle 0|q_3|0\rangle \\ e_0^{(2)} &= \xi^2 e^2 E_0^2 \sum_{\nu_3} \frac{|\langle \nu_3|q_3|0\rangle|^2}{E_0^{(0)} - E_{\nu_3}^{(0)}} \end{aligned}$$

onde os bras e kets correspondem aos autovetores de H_0 . Essas expressões podem ser facilmente calculadas. De fato, escrevendo q_3 como

$$q_3 = \sqrt{\frac{\hbar}{2M\omega_0}}(a^\dagger + a)$$

é evidente que $e_0^{(1)} = 0$ e que apenas o autovetor não perturbado $|\nu_3 = 1\rangle$ dá uma contribuição não nula para $e_0^{(2)}$. Esta última é, desse modo, dada por

$$e_0^{(2)} = \xi^2 e^2 E_0^2 \frac{|\langle \nu_3 = 1 | q_3 | 0 \rangle|^2}{\frac{\hbar\omega_0}{2} - \frac{3\hbar\omega_0}{2}} = \xi^2 e^2 E_0^2 \frac{\hbar}{2M\omega_0} \frac{-1}{\hbar\omega_0} = -\xi^2 \frac{e^2 E_0^2}{2M\omega_0^2}. \quad (4.9)$$

A correção de primeira ordem para o vetor de estado é semelhante à correção de segunda ordem para o autovalor e também pode ser facilmente obtida como

$$\langle \nu_3 | \chi_0 \rangle = -\xi e E_0 \frac{\langle \nu_3 = 1 | q_3 | 0 \rangle}{-\hbar\omega_0} \delta_{\nu_3 1} + \mathcal{O}(\xi^2) = \xi \sqrt{\frac{\hbar}{2M\omega_0}} \frac{e E_0}{\hbar\omega_0} \delta_{\nu_3 1} + \mathcal{O}(\xi^2)$$

de modo que o autovetor perturbado $|\tilde{0}_3\rangle$ correspondente ao estado fundamental de H_3 é dado até a primeira ordem em ξ por

$$|\tilde{0}_3\rangle \simeq |0_3\rangle + \xi |1_3\rangle \sqrt{\frac{\hbar}{2M\omega_0}} \frac{e E_0}{\hbar\omega_0}. \quad (4.10)$$

Como o estado não perturbado $|0_3\rangle$ é normalizado, a adição da correção de primeira ordem introduz na realidade um erro na normalização do estado perturbado. Devido à ortogonalidade da correção ao estado não perturbado, no entanto, essa correção é de segunda ordem em ξ e por isso é normalmente desprezada.

A polarizabilidade do oscilador carregado pode ser obtida, dentro desse esquema perturbativo, calculando o valor médio do momento de dipolo elétrico do oscilador como função do campo elétrico aplicado. O operador que corresponde ao momento de dipolo elétrico é

$$\vec{d} = e\vec{q}$$

cujo valor médio no estado fundamental perturbado $|\tilde{0}\rangle$ do oscilador tridimensional é

$$\langle \vec{d} \rangle_{\tilde{0}} \equiv \langle \tilde{0} | \vec{d} | \tilde{0} \rangle = e \langle \tilde{0}_3 | q_3 | \tilde{0}_3 \rangle \hat{x}_3$$

onde \hat{x}_3 é um versor na direção do eixo cartesiano escolhido como tendo a direção do campo elétrico. De fato, o estado fundamental perturbado é o produto dos estados fundamentais

perturbados dos tres osciladores unidimensionais independentes H_j , e os osciladores H_1 e H_2 são imunes aos efeitos do campo elétrico. Usando mais uma vez o fato de que q_3 só tem elementos de matriz diferentes de zero entre estados não perturbados com valores de ν_3 que diferem por ± 1 , resulta (com $\xi = 1$)

$$\langle \vec{d} \rangle_{\tilde{0}} \simeq \frac{e^2 E_0}{M\omega_0^2} \hat{x}_3 \equiv \chi E_0 \hat{x}_3$$

onde

$$\chi = \frac{e^2}{M\omega_0^2}$$

é a *polarizabilidade estática* do oscilador. Essa quantidade determina a magnitude do momento de dipolo induzido, que nesse caso varia *linearmente* com a magnitude do campo elétrico aplicado, no caso de campos externos “fracos” (no sentido específico de que efeitos de ordem superior sejam desprezíveis) e independentes do tempo.

Esse resultado pode ser entendido de forma simples em termos completamente *clássicos*. De fato, a distância x_3 da carga à posição de equilíbrio do oscilador quando o campo aplicado tem magnitude E_0 é dada classicamente pela condição de equilíbrio da força devida ao campo com a força de restituição do oscilador

$$eE_0 = M\omega_0^2 x_{3\text{clas.}}$$

que corresponde a um momento de dipolo elétrico

$$|\vec{d}| = e|x_{3\text{clas.}}| = \frac{e^2 E_0}{M\omega_0^2}$$

em acordo com o resultado perturbativo quântico.

O resultado clássico que corresponde ao resultado quântico perturbativo de primeira ordem, porém, *não é perturbativo*, e vale para qualquer intensidade do campo externo. Por isso é interessante averiguar se uma correção perturbativa quântica de ordem maior introduz efeitos novos na polarizabilidade do oscilador. A correção de segunda ordem para o vetor de estado pode ser calculada com relativa facilidade com o resultado

$$|\chi_0\rangle_{2a.\text{ ordem}} = \xi^2 |2_3\rangle \frac{1}{\sqrt{2}} \frac{\hbar}{2M\omega_0} \frac{e^2 E_0^2}{\hbar^2 \omega_0^2}$$

de modo que o autovetor perturbado completo é dado até essa ordem por

$$|\tilde{0}_3\rangle \simeq N_0 \left(|0_3\rangle + \xi |1_3\rangle \sqrt{\frac{\hbar}{2M\omega_0}} \frac{eE_0}{\hbar\omega_0} + \xi^2 |2_3\rangle \frac{1}{\sqrt{2}} \frac{\hbar}{2M\omega_0} \frac{e^2 E_0^2}{\hbar^2 \omega_0^2} \right).$$

Nesse caso é preciso corrigir a normalização até a ordem apropriada calculando o coeficiente N_0 também até segunda ordem em ξ . O resultado desse cálculo é

$$N_0 = \left(1 + \xi^2 \frac{\hbar}{2M\omega_0} \frac{e^2 E_0^2}{\hbar^2 \omega_0^2} + \mathcal{O}(\xi^4)\right)^{-\frac{1}{2}} = \left(1 - \xi^2 \frac{e^2 E_0^2}{4M\hbar\omega_0^3} + \mathcal{O}(\xi^4)\right).$$

Com esses resultados o valor médio do operador dipolo elétrico é dado por

$$\langle d_3 \rangle_{\bar{0}} \simeq N_0^2 \xi \frac{e^2 E_0}{M\omega_0^2} \left(1 + \xi^2 \frac{1}{2} \frac{e^2 E_0^2}{M\hbar\omega_0^3}\right) = \xi \frac{e^2 E_0}{M\omega_0^2} \left(1 - \xi^2 \frac{e^2 E_0^2}{2M\hbar\omega_0^3}\right) \left(1 + \xi^2 \frac{1}{2} \frac{e^2 E_0^2}{M\hbar\omega_0^3}\right).$$

Os termos em ξ^2 que aparecem dentro dos parênteses dão contribuições de terceira e quinta ordens para o momento de dipolo elétrico. A contribuição vinda do autovetor perturbado é no entanto ainda cancelada por uma contribuição igual mas de sinal contrário proveniente do fator de normalização N_0^2 , de modo que o resultado obtido em primeira ordem é mantido a menos de contribuições não significativas de quinta ordem. O resultado obtido em primeira ordem para a polarizabilidade do oscilador não é portanto modificado quando se introduz a correção de segunda ordem no vetor de estado.

O problema do oscilador quântico carregado num campo elétrico externo uniforme é na realidade exatamente solúvel, de modo que o valor médio do momento de dipolo elétrico no seu estado fundamental pode ser calculado também *exatamente*. Para obter a solução exata, basta notar que o termo de energia potencial de H_3 pode ser escrito como

$$\frac{M\omega_0^2}{2} q_3^2 - eE_0 q_3 = \frac{M\omega_0^2}{2} \left(q_3 - \frac{eE_0}{M\omega_0^2}\right)^2 - \frac{e^2 E_0^2}{2M\omega_0^2}$$

que corresponde ainda a um potencial harmônico com o ponto de mínimo deslocado da origem e com o valor dado por $-e^2 E_0^2 / 2M\omega_0^2$, que coincide com a correção perturbativa de segunda ordem para a energia do estado fundamental. Não é difícil verificar explicitamente que a correção perturbativa de terceira ordem para essa energia é nula, e perceber ainda que o mesmo deve ocorrer, ordem por ordem, também em todas as ordens superiores de teoria de perturbações. De fato, cada ordem é caracterizada por uma potência diferente de E_0 , de modo que não pode haver cancelamentos envolvendo ordens perturbativas diferentes. Para obter o estado fundamental exato na presença do campo externo, basta introduzir o operador deslocado $Q_3 \equiv q_3 - eE_0/M\omega_0^2$, cujas relações de comutação com os demais operadores canônicos são idênticas às de q_3 . Dessa forma, escrevendo Q_3 como

$$Q_3 = \sqrt{\frac{\hbar}{2M\omega_0}} (A^\dagger + A)$$

resulta que o estado fundamental na presença do campo externo $|0_3^{(E_0)}\rangle$ é caracterizado por

$$A|\tilde{0}_3\rangle = 0 \quad (4.11)$$

donde

$$\langle\tilde{0}_3|Q_3|\tilde{0}_3\rangle = \langle\tilde{0}_3|q_3 - \frac{eE_0}{M\omega_0^2}|\tilde{0}_3\rangle = 0$$

e portanto o valor médio do operador de dipolo elétrico é dado por

$$\langle\tilde{0}_3|eq_3|\tilde{0}_3\rangle = \frac{e^2E_0}{M\omega_0^2},$$

que coincide com o resultado obtido em primeira ordem de teoria de perturbações e também com o resultado clássico. A natureza do estado fundamental exato na presença do campo externo é revelada observando que

$$A = \frac{1}{\sqrt{2}} \left(\frac{Q_3}{b} + i \frac{bp_3}{\hbar} \right) = \frac{1}{\sqrt{2}} \left(\frac{q_3}{b} - \frac{1}{b} \frac{eE_0}{M\omega_0^2} + i \frac{bp_3}{\hbar} \right) = a - \frac{1}{b\sqrt{2}} \frac{eE_0}{M\omega_0^2}$$

onde $b = \sqrt{\hbar/M\omega_0}$, e portanto a relação (4.11) implica em

$$a|\tilde{0}_3\rangle = \sqrt{\frac{\hbar}{2M\omega_0}} \frac{eE_0}{\hbar\omega_0} |\tilde{0}_3\rangle$$

que mostra que $|\tilde{0}_3\rangle$ é um estado coerente do oscilador não perturbado.

Como uma observação final, vale a pena notar que, apesar da concordância do resultado quântico para o valor médio do operador de dipolo elétrico com o resultado clássico, a solução quântica para o estado fundamental do oscilador submetido ao campo externo é muito diferente da solução clássica. De fato, enquanto classicamente a cada valor da intensidade do campo aplicado corresponde um valor bem definido do momento de dipolo elétrico, a solução quântica dá para os valores possíveis dessa mesma quantidade uma *distribuição de probabilidades*, cuja *média* coincide com o valor clássico. Essa distribuição pode ser escrita explicitamente em termos da função de onda $\phi_0^{(E_0)}(x_3) = \langle x_3|\tilde{0}_3\rangle$ como

$$\rho^{(E_0)}(d_3) = \frac{1}{e} \left| \phi_0^{(E_0)}(d_3/e) \right|^2.$$

Essa densidade foi definida de modo que $\rho^{(E_0)}(d_3)d(d_3)$ seja a probabilidade de que o valor do momento de dipolo elétrico esteja no intervalo pequeno $d(d_3)$ em torno do valor d_3 . O valor médio que coincide com o resultado clássico é dado em termos dessa distribuição como

$$\langle d_3 \rangle_{\tilde{0}} = \int_{-\infty}^{+\infty} \frac{d(d_3)}{e} d_3 \rho^{(E_0)}(d_3).$$

Ao contrário da solução clássica, a solução quântica (exata) associa ao valor médio $\langle d_3 \rangle_{\tilde{0}}$ uma *dispersão* dada por $\sigma_{d_3} = eb$, onde b é o parâmetro do oscilador. Essa dispersão é determinada pelo estado coerente que representa o estado fundamental na presença do campo externo, e provém portanto da deslocalização mínima da carga associada à relação de incerteza. Essa deslocalização mínima é independente da intensidade do campo aplicado e se torna relativamente pouco importante para valores do momento médio de dipolo muito maiores que eb .

2. Oscilador harmônico com perturbação cúbica. Muitas vezes um Hamiltoniano de oscilador harmônico é usado como aproximação para a dinâmica de pequena amplitude em torno de um mínimo da energia potencial $V(q)$, levando em conta apenas o primeiro termo não nulo de uma expansão desse potencial em série de potências em torno do ponto de mínimo

$$V(q) = V(q_0) + \frac{1}{2!}(q - q_0)^2 \frac{d^2 V}{dq^2}(q_0) + \frac{1}{3!} \frac{d^3 V}{dq^3}(q_0)(q - q_0)^3 + \dots$$

onde o termo linear está ausente e $d^2 V/dq^2(q_0) > 0$ devido ao fato de que $V(q)$ tem supostamente um mínimo em q_0 . As modificações introduzidas no espectro pelos termos anarmônicos da expansão em série de potências do potencial são geralmente avaliadas perturbativamente, até ordens perturbativas e graus de anharmonicidade apropriados. Com uma translação da coordenada q que leve o ponto de mínimo à origem, a inclusão da anharmonicidade cúbica conduz a um Hamiltoniano do tipo

$$H = H_0 + \xi H_1; \quad H_0 = \frac{p^2}{2M} + \frac{M\omega_0^2}{2}q^2; \quad H_1 = \gamma q^3 \quad (4.12)$$

onde a relação com a expansão do potencial $V(q)$ é evidentemente

$$M\omega_0^2 = \frac{d^2 V}{dq^2}(q_0), \quad \gamma = \frac{1}{3!} \frac{d^3 V}{dq^3}(q_0).$$

Ignorando por um momento o contexto em que ele foi obtido, é fácil demonstrar variacionalmente que o espectro do Hamiltoniano (4.12) não tem um limite inferior. Para isto basta tomar como estado de prova um estado coerente real do oscilador, $|Q\rangle$ tal que $a|Q\rangle = (Q/\sqrt{2})|Q\rangle$. O valor médio da energia cinética nesse estado não depende de Q e tem o mesmo valor que no estado fundamental do oscilador. O valor médio da energia potencial tem duas contribuições, que provém respectivamente do potencial quadrático do oscilador e do termo de perturbação

cúbico. Para valores suficientemente grandes de Q essas contribuições são dominadas respectivamente por termos proporcionais a Q^2 e a Q^3 , de modo que sua soma pode ser sempre tornada tão negativa quanto se queira, escolhendo para isso um valor suficientemente grande de $|Q|$, com sinal apropriado (sinal *oposto* ao de γ). De fato, para qualquer valor E da energia total do sistema, o potencial que consiste na soma de um termo quadrático e de um termo cúbico tem uma região “classicamente permitida” de extensão infinita, na qual a energia potencial total é *menor* que E . O espectro exato do Hamiltoniano com a perturbação cúbica é por isso de fato *contínuo* e seus autovetores são autovetores impróprios, normalizáveis apenas em termos de funções delta de Dirac. Os métodos adequados para o tratamento desses autovetores são os métodos da teoria do espalhamento.

Isso *não* significa, no entanto, que os resultados perturbativos que podem ser obtidos para esse sistema sejam desprovidos de sentido. Um primeiro ponto a observar é o de que os estados não perturbados envolvidos no cálculo das primeiras várias ordens da expansão perturbativa têm densidade de probabilidade apreciavelmente diferente de zero apenas em uma região espacial que se concentra no “bolso” de potencial criado pelo termo harmônico, onde o termo cúbico assume valores comparativamente pequenos, para valores também suficientemente pequenos de γ . O cálculo dessas correções perturbativas leva então a resultados perfeitamente razoáveis. Esses resultados são ainda adequados ao problema anarmônico completo, envolvendo o potencial $V(q)$, desde que as contribuições anarmônicas de ordem superior sejam desprezíveis com relação ao termo cúbico na região espacial relevante. Esta última observação faz ver ainda que as conclusões obtidas por via variacional acerca do espectro de (4.12) *podem não ser relevantes* para o problema completo envolvendo $V(q)$, pois é preciso ter em conta as *contribuições anarmônicas de ordem superior* na região em que o Hamiltoniano com termo cúbico é dominado por este termo, e estas contribuições podem invalidar inteiramente as conclusões obtidas com base na expansão truncada de $V(q)$. Em outras palavras, com relação ao problema completo envolvendo $V(q)$, a validade de uma descrição em termos do Hamiltoniano (4.12), e conseqüentemente também a validade de resultados exatos relativos a esse Hamiltoniano, são limitadas pela validade do truncamento adotado.

Uma segunda observação, de outra natureza, é a de que mesmo no caso em que o problema completo consista apenas no Hamiltoniano cúbico (4.12), as correções perturbativas das primeiras várias ordens são significativas, para valores suficientemente pequenos de γ , desde que corretamente interpretadas. Nesse caso, de fato, as funções de onda do espectro contínuo terão amplitude apreciável *dentro* do bolso de potencial criado pelo termo harmônico em intervalos estreitos de energia em torno de determinados valores, que em termos da teoria de espalhamento correspondem a *ressonâncias* nesse bolso de potencial. O espectro discreto perturbado que resulta da aplicação de correções perturbativas ao espectro harmônico aproxima, nesse caso, o “espectro de ressonâncias” do problema exato. Em particular, as funções de onda perturbadas aproximam as funções de onda ressonantes do espectro contínuo *na região interna* do bolso de

potencial.

O procedimento de cálculo para as correções perturbativas para o Hamiltoniano (4.12) não difere essencialmente daquele dado para o caso da perturbação linear. A perturbação H_1 pode ser escrita sob a forma

$$H_1 = \gamma \sqrt{\frac{\hbar}{2M\omega_0}} (a^\dagger + a)^3$$

com o que o cálculo dos elementos de matriz necessários para a obtenção das correções perturbativas de Rayleigh-Schrödinger para autovalores e autovetores não apresenta maiores problemas. Como no caso linear, as correções não nulas de ordem mais baixa são as de segunda ordem para os autovalores e as de primeira ordem para os autovetores.

4.1.2 Perturbação de um estado degenerado ou quase-degenerado.

O tratamento desenvolvido na seção anterior deixa evidentemente de ser aplicável quando se trate de avaliar os efeitos de uma perturbação H_1 sobre os estados associados a um eventual autovalor degenerado de H_0 , pois nesse caso os denominadores das expressões obtidas poderão ser nulos. Concomitantemente com este problema, no caso em que o autovalor $E_{n_0}^{(0)}$ tenha degenerescência N , existe uma ambiguidade unitária na escolha do conjunto de N autovetores ortonormais $|\phi_{n_0,k}^{(0)}\rangle$, $k = 1, \dots, N$ associados a esse autovalor (cf. Problema 2.4 f). Enquanto, no caso não degenerado, o autovetor *de ordem zero* é definido sem nenhuma ambigüidade, neste caso existe o problema adicional de definir esse autovetor, o que deve ser feito de forma apropriada à natureza da perturbação H_1 a ser tratada. O problema da degenerescência e o problema da escolha dos autovetores de ordem zero são de fato relacionados, e podem ser tratados conjuntamente, adaptando de forma apropriada o procedimento utilizado para o caso não degenerado.

Em primeiro lugar, a ambiguidade na determinação do vetor de estado em ordem zero é levada em conta reescrevendo a decomposição (4.4) do autovetor perturbado $|\phi_{n_0}\rangle$ numa forma em que o autovetor não perturbado é tomado como uma combinação linear a determinar dos autovetores degenerados. Desse modo essa decomposição é escrita agora sob a forma

$$|\phi_{n_0}\rangle = N_{n_0} \left(\sum_{k=1}^N |\phi_{n_0,k}^{(0)}\rangle c_k + |\chi_{n_0}\rangle \right) \quad (4.13)$$

onde N_{n_0} é uma constante de normalização, os c_k são coeficientes a determinar e o vetor $|\chi_{n_0}\rangle$ é, sem perda de generalidade, ortogonal a *todos* os autovetores degenerados $|\phi_{n_0,k}^{(0)}\rangle$. Uma forma conveniente de exprimir essa condição é exigir que $P_{n_0}|\chi_{n_0}\rangle = 0$, onde P_{n_0} é o operador de projeção sobre o sub-espaço associado aos autovetores degenerados

$$P_{n_0} \equiv \sum_{k=1}^N |\phi_{n_0,k}^{(0)}\rangle \langle \phi_{n_0,k}^{(0)}|.$$

As mesmas etapas seguidas no caso não degenerado podem agora ser repetidas aqui. Substituindo a decomposição (4.13) na equação de autovalores (4.3) se obtém

$$(E_{n_0} - E_{n_0}^{(0)} - \xi H_1) \sum_{k=1}^N c_k |\phi_{n_0,k}^{(0)}\rangle + (E_{n_0} - H_0 - \xi H_1) |\chi_{n_0}\rangle = 0,$$

e tomando o produto escalar dessa relação sucessivamente com um $|\phi_{n_0,l}^{(0)}\rangle$ e com $|\phi_m^{(0)}\rangle$, que são autovetores associados respectivamente ao autovalor degenerado $E_{n_0}^{(0)}$ e ao autovalor $E_m^{(0)} \neq E_{n_0}^{(0)}$ de H_0 , resultam as expressões correspondentes a (4.5) e a (4.6)

$$(E_{n_0} - E_{n_0}^{(0)}) c_l - \xi \sum_{k=1}^{\infty} \langle \phi_{n_0,l}^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k = \xi \langle \phi_{n_0,l}^{(0)} | H_1 | \chi_{n_0} \rangle \quad (4.14)$$

e

$$\langle \phi_m^{(0)} | \chi_{n_0} \rangle = \xi \sum_{k=1}^N \frac{\langle \phi_m^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k}{E_{n_0} - E_m^{(0)}} + \xi \frac{\langle \phi_m^{(0)} | H_1 | \chi_{n_0} \rangle}{E_{n_0} - E_m^{(0)}}. \quad (4.15)$$

Da mesma forma que no tratamento do caso não degenerado, a segunda dessas duas expressões pode ser iterada de forma a gerar uma série para as amplitudes $\langle \phi_m^{(0)} | \chi_{n_0} \rangle$. Neste caso, no entanto, todos os termos da série dependem linearmente dos coeficientes c_k que definem as amplitudes dos estados degenerados no autovetor perturbado. Os três primeiros termos da série são, explicitamente,

$$\begin{aligned} \langle \phi_m^{(0)} | \chi_{n_0,k} \rangle &= \xi \sum_{k=1}^N \frac{\langle \phi_m^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k}{E_{n_0} - E_m^{(0)}} + \\ &+ \xi^2 \sum_{m' \neq n_0} \sum_{k=1}^N \frac{\langle \phi_m^{(0)} | H_1 | \phi_{m'}^{(0)} \rangle \langle \phi_{m'}^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k}{(E_{n_0} - E_m^{(0)})(E_{n_0} - E_{m'}^{(0)})} + \\ &+ \xi^3 \dots \end{aligned} \quad (4.16)$$

Usando essa série para exprimir o último termo de (4.14) resulta

$$\begin{aligned}
(E_{n_0} - E_{n_0}^{(0)})c_l &= \xi \sum_{k=1}^N \langle \phi_{n_0,l}^{(0)} | H_1 | \phi_{n_0}^{(0)} \rangle c_k + \\
&+ \xi^2 \sum_{m \neq n_0} \sum_{k=1}^N \frac{\langle \phi_{n_0,l}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k}{E_{n_0} - E_m^{(0)}} + \\
&+ \xi^3 \sum_{m,m' \neq n_0} \sum_{k=1}^N \frac{\langle \phi_{n_0,l}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{m'}^{(0)} \rangle \langle \phi_{m'}^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k}{(E_{n_0} - E_m^{(0)})(E_{n_0} - E_{m'}^{(0)})}.
\end{aligned} \tag{4.17}$$

As equações (4.16) e (4.17) correspondem, no caso não degenerado, à expansão perturbativa de Brillouin-Wigner. O tratamento dado ao grupo de estados degenerados introduziu no entanto as amplitudes c_k , que devem ser determinadas pela equação (4.17), juntamente com as aproximações para o autovalor perturbado E_{n_0} . Como a dependência de E_{n_0} é não linear, o cálculo de aproximações dentro de um esquema desse tipo não é conveniente, além de não ser particularmente valioso. O procedimento usual consiste por isso em buscar um esquema semelhante à expansão de Rayleigh-Schrödinger, no sentido de poder ser calculado ordem por ordem tanto no que se refere às correções para o autovetor como às correções para o autovalor. Isso pode ser conseguido facilmente para as ordens mais baixas, que são as que são tipicamente usadas neste contexto.

Em primeiro lugar, as correções de primeira ordem para o autovalor degenerado são obtidas truncando nessa ordem a série (4.17). Disso resulta a equação

$$(E_{n_0}^{(1)} - E_{n_0}^{(0)})c_l = \sum_{k=1}^N \langle \phi_{n_0,l}^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle c_k$$

que corresponde simplesmente à diagonalização da perturbação H_1 no sub-espço degenerado. As correções de primeira ordem para $E_{n_0}^{(0)}$ são portanto dadas pelos autovalores da matriz que representa a interação H_1 no sub-espço dos autovetores degenerados. A solução desse problema de diagonalização determina também os autovetores de ordem zero adequados à perturbação H_1 como autovetores $\{c_k^{(j)}\}$, $j = 1, \dots, N$ dessa mesma matriz. Eles podem ser escritos mais explicitamente como

$$|\phi_{n_0,j}^{(0)}\rangle = \sum_{k=1}^N |\phi_{n_0,k}^{(0)}\rangle c_k^{(j)}$$

e as correções de primeira ordem para o autovalor associado a cada um desses autovetores é simplesmente o respectivo valor médio de H_1 , em completa analogia com o resultado obtido no caso não degenerado, isto é

$$E_{n_0,j}^{(1)} - E_{n_0}^{(0)} = \langle \phi_{n_0,j}^{(0)} | H_1 | \phi_{n_0,j}^{(0)} \rangle.$$

Se a equação característica que determina os autovalores de H_1 no sub-espço degenerado tiver raízes múltiplas, a diagonalização de H_1 não removerá completamente a degenerescência do problema não perturbado. Em particular, se a perturbação H_1 tem elementos de matriz nulos entre os autovetores degenerados $|\phi_{n_0,k}^{(0)}\rangle$, não haverá correção de primeira ordem para $E_{n_0}^{(0)}$ e a ambigüidade na escolha do autovetores de ordem zero não é resolvida. Nesses casos é necessário tratar também termos de segunda ordem na série (4.17). Como o autovalor perturbado E_{n_0} difere de $E_{n_0}^{(0)}$ por termos que são pelo menos de primeira ordem, é possível substituí-lo por este último no denominador de energia, donde resulta

$$(E_{n_0}^{(2)} - E_{n_0}^{(0)})c_l = \sum_{k=1}^N \left[\langle \phi_{n_0,l}^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle + \sum_{m \neq n_0} \frac{\langle \phi_{n_0,l}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{n_0,k}^{(0)} \rangle}{E_{n_0}^{(0)} - E_m^{(0)}} \right] c_k \quad (4.18)$$

que corresponde à diagonalização da perturbação H_1 modificada por uma *interação efetiva* de segunda ordem

$$H_{1 \text{ ef.}}^{(2)} \equiv H_1 + \sum_{m \neq n_0} \frac{H_1 |\phi_m^{(0)}\rangle \langle \phi_m^{(0)}| H_1}{E_{n_0}^{(0)} - E_m^{(0)}}.$$

Os autovetores $\{c_k^{(j)}\}$, $j = 1, \dots, N$ desse objeto definem novos autoestados de ordem zero em termos dos quais as correções *até segunda ordem* do autovalor degenerado são dadas por

$$\begin{aligned} E_{n_0,j}^{(2)} - E_{n_0}^{(0)} &= \sum_{k,l=1}^N c_k^{(j)*} \langle \phi_{n_0,k}^{(0)} | H_{1 \text{ ef.}}^{(2)} | \phi_{n_0,l}^{(0)} \rangle c_l^{(j)} = \\ &= \sum_{k,l=1}^N c_k^{(j)*} \langle \phi_{n_0,k}^{(0)} | H_1 | \phi_{n_0,l}^{(0)} \rangle c_l^{(j)} + \sum_{k,l=1}^N \sum_{m \neq n_0} \frac{c_k^{(j)*} \langle \phi_{n_0,k}^{(0)} | H_1 | \phi_m^{(0)} \rangle \langle \phi_m^{(0)} | H_1 | \phi_{n_0,l}^{(0)} \rangle c_l^{(j)}}{E_{n_0}^{(0)} - E_m^{(0)}} \end{aligned}$$

mais uma vez em completa analogia com a expressão correspondente para o caso não degenerado.

À inclusão de efeitos de segunda ordem nos autovalores corresponde a inclusão de correções de primeira ordem para os autovetores degenerados o que, de acordo com as equações (4.17) e (4.13), dá

$$|\phi_{n_0,j}^{(0)}\rangle \simeq \sum_{k=1}^N |\phi_{n_0,k}^{(0)}\rangle c_k^{(j)} + \sum_{m \neq n_0} \sum_{k=1}^N \frac{|\phi_m^{(0)}\rangle \langle \phi_m^{(0)}| H_1 |\phi_{n_0,k}^{(0)}\rangle c_k^{(j)}}{E_{n_0}^{(0)} - E_m^{(0)}} \quad (+\mathcal{O}(\xi^2))$$

onde os coeficientes $c_k^{(j)}$ são os correspondentes às soluções da equação secular de segunda ordem (4.18).

A extensão do cálculo de correções perturbativas para ordens mais altas encontra a questão adicional dos termos provenientes da expansão dos denominadores de energia em torno dos autovalores não perturbados, que passam a estar presentes a partir da segunda e da terceira ordens, respectivamente para os autovetores e para os autovalores. Elas podem ser obtidas sempre que necessário, mas são de uso muito pouco freqüente.

Oscilador bi-dimensional carregado num campo magnético externo. Um exemplo simples da aplicação dos resultados acima consiste no movimento de uma partícula carregada cujo movimento é confinado a um domínio bi-dimensional no qual ela é sujeita a um potencial harmônico e à ação de um campo magnético externo, perpendicular ao plano em que se dá o movimento da partícula. O Hamiltoniano que descreve essa situação é

$$H = H_0 + \xi H_1; \quad H_0 = \frac{\vec{p}^2}{2M} + \frac{M\omega_0^2}{2}\vec{q}^2, \quad H_1 = gl_3$$

onde $\vec{p} = \{p_1, p_2\}$, $\vec{q} = \{q_1, q_2\}$ e $l_3 \equiv (q_1 p_2 - q_2 p_1)$ é a componente perpendicular ao plano do momento angular da partícula. Este último termo representa a interação da partícula com o campo magnético externo, que classicamente é dada por $\vec{\mu} \cdot \vec{B}$, sendo $\vec{\mu}$ o momento magnético associado à partícula, que é proporcional ao seu momento angular. O Hamiltoniano H_0 pode ser escrito como a soma de dois hamiltonianos unidimensionais de oscilador harmônico, que são constantes do movimento devido às relações canônicas de comutação

$$[q_j, q_j] = [p_j, p_j] = 0, \quad [q_j, p_j] = i\hbar \delta_{jk}.$$

Desse modo o problema de autovalores para H_0 pode ser resolvido como

$$H_0 |\nu_1, \nu_1\rangle = \hbar\omega_0(\nu_1 + \nu_2 + 1) |\nu_1, \nu_2\rangle$$

onde os autovetores $|\nu_1, \nu_2\rangle$ são produtos de dois autovetores de oscilador unidimensional h_1 e h_2 correspondentes respectivamente às variáveis dinâmicas com índices 1 e 2. Como a energia depende apenas da soma dos dois inteiros ν_1 e ν_2 , todos os autovalores são degenerados, com exceção do que corresponde ao estado fundamental.

Uma forma conveniente da perturbação H_1 para trabalhar nessa representação dos autovetores não perturbados é obtida exprimindo os operadores p_j e q_j como

$$q_j = \frac{b}{\sqrt{2}}(a_j^\dagger + a_j), \quad p_j = \frac{i\hbar}{b\sqrt{2}}(a_j^\dagger - a_j), \quad b \equiv \sqrt{\frac{\hbar}{M\omega_0}}$$

o que leva a

$$H_1 = i\hbar g(a_2^\dagger a_1 - a_1^\dagger a_2)$$

que mostra claramente que H_1 não é compatível com h_1 e h_2 , e portanto que a perturbação não é diagonal na representação adotada para o sistema não perturbado. Desse modo, considerando para fixar as idéias o autovalor $3\hbar\omega_0$, que corresponde a $\nu_1 + \nu_2 = 2$, as energias corrigidas até primeira ordem e os autovetores de ordem zero adaptados a essa perturbação são determinados resolvendo o problema de autovalores para H_1 restrito ao sub-espço tridimensional gerado pelos vetores degenerados $|\phi_1\rangle = |2, 0\rangle$, $|\phi_2\rangle = |1, 1\rangle$, $|\phi_3\rangle = |0, 2\rangle$. Os elementos de matriz necessários podem ser calculados facilmente, levando a

$$e^{(j)} \begin{pmatrix} c_1^{(j)} \\ c_2^{(j)} \\ c_3^{(j)} \end{pmatrix} = \begin{pmatrix} 0 & -i\sqrt{2}\hbar g & 0 \\ i\sqrt{2}\hbar g & 0 & -i\sqrt{2}\hbar g \\ 0 & i\sqrt{2}\hbar g & 0 \end{pmatrix} \begin{pmatrix} c_1^{(j)} \\ c_2^{(j)} \\ c_3^{(j)} \end{pmatrix}.$$

Nessa equação os autovalores $e^{(j)}$ correspondem às correções de primeira ordem para a energia $E_{3\hbar\omega_0,j}^{(1)} - E_{3\hbar\omega_0}^{(0)}$ e os autovetores correspondentes $\{c_k^{(j)}\}$ determinam os autovetores de ordem zero no subespaço degenerado considerado.

A equação característica para a matriz a diagonalizar é

$$\lambda^3 - 4\hbar^2 g^2 \lambda = 0$$

cujas raízes são 0 e $\pm 2\hbar g$, o que mostra que a degenerescência é completamente removida em primeira ordem, neste caso. Os autovetores associados a cada um desses autovalores são, por outro lado

$$\{c_k^{(0)}\} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \{c_k^{(+)}\} = \frac{1}{2} \begin{pmatrix} 1 \\ i\sqrt{2} \\ -1 \end{pmatrix} \quad \text{e} \quad \{c_k^{(-)}\} = \frac{1}{2} \begin{pmatrix} 1 \\ -i\sqrt{2} \\ -1 \end{pmatrix}$$

que correspondem, na notação de kets, aos vetores de estado

$$|e^{(0)}\rangle = \frac{1}{\sqrt{2}} (|2, 0\rangle + |0, 2\rangle) \quad \text{e} \quad |e^{(\pm)}\rangle = \frac{1}{2} \left(|2, 0\rangle \pm \frac{i}{\sqrt{2}} |1, 1\rangle - |0, 2\rangle \right).$$

O primeiro desses autovetores permanece à energia $3\hbar\omega_0$, enquanto a energia dos dois últimos é $3\hbar\omega_0 \pm 2\hbar g$. É fácil ainda verificar que o operador l_3 , e portanto também o Hamiltoniano de perturbação H_1 , são constantes do movimento (isto é, comutam com H_0). Isso é suficiente para garantir que esse resultado de primeira ordem é na realidade exato. O fato de que H_1 é uma constante do movimento garante que os elementos de matriz de H_1 entre estados não degenerados de H_0 se anulam, o que por sua vez leva ao anulamento de todas as correções perturbativas de ordem superior.

4.1.3 Perturbação da equação de movimento.

A questão a ser examinada nesta seção trata do desenvolvimento de técnicas sistemáticas de aproximação para a solução do problema de condições iniciais para a equação de movimento quântica

$$(H_0 + \xi H_1) |\varphi(t)\rangle = i\hbar \frac{d|\varphi(t)\rangle}{dt}, \quad (4.19)$$

correspondente à descrição de Schrödinger. O problema consiste em obter a solução dessa equação que satisfaz a uma condição inicial dada $|\varphi(t=0)\rangle = |\varphi_0\rangle$, e as técnicas de aproximação serão desenvolvidas com base na suposição de que o problema de autovalores para H_0

$$H_0|\phi_n\rangle = E_n|\phi_n\rangle$$

foi completamente resolvido, isto é, os autovalores e autovetores¹ de H_0 são conhecidos. Isso certamente implica que H_0 deve ser independente do tempo, mas o tratamento que segue é também aplicável para perturbações H_1 que dependem explicitamente do tempo. As técnicas a serem desenvolvidas têm utilidade para um grande número de aplicações, incluindo especialmente o cálculo perturbativo da evolução temporal de estados estacionários do hamiltoniano não perturbado sob a ação dos termos adicionais, tratados perturbativamente, e, em particular, o cálculo de “taxas de decaimento” desses estados para outros estados do espectro não perturbado, a preparação de estados no sistema não perturbado através da ação de perturbações adicionadas externamente, o tratamento perturbativo de problemas de espalhamento, etc.

Como visto na seção 2.3, o conhecimento dos autovalores e autovetores de H_0 é suficiente para resolver completamente o problema de condições iniciais na ausência da perturbação H_1 . Nesse caso, de fato, a solução pode ser expressa por

$$|\varphi(t)\rangle = \sum_n e^{-\frac{i}{\hbar}E_n t} |\phi_n\rangle c_n, \quad (\xi = 0)$$

¹Deve ser notada aqui uma simplificação da notação com relação às seções anteriores, onde os autovalores e autovetores de H_0 foram designados como $E_n^{(0)}$ e $|\phi_n^{(0)}\rangle$ respectivamente.

onde os coeficientes $c_n \equiv \langle \phi_n | \varphi_0 \rangle$ são determinados pela condição inicial. Por outro lado, usando a completeza dos autovetores de H_0 é possível escrever as soluções da equação (4.19) também sob a forma

$$|\varphi(t)\rangle = \sum_n |\phi_n\rangle A_n(t)$$

onde a dependência com o tempo dos coeficientes $A_n(t)$ deve ser determinada usando a equação de movimento, enquanto a condição inicial é levada em conta exigindo que $A_n(t = 0) = c_n$. Uma forma mais conveniente de escrever esses coeficientes é

$$A_n(t) \equiv a_n(t) e^{-\frac{i}{\hbar} E_n t} \quad (4.20)$$

com a condição inicial $a_n(t = 0) = c_n$, pois dessa forma os novos coeficientes $a_n(t)$ se tornam independentes do tempo no caso “não perturbado” $\xi \rightarrow 0$. Adotando, então, a forma

$$|\varphi(t)\rangle = \sum_n |\phi_n\rangle a_n(t) e^{-\frac{i}{\hbar} E_n t} \quad (4.21)$$

para a solução do problema de condição inicial perturbado, é fácil obter a partir de (4.19) equações que determinam a dependência dos $a_n(t)$ com o tempo. De fato, substituindo (4.21) em (4.19) resulta, tendo em conta que os $|\phi_n\rangle$ são autovetores de H_0 ,

$$\sum_n (E_n + \xi H_1) |\phi_n\rangle a_n(t) e^{-\frac{i}{\hbar} E_n t} = \sum_n |\phi_n\rangle \left(i\hbar \frac{da_n(t)}{dt} + E_n a_n(t) \right) e^{-\frac{i}{\hbar} E_n t}.$$

Os termos que contém E_n como fator se cancelam, consistentemente com a observação já feita de que no limite $\xi \rightarrow 0$ deve-se ter $da_n/dt \equiv 0$. Tomando ainda o produto escalar dessa equação com $|\phi_m\rangle$, e tendo em conta a ortonormalidade desses autovetores, resulta o sistema de equações diferenciais acopladas para os $a_n(t)$

$$i\hbar \frac{da_m(t)}{dt} = \sum_n \xi \langle \phi_m | H_1 | \phi_n \rangle a_n(t) e^{i\omega_{mn}t} \quad (4.22)$$

onde foram introduzidas as frequências de Bohr

$$\omega_{mn} = \frac{E_m - E_n}{\hbar}.$$

A passagem dos coeficientes de expansão $A_n(t)$ para os $a_n(t)$ conforme a equação (4.20) está ligada à passagem de uma descrição de Schrödinger da dinâmica para uma descrição de interação, apresentada na seção **2.3**. O vetor de estado $|\varphi_I(t)\rangle$ nesta última descrição

está relacionado com o vetor de estado $|\varphi(t)\rangle$ correspondente à descrição de Schrödinger através da relação

$$|\varphi_I(t)\rangle = e^{\frac{i}{\hbar}H_0t}|\varphi(t)\rangle$$

e satisfaz a equação de movimento

$$H_{1I}(t)|\varphi_I(t)\rangle = i\hbar \frac{d|\varphi_I(t)\rangle}{dt} \quad (4.23)$$

onde $H_{1I}(t) = e^{\frac{i}{\hbar}H_0t}H_1e^{-\frac{i}{\hbar}H_0t}$. De fato, usando a expansão (4.21) resulta que o vetor de estado na descrição de interação é dado simplesmente por

$$|\varphi_I(t)\rangle = \sum_n |\phi_n\rangle a_n(t) \quad (4.24)$$

enquanto que os elementos de matriz do termo de interação, escrito também na descrição de interação, são

$$\langle\phi_m|H_{1I}(t)|\phi_n\rangle = \langle\phi_m|H_1|\phi_n\rangle e^{i\omega_{mn}t}. \quad (4.25)$$

As equações acopladas (4.22) podem portanto ser obtidas também a partir da equação de movimento (4.23), usando a expansão (4.24) e os elementos de matriz (4.25). Os fatores de fase envolvendo as frequências de Bohr resultam, nessa descrição, da transformação unitária envolvendo o operador de evolução não perturbado que a relaciona com a descrição de Schrödinger.

As equações acopladas (4.22) foram obtidas da equação de movimento (4.19) sem qualquer aproximação, e são portanto apenas uma representação alternativa da mesma dinâmica, que é no entanto útil como ponto de partida para um esquema de aproximação baseado na idéia de uma hierarquização dos efeitos da perturbação em termos de potências de ξ . Para implementar um esquema desse tipo, cada um dos coeficientes $a_n(t)$ é escrito sob a forma de uma série de potências de ξ

$$a_n(t) = \sum_{k=0}^{\infty} a_n^{(k)}(t)\xi^k \quad (4.26)$$

cuja substituição em (4.22) dá

$$i\hbar \sum_k \xi^k \frac{da_m^{(k)}(t)}{dt} = \sum_k \sum_n \xi^{k+1} \langle\phi_m|H_1|\phi_n\rangle e^{i\omega_{mn}t} a_n^{(k)}(t).$$

Esse novo sistema pode agora ser analisado separando ordens sucessivas no parâmetro de expansão ξ . Os termos de ordem zero dão simplesmente

$$\frac{da_m^{(0)}(t)}{dt} = 0,$$

o que significa que os termos de ordem zero em ξ nas séries (4.26) são independentes do tempo. Eles correspondem, de fato, ao problema de condição inicial *não perturbado*, de modo que a escolha natural para os valores dos $a_m^{(0)}$ é

$$a_m^{(0)} = c_m$$

com o que a condição inicial para os $a_n^{(k)}(t)$ com $k > 0$ é simplesmente $a_n^{(k>0)}(t = 0) = 0$. Colecionando em seguida os termos de primeira ordem em ξ se obtém a expressão

$$i\hbar \frac{da_m^{(1)}(t)}{dt} = \sum_n \langle \phi_m | H_1 | \phi_n \rangle e^{i\omega_{mn}t} a_n^{(0)}$$

que pode ser integrada, com a condição inicial homogênea, dando

$$a_m^{(1)}(t) = \frac{1}{i\hbar} \int_0^t dt' \sum_n \langle \phi_m | H_1 | \phi_n \rangle e^{i\omega_{mn}t'} c_n. \quad (4.27)$$

Essa expressão determina as contribuições *de primeira ordem em ξ* para os coeficientes $a_m(t)$ em termos da condição inicial. Os efeitos de H_0 são tratados através dos fatores de fase envolvendo as frequências de Bohr ω_{mn} do problema não perturbado. Os termos correspondentes a uma potência genérica do parâmetro de expansão fornecem, por outro lado, uma relação de recorrência que permite obter as contribuições de ordens superiores em ξ que, integrada com a condição inicial homogênea, é dada por

$$a_m^{(k)}(t) = \frac{1}{i\hbar} \int_0^t dt' \sum_n \langle \phi_m | H_1 | \phi_n \rangle e^{i\omega_{mn}t'} a_n^{(k-1)}(t'). \quad (4.28)$$

Utilizando essa relação recursivamente é possível exprimir as contribuições de uma ordem qualquer k também em termos da condição inicial através de k integrações sobre o tempo e de k somas sobre autovetores de H_0 envolvendo produtos de k elementos de matriz de H_1 com os respectivos fatores de fase contendo frequências de Bohr. As contribuições de segunda ordem, por exemplo, são dadas por

$$a_m^{(2)}(t) = \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \sum_{n_1} \sum_{n_2} \langle \phi_m | H_1 | \phi_{n_1} \rangle e^{i\omega_{mn_1}t_1} \langle \phi_{n_1} | H_1 | \phi_{n_2} \rangle e^{i\omega_{n_1n_2}t_2} c_{n_2}. \quad (4.29)$$

A aproximação perturbativa de ordem k calcula os coeficientes $a_n(t)$ ignorando as contribuições de ordem maior que k na expansão (4.26), em potências de ξ . Uma característica importante das contribuições de ordem $k > 0$ é a de que elas satisfazem condições iniciais homogêneas, e são portanto de fato pequenas (no sentido de serem muito menores que 1) *para tempos suficientemente curtos*. Além disso, a contribuição de ordem k é alimentada apenas pela contribuição de ordem $k - 1$, de modo que contribuições de ordem mais alta tendem a ser pequenas por tempos mais longos que as de ordem mais baixa. Uma condição *necessária* para a validade de uma correção perturbativa num tempo T é a de que para qualquer $t < T$ se tenha

$$|a_n^{(k)}(t)|^2 \ll 1, \quad k \geq 1. \quad (4.30)$$

A relação estreita entre as equações acopladas (4.22) e a descrição de interação apontada acima pode ser usada para estabelecer a hierarquização em potências de ξ de outro modo. Escrevendo a solução da equação de movimento (4.23), com a condição inicial prescrita, sob a forma

$$|\varphi_I(t)\rangle = U_I(t, 0)|\varphi_I(t=0)\rangle = U_I(t, 0)|\varphi_0\rangle, \quad (4.31)$$

o operador de evolução da descrição de interação $U_I(t, 0)$ é determinado pela equação diferencial

$$H_{1I}(t)U_I(t, 0) = i\hbar \frac{dU_I(t, 0)}{dt}$$

com a condição inicial $U_I(t=0, 0) = \hat{1}$. Integrando os dois membros dessa equação diferencial e incorporando a condição inicial que deve ser satisfeita por $U_I(t, 0)$, ela se transforma na equação integral

$$U_I(t, 0) = \hat{1} + \frac{1}{i\hbar} \int_0^t dt' H_{1I}(t') U_I(t', 0) \quad (4.32)$$

que pode ser tratada recursivamente levando à chamada *série de Dyson*

$$\begin{aligned} U_I(t, 0) = \hat{1} &+ \frac{1}{i\hbar} \int_0^t dt_1 H_{1I}(t_1) + \\ &+ \frac{1}{(i\hbar)^2} \int_0^t dt_1 H_{1I}(t_1) \int_0^{t_1} dt_2 H_{1I}(t_2) + \\ &+ \frac{1}{(i\hbar)^3} \int_0^t dt_1 H_{1I}(t_1) \int_0^{t_1} dt_2 H_{1I}(t_2) \int_0^{t_2} dt_3 H_{1I}(t_3) + \\ &+ \dots \end{aligned} \quad (4.33)$$

Usando agora a relação (4.31) juntamente com esta expansão do operador de evolução, o vetor de estado $|\varphi_I(t)\rangle$ aparece escrito sob a forma de uma série que, como pode ser facilmente verificado introduzindo explicitamente a representação em termos dos autovetores de H_0 com o uso da resolução da unidade

$$\hat{1} = \sum_n |\phi_n\rangle\langle\phi_n|,$$

coincide com a expansão (4.26). É claro que a passagem da equação (4.32) à equação (4.33) e suas conseqüências envolve questões não triviais de convergência, como acontece também quando as equações acopladas (4.22) são tratadas em termos da expansão (4.26).

1. Oscilador sujeito a uma força externa constante. O problema solúvel do oscilador harmônico forçado tratado na seção **3.1.2** pode ser útil para ilustrar o uso da expansão perturbativa para problemas de condição inicial. Tomando o termo do Hamiltoniano (3.2) correspondente à força externa como perturbação, e supondo inicialmente, para maior simplicidade, que essa força externa seja constante, isto é

$$H_0 = \hbar\omega_0(a^\dagger a + \frac{1}{2}), \quad H_1 = -F_0(a^\dagger + a),$$

as contribuições perturbativas de primeira e segunda ordens para o vetor de estado dependente do tempo são facilmente obtidas das expressões (4.27) e (4.29), tomando como condição inicial o estado fundamental $|\nu = 0\rangle$ de H_0 . Essa condição inicial corresponde a $c_\nu = \delta_{\nu 0}$, e as contribuições perturbativas podem ser calculadas sem dificuldade dando

$$a_\nu^{(1)}(t) = -\frac{F_0}{i\hbar} \int_0^t dt' \delta_{\nu 1} e^{i\omega_0 t'} = \frac{F_0}{\hbar\omega_0} (e^{i\omega_0 t} - 1) \delta_{\nu 1}$$

e

$$\begin{aligned} a_\nu^{(2)}(t) &= \left(\frac{F_0}{i\hbar}\right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 (\delta_{\nu 2} \sqrt{2} e^{i\omega_0 t_1} + \delta_{\nu 0} e^{i\omega_0 t_1}) e^{i\omega_0 t_2} \\ &= \frac{F_0^2}{\hbar^2 \omega_0^2} \left[\frac{\delta_{\nu 2}}{\sqrt{2}} (e^{i\omega_0 t} - 1)^2 + \delta_{\nu 0} (i\omega_0 t + e^{-i\omega_0 t} - 1) \right] \end{aligned}$$

onde os deltas de Kronecker provém dos elementos de matriz de H_1 , que são diferentes de zero apenas para autovetores de H_0 cujas energias difiram por $\pm\hbar\omega_0$. A aproximação perturbativa até segunda ordem para o vetor de estado dependente do tempo é portanto

$$\begin{aligned}
|t\rangle &\simeq (1 + a_0^{(2)})|0\rangle + a_1^{(1)}|1\rangle + a_2^{(2)}|2\rangle = \\
&= \left[1 + \frac{F_0^2}{\hbar^2 \omega_0^2} (i\omega_0 t + e^{-i\omega_0 t} - 1) \right] |0\rangle + \frac{F_0}{\hbar \omega_0} (e^{i\omega_0 t} - 1) |1\rangle + \frac{F_0^2}{\sqrt{2} \hbar^2 \omega_0^2} (e^{i\omega_0 t} - 1)^2 |2\rangle.
\end{aligned}$$

No caso da contribuição de primeira ordem para a componente de $|t\rangle$ que é proporcional ao autovetor não perturbado $|1\rangle$ a restrição (4.30) assume a forma

$$2 \frac{F_0^2}{\hbar^2 \omega_0^2} (1 - \cos \omega_0 t) \ll 1.$$

Para $1/\omega_0 \gg t$ o lado esquerdo dessa desigualdade se comporta como $F_0^2 t^2 / \hbar^2$, o que corresponde à limitação $t \ll \hbar / |F_0|$. Essa limitação prevalece sobre a suposição de que $t \ll 1/\omega_0$ quando $|F_0| > \hbar \omega_0$. Se, por outro lado, $|F_0| \ll \hbar \omega_0$, o uso de (4.30) para o termo de primeira ordem não leva a qualquer condição sobre t , devido ao caráter limitado de $\cos \omega_0 t$.

Uma discussão inteiramente análoga se aplica também à contribuição de segunda ordem $a_2^{(2)}(t)$. O mesmo não acontece, no entanto, com relação a $a_0^{(2)}(t)$, devido à ocorrência de um termo linear em t nessa amplitude. Nesse caso a condição (4.30) dá

$$|a_0^{(2)}(t)|^2 = \frac{F_0^4}{\hbar^4 \omega_0^4} |i\omega_0 t + e^{-i\omega_0 t} - 1|^2 \ll 1.$$

Para $t \ll 1/\omega_0$ essa quantidade se comporta como $F_0^4 t^4 / 4 \hbar^4$, mas para tempos maiores a contribuição que se torna dominante é do tipo $F_0^4 t^2 / \hbar^4 \omega_0^2$, que restringe t mesmo para valores pequenos de F_0 .

Esses resultados podem ser comparados com a solução exata obtida na seção **3.1.2**. Em termos da descrição de Schrödinger, o vetor de estado dependente do tempo $|t\rangle$ foi identificado ali como sendo um estado coerente que, para a condição inicial e para a força externa constante que foram consideradas no tratamento perturbativo, é dado por

$$|t\rangle = |z(t)\rangle = e^{-\frac{|z(t)|^2}{2}} \sum_{\nu} \frac{z(t)^{\nu}}{\sqrt{\nu!}} |\nu\rangle$$

com

$$z(t) = \frac{iF_0}{\hbar} e^{-i\omega_0 t} \int_0^t dt' e^{i\omega_0 t'} = \frac{iF_0}{\hbar} e^{-i\omega_0 t} \frac{e^{i\omega_0 t} - 1}{i\omega_0} \quad (4.34)$$

e portanto

$$\frac{|z(t)|^2}{2} = \frac{F_0^2}{\hbar^2 \omega_0^2} (1 - \cos \omega_0 t).$$

No tratamento dado na seção **3.1.2**, essa identificação se baseou na observação de que o vetor de estado no tempo t é um autovetor do operador a com autovalor $z(t)$, que no entanto determina $|t\rangle$ a menos de um fator de fase que pode depender do tempo. Para fins da comparação a ser feita aqui é importante determinar explicitamente esse fator resolvendo a equação de movimento

$$i\hbar \frac{d|t\rangle}{dt} = (H_0 + H_1)|t\rangle$$

com a condição inicial $|t=0\rangle = |0\rangle$. Isso pode ser feito escrevendo a solução sob a forma

$$|t\rangle = N(t) \sum_{\nu} \frac{z(t)^{\nu}}{\sqrt{\nu!}} |\nu\rangle, \quad \text{com } N(t=0) = 1, \quad z(t=0) = 0. \quad (4.35)$$

A substituição dessa expressão na equação de movimento dá equações para $z(t)$ e $A(t)$ que podem ser resolvidas dando respectivamente o resultado anterior (4.34) e

$$N(t) = \exp \left(-i \frac{\omega_0 t}{2} + i \frac{F_0^2}{\hbar^2 \omega_0^2} [\omega_0 t - \sin \omega_0 t] - \frac{|z(t)|^2}{2} \right).$$

Alguns passos na obtenção da solução exata da equação de movimento por meio do “ansatz” (4.35) podem ser úteis. A substituição na equação de movimento dá, usando o fato de que $a|t\rangle = z(t)|t\rangle$,

$$\begin{aligned} i\hbar \frac{\dot{N}(t)}{N(t)} |t\rangle + i\hbar \dot{z}(t) N(t) \sum_{\nu} \frac{z(t)^{(\nu-1)}}{\sqrt{(\nu-1)!}} \sqrt{\nu} |\nu\rangle = \\ = \frac{\hbar \omega_0}{2} |t\rangle + \hbar \omega_0 z(t) a^{\dagger} |t\rangle - F_0 \left(z(t) a^{\dagger} |t\rangle + z(t) |t\rangle \right). \end{aligned}$$

A soma que aparece no segundo termo com o respectivo fator $N(t)$ pode ser facilmente identificada como $a^{\dagger}|t\rangle$. Usando o fato de que esse vetor e o vetor $|t\rangle$ são linearmente independentes, para que essa equação seja satisfeita é preciso que

$$i\hbar \frac{\dot{N}(t)}{N(t)} = \frac{\hbar \omega_0}{2} - F_0 z(t) \quad \text{e} \quad i\hbar \dot{z}(t) = \hbar \omega_0 z(t) - F_0.$$

Esta última equação pode ser integrada com a condição inicial $z(t=0) = 0$ reproduzindo o resultado (4.34) para $z(t)$, enquanto a equação que envolve $\dot{N}(t)$ com a condição inicial $N(t=0) = 1$ dá

$$N(t) = \exp\left(-i\frac{\omega_0 t}{2} + i\frac{F_0}{\hbar} \int_0^t dt' z(t')\right)$$

que, com o resultado obtido para $z(t)$, leva à expressão dada para $N(t)$.

A forma exata dos coeficientes de expansão do vetor de estado $|t\rangle$ na base dos autovetores $|\nu\rangle$ de H_0 é portanto

$$|t\rangle = \sum_{\nu} A_{\nu}(t) |\nu\rangle = \sum_{\nu} a_{\nu}(t) e^{-i(\nu+\frac{1}{2})\omega_0 t} |\nu\rangle; \quad A_{\nu}(t) = N(t) \frac{z(t)^{\nu}}{\sqrt{\nu!}}, \quad a_{\nu}(t) = e^{i(\nu+\frac{1}{2})\omega_0 t} A_{\nu}(t)$$

de modo que

$$\begin{aligned} a_0(t) &= e^{i\frac{\omega_0 t}{2}} N(t), \\ a_1(t) &= e^{i\frac{\omega_0 t}{2}} N(t) \frac{F_0}{\hbar\omega_0} (e^{i\omega_0 t} - 1) \quad e \\ a_2(t) &= e^{i\frac{\omega_0 t}{2}} N(t) \frac{F_0^2}{\sqrt{2}\hbar^2\omega_0^2} (e^{i\omega_0 t} - 1)^2. \end{aligned}$$

Os coeficientes perturbativos $a_1^{(1)}(t)$ e $a_2^{(2)}(t)$ coincidem com $a_1(t)$ e $a_2(t)$ expandidos respectivamente até primeira e segunda ordem, o que implica apenas em ignorar termos dependentes de F_0^2 provenientes da expansão da gaussiana de normalização $\exp(-|z(t)|^2/2)$. Em particular, a ausência de restrições sobre t no caso em que $F_0^2 \ll \hbar^2\omega^2$ pode ser reconhecida como proveniente de que a amplitude de oscilação da solução exata dada pelo estado coerente $|z(t)\rangle$ é nesse caso suficientemente pequena. A expansão até segunda ordem de $a_0(t)$ dá, por outro lado

$$a_0(t) = 1 + \frac{F_0^2}{\hbar^2\omega_0^2} (i\omega_0 t + \cos\omega_0 t - i\sin\omega_0 t - 1) + \mathcal{O}\left(\frac{F_0^4}{\hbar^4\omega_0^4}\right)$$

que também coincide com o resultado perturbativo até segunda ordem

$$a_0^{(0)} + a_0^{(2)}(t) = 1 + \frac{F_0^2}{\hbar^2\omega_0^2} (i\omega_0 t + e^{-i\omega_0 t} - 1).$$

Vale a pena notar que os termos imaginários, e em particular o termo linear em t , provém da expansão do fator de fase da solução exata, contido em $N(t)$.

2. Oscilador sujeito a uma força externa harmônica. Um caso relevante de perturbação explicitamente dependente do tempo que pode ser tratado no mesmo contexto do caso anterior é o da força externa com dependência temporal harmônica

$$H_1(t) = 2f_0 \cos \omega t (a^\dagger + a),$$

cujas frequência difere da frequência ω_0 do oscilador por uma dessintonia Δ , $\omega = \omega_0 + \Delta$. Tomando novamente o estado fundamental $|0\rangle$ de H_0 como condição inicial, a correção perturbativa de ordem mais baixa popula apenas o primeiro estado excitado através da amplitude

$$a_\nu^{(1)}(t) = \frac{2f_0}{i\hbar} \delta_{\nu 1} \int_0^t dt' e^{i\omega_0 t'} \cos \omega t = -\frac{f_0}{\hbar} \delta_{\nu 1} \left(\frac{e^{i(\omega_0 - \omega)t} - 1}{\omega_0 - \omega} + \frac{e^{i(\omega_0 + \omega)t} - 1}{\omega_0 + \omega} \right).$$

Os termos que contêm $\omega_0 - \omega = -\Delta$ e $\omega_0 + \omega$ nos denominadores são usualmente chamados respectivamente termo *ressonante* e termo *anti-ressonante*, e a probabilidade $|a_1^{(1)}(t)|^2$ contém efeitos de interferência entre esses dois termos. A amplitude do termo ressonante domina no caso em que a dessintonia Δ é pequena na escala de ω_0 , e no limite em que $\Delta \rightarrow 0$ deixa de ter comportamento oscilatório, adquirindo uma dependência *linear* com o tempo. A condição (4.30) aplicada a $|a_1^{(1)}(t)|^2$ pode ser escrita em geral sob a forma

$$|a_1^{(1)}(t)|^2 = \frac{f_0^2}{\hbar^2 \omega_0^2} \left| \frac{\omega_0}{-\Delta} (e^{-i\Delta t} - 1) + \frac{\omega_0}{2\omega_0 + \Delta} (e^{i(2\omega_0 + \Delta)t} - 1) \right|^2 \ll 1.$$

Essa condição não traz nenhuma limitação para t quando a razão $(f_0/\hbar\Delta)^2$ for suficientemente pequena. No caso em que $\Delta = 0$ ela dá $t \ll \hbar/f_0$, supondo que f_0/ω_0 e conseqüentemente as contribuições envolvendo o termo anti-ressonante sejam suficientemente pequenos e não precisem ser levados em conta nesta estimativa.

A solução exata da equação de movimento com a condição inicial adotada tem novamente a forma (4.35) de um estado coerente normalizado e com fatores de fase apropriados. A amplitude $z(t)$ é dada por

$$z(t) = -\frac{f_0}{\hbar} e^{-i\omega_0 t} \left(\frac{e^{i(\omega_0 - \omega)t} - 1}{\omega_0 - \omega} + \frac{e^{i(\omega_0 + \omega)t} - 1}{\omega_0 + \omega} \right)$$

enquanto o fator de normalização com a fase apropriada é

$$N(t) = \exp \left(i \frac{\omega_0 t}{2} - i \frac{2f_0}{\hbar} \int_0^t dt' \cos \omega t' z(t') \right).$$

A dependência desse fator com f_0 é pelo menos de segunda ordem devido à presença de $z(t')$ no integrando. A expansão até primeira ordem em f_0 do resultado exato reproduz o resultado do cálculo perturbativo de primeira ordem.

4.1.4 Transições num espectro contínuo ou quase-contínuo: a “regra áurea” de Fermi.

Uma especialização importante do tratamento perturbativo do problema de condições iniciais consiste no cálculo da probabilidade de transição, induzida por uma perturbação independente do tempo, de um estado estacionário não perturbado inicial para estados estacionários finais, também não perturbados, que fazem parte de um espectro contínuo ou “quase-contínuo”, no sentido que a separação entre diferentes estados finais é pequena na escala \hbar/τ , sendo τ a escala de tempo em que se dá a transição. A expressão para essa probabilidade é a soma das probabilidades de transição para cada um dos níveis do conjunto relevante de estados finais $\{f\}$, que em primeira ordem de teoria de perturbações pode ser escrita, usando a expressão (4.27), como

$$P_{i \rightarrow \{f\}} = \frac{1}{\hbar^2} \sum_{j \in \{f\}} |\langle j | H_1 | i \rangle|^2 \frac{4 \operatorname{sen}^2 \left(\frac{\omega_{ji} t}{2} \right)}{\omega_{ji}^2}. \quad (4.36)$$

Essa expressão mostra que a contribuição de cada estado final j depende de *dois* fatores, que são respectivamente o módulo quadrado do elemento de matriz que o acopla ao estado inicial através da perturbação H_1 e a função dependente do tempo

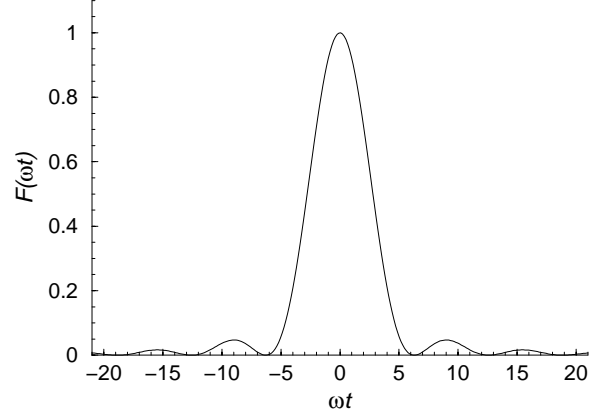
$$f(\omega, t) = \frac{4 \operatorname{sen}^2 \left(\frac{\omega t}{2} \right)}{\omega^2} \quad (4.37)$$

com $\omega \rightarrow \omega_{ji}$, que é a frequência de Bohr associada à diferença de energia entre os estados estacionários inicial e final. Essa função contém, através das frequências de Bohr, *toda* a dependência da energia dos estados estacionários não perturbados que aparece na probabilidade de transição, e uma primeira observação que pode ser feita com relação a ela é a de que para tempos *curtos*, no sentido que $\omega_{ji} t \ll 1$, ela se torna de fato *independente de ω_{ji}* . Em outras palavras, para tais tempos a probabilidade de transição se torna *independente da energia dos estados não perturbados envolvidos na transição*. Isso pode ser interpretado em termos da relação de incerteza energia-tempo, no sentido de que processos sensíveis a diferenças de energia da ordem de $|E_j - E_i|$ são necessariamente associados a escalas de tempo maiores que $\hbar/|E_j - E_i| = 1/|\omega_{ji}|$.²

A validade da expressão perturbativa (4.36) é limitada pela condição de que $P_{i \rightarrow \{f\}} \ll 1$, cujo cumprimento depende portanto também de dois ingredientes distintos. Por um lado essa condição é satisfeita, para tempos suficientemente curtos, independentemente do valor dos elementos de matriz que acoplam o estado inicial aos estados finais. Por outro lado, ela *pode*

²cf. Problema 3.9.

Figura 4.1: Gráfico da função adimensional $F(\omega t) \equiv f(\omega, t)/t^2$ como função de ωt . Quando $f(\omega, t)$ é vista como função de ω com t como parâmetro, o valor de $f(0, t)$ depende de t como t^2 , e os zeros mais próximos de $\omega = 0$ aparecem em $\omega = \pm 2\pi/t$.



estar satisfeita também para tempos muito maiores que $2\pi/|\omega_{ji}|$, desde que os respectivos elementos de matriz de acoplamento sejam suficientemente pequenos. Neste último caso se torna relevante examinar o comportamento de $f(\omega, t)$ como função de ω , para um valor fixo de t . Essa função assume o valor máximo t^2 para $\omega = 0$, se anula para $\omega = \pm 2n\pi/t$, $n = 1, 2, \dots$ e tem máximos secundários entre zeros consecutivos, cujo valor decresce como $1/\omega^2$. Esse comportamento pode ser obtido a partir da figura 4.1, que mostra o gráfico da função adimensional

$$F(\omega t) = \frac{f(\omega, t)}{t^2} = \frac{4 \sin^2\left(\frac{\omega t}{2}\right)}{\omega^2 t^2}.$$

Portanto, se (4.36) permanece válida para tempos suficientemente longos, as contribuições relevantes para a probabilidade de transição total são dominadas por transições que alimentam estados finais tais que $|\omega_{ji}| < 2\pi/t$. Nesse intervalo, de fato, o fator (4.37) tem um pico pronunciado cuja integral com relação a ω pode ser estimada como $2\pi t^2/t = 2\pi t$. Para um espectro de estados finais nessa região com uma densidade de dn/dE níveis por unidade de energia, a relação (4.36) pode ser reescrita como

$$P_{i \rightarrow \{f\}} \rightarrow \frac{1}{\hbar^2} \int d\omega_{ji} \hbar \left(\frac{dn}{dE} \right)_{E_i} |\langle j|H_1|i \rangle|^2 f(\omega_{ji}, t) = \frac{2\pi}{\hbar} \overline{|\langle j|H_1|i \rangle|^2} \left(\frac{dn}{dE} \right)_{E_i} t$$

onde a soma sobre estados foi substituída por uma integral em ω_{ji} e o valor médio do quadrado do elemento de matriz de acoplamento entre estados iniciais e finais foi extraído da integral. Esse resultado é usualmente expresso em termos de uma *probabilidade de transição por unidade de tempo* dada por

$$W \equiv \frac{1}{\tau} = \frac{2\pi}{\hbar} \overline{|\langle j|H_1|i \rangle|^2} \left(\frac{dn}{dE} \right)_{E_i} \quad (4.38)$$

que é uma das formas da *Regra Áurea* de Fermi. Essa expressão é válida para tempos *curtos* no sentido que $Wt \ll 1$. Esses tempos devem no entanto ser *longos* no sentido *diferente* de que as contribuições para W provenham de uma faixa estreita de energia em torno da energia E_i do estado inicial devido às restrições impostas por $f(\omega, t)$. A relação de incerteza energia-tempo permite estimar a largura dessa faixa como sendo da ordem de $\hbar/t \ll \hbar/W$, que deve no entanto ser maior que a separação em energia dos estados finais envolvidos na transição, dE/dn . É nesse sentido que os estados finais devem pertencer a um espectro “quase contínuo”.

Em diversas aplicações da regra áurea o espectro de estados finais não perturbados é não só denso, mas também fortemente degenerado, e os elementos de matriz envolvidos na transição dependem muito fracamente da energia na escala \hbar/W . Um caso simples mas típico em que isso acontece é o do tratamento perturbativo do espalhamento de um feixe de partículas por um potencial, no qual o Hamiltoniano não perturbado é o Hamiltoniano livre $H_0 = \vec{p}^2/2M$. Nesse caso a degenerescência corresponde às diferentes direções do momento para um determinado valor da energia, e uma forma conveniente de escrever a (4.38) é ignorar a dependência dos elementos de matriz com a energia na faixa em que $f(\omega, t)$ é importante e tomar para esta função o limite formal $t \rightarrow \infty$ que dá $f(\omega, t) \rightarrow 2\pi\hbar t \delta(E_j - E_i)$, onde esta última função é uma função delta de Dirac. O que se obtém desse modo é a expressão alternativa

$$W \equiv \frac{1}{\tau} = \frac{2\pi}{\hbar} \sum_{j \in \{f\}} |\langle j|H_1|i \rangle|^2 \delta(E_j - E_i) \quad (4.39)$$

cujo uso é ilustrado nos exemplos abaixo.

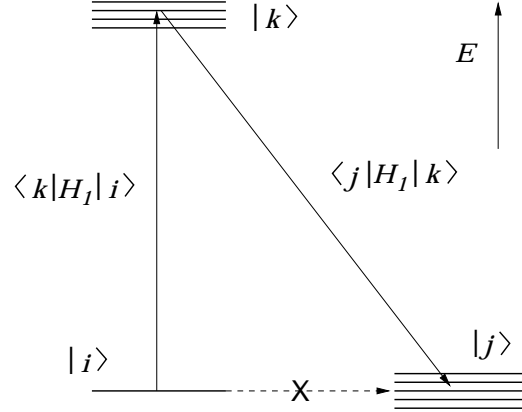
A regra áurea (4.38) e (4.39) foi obtida a partir da expressão perturbativa *de primeira ordem* para $P_{i \rightarrow \{f\}}$, equação (4.36). Existem casos, no entanto, em que os elementos de matriz de transição $\langle j|H_1|i \rangle$ se anulam para todos os estados finais dentro da faixa relevante de energia, embora sendo diferentes de zero para estados com energia diferente da energia inicial E_i . Quando estes últimos, por sua vez, se conectam com estados finais com energia próxima à do estado inicial, a transição pode ocorrer em duas etapas, do estado inicial para os estados distantes em energia e daí para os estados finais de energia próxima à do estado inicial (v. figura 4.2). Nesse caso é necessário usar a segunda ordem de teoria de perturbações. A expressão (4.29) para os coeficientes de segunda ordem dá

$$a_j^{(2)}(t) = -\frac{1}{\hbar^2} \sum_k \frac{\langle j|H_1|k \rangle \langle k|H_1|i \rangle}{\omega_{ki}} \left(\frac{e^{i\omega_{ji}t} - 1}{\omega_{ji}} - \frac{e^{i\omega_{jk}t} - 1}{\omega_{jk}} \right).$$

De forma análoga ao que acontece no cálculo de primeira ordem, os fatores

$$\frac{e^{i\omega t} - 1}{\omega}$$

Figura 4.2: Situação em que transições entre $|i\rangle$ e $\{|j\rangle\}$ são “proibidas” em primeira ordem mas permitidas em segunda ordem, através de processos em duas etapas que passam por estados intermediários $\{|k\rangle\}$. A energia é representada no diagrama como crescendo verticalmente, de baixo para cima.



só dão contribuições importantes (para t suficientemente grande na escala $1/\omega$) para ω suficientemente próximo de zero. Desse modo, o segundo termo entre parênteses seleciona estados k com energia próxima à de j , para os quais o elemento de matriz $\langle k|H_1|i\rangle$, que aparece como fator, supostamente se anula. Nessas condições, portanto, apenas o primeiro termo entre parênteses contribui para a transição de i para j , e a probabilidade de transição no tempo t é dada por

$$|a_j^{(2)}(t)|^2 = \frac{1}{\hbar^2} \left| \sum_k \frac{\langle j|H_1|k\rangle \langle k|H_1|i\rangle}{\hbar\omega_{ki}} \right|^2 f(\omega_{ji}, t)$$

que tem a mesma estrutura que cada um dos termos da (4.36), com a substituição

$$|\langle j|H_1|i\rangle|^2 \rightarrow \left| \sum_k \frac{\langle j|H_1|k\rangle \langle k|H_1|i\rangle}{\hbar\omega_{ki}} \right|^2.$$

Dessa forma, nas mesmas condições em que foram obtidas as expressões (4.38) e (4.39) se obtém agora expressões correspondentes de segunda ordem com essa mesma substituição. A expressão (4.39), por exemplo, fica substituída por

$$W^{(2)} \equiv \frac{1}{\tau} = \frac{2\pi}{\hbar} \sum_{j \in \{f\}} \left| \sum_k \frac{\langle j|H_1|k\rangle \langle k|H_1|i\rangle}{\hbar\omega_{ki}} \right|^2 \delta(E_j - E_i). \quad (4.40)$$

No caso em que os elementos de matriz que contribuem em primeira ordem não se anulam, essa expressão tem uma singularidade em $\omega_{ki} = 0$, ou seja para $E_k = E_i$. Ao mesmo tempo, a contribuição envolvendo ω_{jk} da amplitude de segunda ordem não se anula e é também singular em $\omega_{jk} = 0$. É possível mostrar³ que o tratamento apropriado dessas singularidades leva à expressão

³v. R. P. Feynman e A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, 1965, Seção 6-5, pag. 154.

$$W^{(2)} \equiv \frac{1}{\tau} = \frac{2\pi}{\hbar} \sum_{j \in \{f\}} \left| \langle j|H_1|i\rangle + \lim_{\eta \rightarrow 0} \sum_k \frac{\langle j|H_1|k\rangle \langle k|H_1|i\rangle}{\hbar\omega_{ki} - i\eta} \right|^2 \delta(E_j - E_i) \quad (4.41)$$

que inclui contribuições de primeira e de segunda ordem. Como pode ser notado, as expressões (4.40) e (4.41) têm ainda a mesma estrutura que a versão de primeira ordem da regra áurea de Fermi. A inclusão dos efeitos de segunda ordem apenas modifica a expressão para a amplitude de transição.

1. Tratamento perturbativo do espalhamento por um potencial externo (aproximação de Born). A ação de um potencial externo sobre uma partícula cujo momento inicial é suficientemente grande para que essa ação possa ser bem representada em primeira ordem de teoria de perturbações é uma aplicação ao mesmo tempo simples e importante da regra áurea de Fermi. Essa aproximação para o problema de espalhamento é conhecida como *aproximação de Born*. O Hamiltoniano não perturbado é o Hamiltoniano livre, e o termo perturbativo é o potencial externo que age sobre a partícula de massa m

$$H_0 = \frac{p^2}{2m}; \quad H_1 = V(\vec{q})$$

de modo que os estados não perturbados são ondas planas que será conveniente tratar em termos de condições de contorno periódicas num volume de quantização L^3 . Desse modo o espectro do momento é discreto e as autofunções de onda simultâneas de \vec{p} e de H_0 são

$$\phi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{L^3}} e^{i\vec{k} \cdot \vec{r}}$$

com autovalores $\hbar\vec{k}$ e $\hbar^2 k^2/2m$ respectivamente.

Supondo que o estado inicial da partícula seja o estado de momento $\hbar\vec{k}_0$, os elementos de matriz de transição para estados finais de momento $\hbar\vec{k}$ são

$$\langle \vec{k}|H_1|\vec{k}_0\rangle = \frac{1}{L^3} \int d^3r e^{i(\vec{k}_0 - \vec{k}) \cdot \vec{r}} V(\vec{r}) = \frac{(2\pi)^{3/2}}{L^3} \tilde{V}(\vec{k}_0 - \vec{k})$$

onde $\tilde{V}(\vec{k})$ é a transformada de Fourier do potencial, definida como

$$\tilde{V}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int d^3r e^{i\vec{k} \cdot \vec{r}} V(\vec{r}).$$

A probabilidade de transição por unidade de tempo do estado inicial \vec{k}_0 para estados finais \vec{k} é então dada pela forma (4.39) da regra áurea como

$$W_{\vec{k}_0 \rightarrow \{\vec{k}\}} = \frac{2\pi}{\hbar} \sum_{\vec{k}} \frac{(2\pi)^3}{L^6} |\tilde{V}(\vec{k}_0 - \vec{k})|^2 \delta \left(\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 k_0^2}{2m} \right).$$

A soma sobre os estados finais pode ser transformada numa integral da forma usual, isto é, usando

$$\sum_{\vec{k}} \rightarrow \frac{L^3}{(2\pi)^3} \int d^3 k$$

e a forma mais conveniente de tratar a integral é usando coordenadas esféricas no espaço de momento. O que se obtém desse modo é

$$W_{\vec{k}_0 \rightarrow \{\vec{k}\}} = \frac{2\pi}{\hbar} \frac{1}{L^3} \int d\Omega_{\hat{k}} \int_0^\infty dk k^2 |\tilde{V}(\vec{k}_0 - \vec{k})|^2 \delta \left(\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 k_0^2}{2m} \right)$$

onde $\int d\Omega_{\hat{k}}$ é a integral sôbre os ângulos que definem a direção do momento final. Essa expressão pode ser reescrita como

$$W_{\vec{k}_0 \rightarrow \{\vec{k}\}} = \int d\Omega_{\hat{k}} \frac{dW}{d\Omega_{\hat{k}}}$$

onde

$$\frac{dW}{d\Omega_{\hat{k}}} = \frac{2\pi}{\hbar} \frac{1}{L^3} \int_0^\infty dk k^2 |\tilde{V}(\vec{k}_0 - \vec{k})|^2 \delta \left(\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 k_0^2}{2m} \right)$$

dá a *distribuição angular* da probabilidade de transição total W . Esta última integral pode ser feita facilmente usando a função delta de Dirac. Para isso é importante que a variável de integração seja a variável que aparece como argumento nessa função, o que é conseguido nesse caso fazendo a transformação

$$dk k^2 = d \left(\frac{\hbar^2 k^2}{2m} \right) \frac{2m}{\hbar^2} \frac{k}{2}$$

que leva ao resultado

$$\frac{dW}{d\Omega_{\hat{k}}} = \frac{2\pi m}{\hbar^3 L^3} k_0 |\tilde{V}(\vec{k}_0 - \vec{k})|_{k=k_0\hat{k}}^2$$

onde $k_0 \hat{k}$ é um vetor de módulo k_0 na direção \hat{k} . Essa expressão contém ainda o volume de quantização L^3 , o que provém de que as ondas planas utilizadas foram normalizadas nesse volume, e conseqüentemente a corrente de partículas incidentes no potencial depende do volume de quantização. Essa dependência é no entanto eliminada pela consideração, em lugar de $dW/d\Omega_{\hat{k}}$, da probabilidade de transição *por unidade de fluxo incidente*, o que corresponde à seção de choque diferencial $d\sigma/d\Omega_{\hat{k}}$. De fato, o fluxo incidente associado à onda plana de momento \vec{k}_0 é o módulo da densidade de corrente de probabilidade dado pela equação (3.18), que é $\hbar k_0/mL^3$. Portanto

$$d\sigma/d\Omega_{\hat{k}} = \frac{2\pi m^2}{\hbar^4} |\tilde{V}(\vec{k}_0 - \vec{k})|_{\vec{k}=k_0\hat{k}}^2.$$

A conservação de energia implica que o módulo do momento final \vec{k} seja igual ao do momento inicial \vec{k}_0 , e a dependência da seção de choque diferencial com o ângulo entre esses dois vetores provém da transformada de Fourier do potencial responsável pelo espalhamento.

Considerando, por exemplo, o potencial coulombiano blindado (repulsivo ou atrativo) entre uma carga elétrica $\pm e$ e outra, positiva, Ze

$$V(\vec{r}) = \pm \frac{Ze^2}{r} e^{-\mu r},$$

onde μ é um “parâmetro de blindagem” que poderá ser feito nulo no final do cálculo (cf. problema 4.8), é fácil obter a transformada de Fourier requerida para o cálculo da seção de choque na aproximação de Born

$$\tilde{V}(\vec{k}) = \pm \frac{Ze^2}{(2\pi)^{3/2}} \int d^3r e^{i\vec{k}\cdot\vec{r}} \frac{e^{-\mu r}}{r} = \pm \sqrt{\frac{2}{\pi}} \frac{Ze^2}{k^2 + \mu^2}.$$

Usando esse resultado, a seção de choque diferencial adquire a forma

$$d\sigma/d\Omega_{\hat{k}} = \frac{4m^2 Z^2 e^4}{\hbar^4} \left(\frac{1}{(\vec{k}_0 - k_0 \hat{k})^2 + \mu^2} \right)^2$$

que pode ainda ser modificada de modo a explicitar a dependência com o ângulo de deflexão sofrido pela partícula, que é na realidade o ângulo θ entre o momento inicial \vec{k}_0 e o momento final $k_0 \hat{k}$. Para isso basta calcular

$$(\vec{k}_0 - k_0 \hat{k})^2 = 2k_0^2 - 2k_0^2 \cos \theta = 4k_0^2 \sin^2 \frac{\theta}{2}$$

que leva finalmente a

$$d\sigma/d\Omega_k = \frac{Z^2}{4} \left(\frac{me^2}{\hbar^2 k_0^2} \right)^2 \left(\frac{1}{\sin^2 \frac{\theta}{2} + \frac{\mu^2}{4k_0^2}} \right)^2.$$

No limite $\mu \rightarrow 0$ essa expressão se reduz à seção de choque de Rutherford.

2. Decaimento dos pions carregados. Os pions carregados, π^+ e π^- , são partículas de massa $m_\pi = 139.6$ MeV cuja vida média é $\tau = 2.55 \times 10^{-8}$ segundos. O modo predominante de decaimento (preferido em 99.99% dos casos) leva a estados finais contendo um muon (μ^+ e μ^- respectivamente, com massa $m_\mu = 105.66$ MeV) e um neutrino (com massa zero)

$$\pi^\pm \longrightarrow \mu^\pm + \nu.$$

Supondo que esse decaimento seja produzido por um termo de interação H_1 que, num tratamento perturbativo, acopla estados iniciais $|i\rangle$ contendo apenas um pion livre a estados finais $|f\rangle$ contendo um muon e um neutrino também livres, é possível escrever a taxa de decaimento do pion, fazendo uso mais uma vez da regra áurea escrita sob a forma (4.39), como

$$W = \frac{2\pi}{\hbar} \sum_f |\langle f|H_1|i\rangle|^2 \delta(E_i - E_f).$$

Os elementos de matriz que aparecem nessa expressão são objetos que devem ter a forma geral

$$\langle f|H_1|i\rangle = \left(\frac{1}{\sqrt{L^3}} \right)^3 \int d^3r e^{-i(\vec{k}_\mu + \vec{k}_\nu) \cdot \vec{r}} \langle \mu, \nu | h_1 | \pi \rangle e^{i\vec{k}_\pi \cdot \vec{r}} = \frac{1}{\sqrt{L^3}} \langle \mu, \nu | h_1 | \pi \rangle \delta_{\vec{p}_\mu + \vec{p}_\nu, \vec{p}_\pi}$$

onde as ondas planas descrevem os estados não perturbados inicial e final, que dependem dos momentos $\vec{p} = \hbar \vec{k}$ de cada uma das partículas envolvidas na transição e o elemento de matriz $\langle \mu, \nu | h_1 | \pi \rangle$ representa a parte de H_1 responsável pela conversão do pion, que foi escrito como independente da posição em que esta ocorre, consistentemente com a conservação do momento total, que de fato aparece explicitamente no delta de Kronecker resultante da integração sobre \vec{r} . Substituindo essa expressão em W , e supondo que $\langle \mu, \nu | h_1 | \pi \rangle$ não dependa dos momentos envolvidos na transição, resulta

$$W = \frac{2\pi}{\hbar} \frac{1}{L^3} |\langle \mu, \nu | h_1 | \pi \rangle|^2 \sum_{\vec{p}_\mu, \vec{p}_\nu} \delta_{\vec{p}_\mu + \vec{p}_\nu, \vec{p}_\pi} \delta(\sqrt{m_\pi^2 c^4 + p_\pi^2 c^2} - \sqrt{m_\mu^2 c^4 + p_\mu^2 c^2} - p_\nu c) \quad (4.42)$$

onde foram utilizadas as expressões relativísticas apropriadas para as energias inicial e final. Esta última expressão dá a taxa de transição em termos de um fator dinâmico, que envolve

o elemento de matriz responsável pela conversão do pion, e um segundo fator, correspondente à soma sobre \vec{p}_μ e \vec{p}_ν , que é essencialmente uma contagem dos estados finais compatíveis com a conservação de momento e de energia, chamado *fator de espaço de fase*. Essa fatoração da taxa de transição depende essencialmente da suposição feita de que o elemento de matriz de transição não depende (ou pelo menos não depende apreciavelmente) dos momentos, pois uma tal dependência não permitiria que o elemento de matriz fosse extraído da soma.

Para calcular o fator de espaço de fase, o procedimento padrão consiste primeiramente em usar a conservação de momento para eliminar um dos momentos finais, e em seguida transformar a soma remanescente em uma integral, que pode ser feita facilmente usando a função delta de Dirac correspondente à conservação de energia. Tomando o momento inicial do pion como igual a zero (isto é, supondo que o “pion decai em repouso”), a primeira dessas operações dá $\vec{p}_\nu = -\vec{p}_\mu \equiv \vec{p}$ de modo que a soma se reduz a

$$\sum_{\vec{p}} \delta(m_\pi c^2 - \sqrt{m_\mu^2 c^4 + p^2 c^2} - pc) \rightarrow \frac{L^3}{(2\pi)^3} \int \frac{d^3 p}{\hbar^3} \delta(m_\pi c^2 - \sqrt{m_\mu^2 c^4 + p^2 c^2} - pc).$$

Como o integrando não depende dos ângulos de \vec{p} (significando que, com as hipóteses feitas, o decaimento deve ser isotrópico), a integração angular dá simplesmente um fator 4π , com o que a expressão para o fator de espaço de fases fica reduzido à integral simples

$$\frac{L^3}{(2\pi)^3} \frac{4\pi}{\hbar^3} \int_0^\infty dp p^2 \delta(m_\pi c^2 - \sqrt{m_\mu^2 c^4 + p^2 c^2} - pc) \equiv \frac{L^3}{(2\pi)^3} \frac{4\pi}{\hbar^3} \int_0^\infty dp p^2 \delta(A(p))$$

que pode ser feita através de uma mudança de variável em que a nova variável de integração seja o argumento da função delta de Dirac, isto é

$$\int_0^\infty dp p^2 \delta(A(p)) = \int dA(p) \left| \frac{dp}{dA(p)} \right| p^2 \delta(A(p)) = \left(p^2 \left| \frac{dp}{dA(p)} \right| \right)_{A(p)=0}.$$

Levando finalmente esses resultados à relação (4.42) resulta

$$W = \frac{|\langle \mu, \nu | h_1 | \pi \rangle|^2}{\pi \hbar^4} \left(p^2 \left| \frac{dp}{dA(p)} \right| \right)_{A(p)=0},$$

que é independente do volume de quantização L^3 . O último fator pode ser calculado facilmente, pois

$$\frac{dp}{dA(p)} = \left(\frac{dA(p)}{dp} \right)^{-1} = -\frac{1}{c} \frac{1}{1 + \frac{pc}{E_\mu}}, \quad E_\mu = \sqrt{m_\mu^2 c^4 + p^2 c^2}$$

o que permite escrever a taxa de transição como

$$W = \frac{|\langle \mu, \nu | h_1 | \pi \rangle|^2}{\pi \hbar^4 c} \left(\frac{p^2}{1 + \frac{pc}{E_\mu}} \right)_{E_\mu + pc = m_\pi c^2}. \quad (4.43)$$

Essa expressão pode ser utilizada para extrair um valor para o módulo quadrado do elemento de matriz responsável pela conversão do pion a partir do valor observado de sua vida média, fazendo $W = 1/\tau$. Um modo conveniente de exprimir o resultado é escrever esse objeto como o produto de um fator dimensional, expresso em unidades convenientes, por um fator numérico ajustado para reproduzir o valor observado de τ . De fato, as dimensões de $|\langle \mu, \nu | h_1 | \pi \rangle|^2$ devem ser, para que W seja o inverso de um tempo,

$$[|\langle \mu, \nu | h_1 | \pi \rangle|^2] = M^2 L^7 T^{-4}$$

onde M , L e T indicam unidades de massa, comprimento e tempo respectivamente. Um sistema conveniente de unidades é aquele em que massas são medidas em unidades de m_π , comprimentos em unidades de $\hbar/m_\pi c$ (chamado o comprimento de onda Compton do pion) e tempos em unidades de $\hbar/m_\pi c^2$. Em termos dessas unidades o fator dimensional requerido se escreve portanto como

$$M^2 L^7 T^{-4} = m_\pi^2 \frac{\hbar^7}{m_\pi^7 c^7} \frac{m_\pi^4 c^8}{\hbar^4} = \frac{c \hbar^3}{m_\pi}$$

e portanto

$$|\langle \mu, \nu | h_1 | \pi \rangle|^2 \equiv \frac{c \hbar^3}{m_\pi} g^2$$

que, com os valores observados, dá $g^2 \sim 10^{-15}$.

O modo de decaimento alternativo, escolhido pelo pion em apenas 0.01% dos casos, tem um elétron (ou pósitron) e um neutrino no estado final

$$\pi^\pm \longrightarrow e^\pm + \nu$$

que, supondo que fosse produzido por um elemento de matriz comparável ao que produz o decaimento muonico, deveria na realidade ser fortemente favorecido pelo fator de espaço de fase, pois a massa muito menor do elétron ($m_e = 0.511$ MeV) implica num número muito maior de estados finais compatíveis com a conservação de energia e de momento. O fato experimental de que o decaimento muonico é na realidade o modo favorecido sugere a intervenção de um mecanismo dinâmico que inibe o decaimento para elétron e neutrino. Esse mecanismo está

de fato ligado à não conservação da paridade nas interações responsáveis por esses processos, associado à conservação do momento angular, tendo em conta os momentos angulares intrínsecos das partículas finais e o momento angular intrínseco nulo do pion⁴.

4.2 Formulação variacional do problema de autovalores para espectros discretos.

Uma variável dinâmica g define, através da relação

$$G[|\varphi\rangle] = \frac{\langle\varphi|g|\varphi\rangle}{\langle\varphi|\varphi\rangle},$$

um funcional real G de vetores de estado $|\varphi\rangle$, o que significa que G assume um valor real para cada vetor $|\varphi\rangle$ do espaço de Hilbert. Como mostrado no início deste capítulo, os valores assumidos por esse funcional são limitados inferiormente pelo menor autovalor da variável dinâmica g . No caso em que o espectro de g seja pelo menos parcialmente discreto, isto é, no caso em que g tenha autovetores de norma finita, é possível mostrar ainda que **a)** o funcional G é *estacionário* para variações arbitrárias de $|\varphi\rangle$ quando esse vetor é um *autovetor* de g e reciprocamente; e ainda que **b)** o autovalor de g correspondente a esse autovetor é o valor estacionário do funcional G . Uma *variação arbitrária* de $|\varphi\rangle$ significa simplesmente a substituição

$$|\varphi\rangle \rightarrow |\varphi\rangle + |\delta\rangle$$

onde $|\delta\rangle$ é um vetor arbitrário do espaço de Hilbert com norma muito pequena em termos da norma de $|\varphi\rangle$, isto é

$$\sqrt{\langle\delta|\delta\rangle} \ll \sqrt{\langle\varphi|\varphi\rangle}.$$

Como se verá em seguida, o propósito desta restrição é permitir que, no cálculo da variação do valor do funcional, termos de segunda ordem na norma de $|\delta\rangle$, como por exemplo os que envolvem $\langle\delta|\delta\rangle$, possam ser ignorados frente a termos de primeira ordem, como os que envolvem $\langle\varphi|\delta\rangle$.

A variação δG do funcional correspondente a uma tal variação de $|\varphi\rangle$ pode ser calculada como sendo

⁴v. F. Halzen e A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, 1984, pag. 264.

$$\begin{aligned}
\delta G &= G[|\varphi\rangle + |\delta\rangle] - G[|\varphi\rangle] \\
&= \frac{\langle\delta|g|\varphi\rangle}{\langle\varphi|\varphi\rangle} + \frac{\langle\varphi|g|\delta\rangle}{\langle\varphi|\varphi\rangle} - G[|\varphi\rangle] \left(\frac{\langle\delta|\varphi\rangle}{\langle\varphi|\varphi\rangle} + \frac{\langle\varphi|\delta\rangle}{\langle\varphi|\varphi\rangle} \right) + \mathcal{O}(|\delta|^2) \\
&= \frac{1}{\langle\varphi|\varphi\rangle} (\langle\delta|(g - G[|\varphi\rangle])|\varphi\rangle + \langle\varphi|(g - G[|\delta\rangle])|\varphi\rangle) + \mathcal{O}(|\delta|^2).
\end{aligned}$$

A estacionaridade de G significa que a variação de primeira ordem δG deve se anular para *qualquer* $|\delta\rangle$ de norma infinitesimal, isto é

$$\langle\delta|(g - G[|\varphi\rangle])|\varphi\rangle + \langle\varphi|(g - G[|\delta\rangle])|\varphi\rangle = 0. \quad (4.44)$$

A arbitrariedade de $|\delta\rangle$ permite, em particular, fazer a substituição $|\delta\rangle \rightarrow i|\delta\rangle$ nessa última relação, o que leva a

$$-\langle\delta|(g - G[|\varphi\rangle])|\varphi\rangle + \langle\varphi|(g - G[|\delta\rangle])|\varphi\rangle = 0$$

mostrando que na realidade os dois termos devem se anular independentemente. Da arbitrariedade de $|\delta\rangle$ resulta então

$$(g - G[|\varphi\rangle])|\varphi\rangle = 0 \quad \text{e} \quad \langle\varphi|(g - G[|\delta\rangle]) = 0$$

o que mostra que $|\varphi\rangle$ é autovetor de g com autovalor $G[|\varphi\rangle]$. Reciprocamente, se

$$g|\varphi\rangle = \gamma|\varphi\rangle,$$

segue imediatamente, para qualquer $|\delta\rangle$, que

$$\langle\delta|(g - \gamma)|\varphi\rangle = \langle\varphi|(g - \gamma)|\delta\rangle = 0.$$

É também trivial verificar que, nesse caso, $G[|\varphi\rangle] = \gamma$, de modo que $\delta G = \mathcal{O}(|\delta|^2)$, isto é, G é estacionário em $|\varphi\rangle$.

É importante notar que a substituição $|\delta\rangle \rightarrow i|\delta\rangle$, que permite concluir que os dois termos de (4.44) devem se anular independentemente, decorre do fato de que o espaço de vetores de estado é um espaço vetorial *complexo*. Isso é, de fato, o que faz com que $i|\delta\rangle$ seja também uma variação admissível se $|\delta\rangle$ for admissível. A variação de valores médios de operadores ou de normas de vetores num espaço vetorial complexo pode por isso ser obtida “variando ’bras’ e ’kets’ independentemente”, ou seja, considerando $|\delta\rangle$ e $\langle\delta|$ como variações independentes.

O problema de autovalores para o espectro discreto de uma variável dinâmica g pode portanto ser formulado em termos variacionais através da equação

$$\delta G[|\varphi\rangle] = 0. \quad (4.45)$$

Essa formulação permite a obtenção de *aproximações* para os autovetores e autovalores de g restringindo o domínio de vetores de prova $|\varphi\rangle$ e de variações admissíveis $|\delta\rangle$ de forma computacionalmente conveniente mas, tanto quanto possível, também de forma a preservar as características relevantes dos autovetores. É claro que o significado desta última condição pressupõe um conhecimento pelo menos qualitativo das propriedades “relevantes” da solução exata, no que reside certamente um ponto crítico para a utilidade do método.

Além dos métodos envolvendo estados de prova parametrizados de várias maneiras, citados no início do capítulo, um procedimento bastante usual de aproximação, baseado na formulação variacional do problema de autovalores, é o que consiste em supor que $|\varphi\rangle$ possa ser bem expresso como uma superposição linear de um conjunto finito de vetores ortonormais $\{|v_n\rangle\}$, $n = 1, \dots, N$, isto é

$$|\varphi\rangle = \sum_{n=1}^N c_n |v_n\rangle,$$

e usar (4.45) para determinar os coeficientes c_i . As variações admissíveis são nesse caso também restritas a vetores infinitesimais da forma $|\delta\rangle = \sum_{n=1}^N \delta c_n |v_n\rangle$. Esse procedimento é usado freqüentemente, por exemplo, para obter autofunções e autovalores aproximados para uma partícula sujeita a um poço anarmônico de potencial usando os primeiros N autovetores do oscilador harmônico como vetores $|v_n\rangle$. O funcional restrito G pode ser escrito nesse caso como uma *função* dos coeficientes c_n

$$G[|\varphi\rangle] \rightarrow G(c_n) = \frac{\sum_{nm} c_m^* \langle v_m | g | v_n \rangle c_n}{\sum_k |c_k|^2}.$$

Como os coeficientes c_n são quantidades complexas, é possível “variar c_n e c_n^* independentemente” (o que corresponde a considerar variações δc_n e $i\delta c_n$). Isso leva facilmente à condição de estacionaridade

$$\sum_{m=1}^N \delta c_m^* \sum_{n=1}^N (\langle v_m | g | v_n \rangle - G \delta_{mn}) c_n = 0$$

que, tendo em vista a arbitrariedade dos δc_m^* , leva por sua vez a

$$\sum_{n=1}^N (\langle v_m | g | v_n \rangle - G \delta_{mn}) c_n = 0. \quad (4.46)$$

Os autovetores aproximados de g são portanto dados nesse esquema como os autovetores da matriz dessa variável dinâmica nos vetores $|v_n\rangle$, e as aproximações correspondentes para os autovalores são os autovalores γ dessa mesma matriz

$$\sum_{n=1}^N \langle v_m | g | v_n \rangle = \gamma c_m.$$

O fato de que esses autovalores são também os valores assumidos pela função G calculada com os $\{c_n\}$ correspondentes aos autovetores decorre diretamente de que

$$\sum_{m,n=1}^N c_m^* \langle v_m | g | v_n \rangle c_n = \gamma \sum_{m=1}^N |c_m|^2$$

ou seja, $\gamma = G(c_n)$.

A ortonormalidade dos vetores $\{|v_n\rangle\}$ que definem o sub-espço variacional, embora conveniente, não é necessária. Ela pode ser substituída simplesmente pela independência linear desses vetores (a escolha de um conjunto não linearmente independente de vetores podendo ser sempre substituída por uma escolha, equivalente para os fins da aproximação a ser obtida, de vetores linearmente independentes, obtidos pela eliminação de vetores redundantes). Nesse caso o delta de Kronecker que aparece em (4.46) é substituído pela matriz de produtos escalares $s_{mn} = \langle v_m | v_n \rangle$.

Não é fácil, em geral, avaliar a qualidade de aproximações para autovalores e autovetores obtidas através do uso de (4.45). No caso do último procedimento descrito, um critério freqüentemente adotado é o de estudar o efeito da inclusão de vetores adicionais no conjunto $\{|v_n\rangle\}$. Uma condição *necessária* (mas não suficiente!) para a boa qualidade da aproximação para determinados autovalores e autovetores é a sua insensibilidade à ampliação do sub-espço variacional. Em situações típicas, diferentes autovetores mostram sensibilidade também diferente a essa ampliação, e aos autovetores mais estáveis é usualmente atribuído um grau maior de confiabilidade que aos menos estáveis.

4.3 Aproximação semi-clássica para a função de onda.

A discussão das propriedades gerais da dinâmica de uma partícula quântica sujeita a um potencial externo, apresentada na seção 3.3.1 pode ser adaptada no sentido de dar um esquema de aproximação útil no regime dito semi-clássico, ou “de grandes números quânticos”, no qual o último termo da equação (3.19) pode ser ignorado frente aos demais. Esta seção trata de desenvolver esse esquema em sua versão usual, que se limita à consideração de soluções estacionárias de problemas unidimensionais. Ele é no entanto aplicável também em situações mais gerais, algumas das quais mencionadas na lista de referências ao final do capítulo.

Um estado estacionário de energia E para um problema unidimensional cujo hamiltoniano é da forma

$$H = \frac{p^2}{2M} + V(x)$$

pode ser descrito por uma função de onda do tipo

$$\psi(x, t) = \phi(x)e^{-\frac{i}{\hbar}Et} \equiv A(x)e^{\frac{i}{\hbar}[S(x)-Et]} \quad (4.47)$$

onde as funções reais $A(x)$ e $S(x)$ correspondem ao módulo e à fase da $\phi(x)$, respectivamente. O procedimento usado anteriormente para se obter a equação (3.19) dá, neste caso particular,

$$\left(\frac{dS}{dx}\right)^2 = 2M[E - V(x)] + \hbar^2 \frac{1}{A(x)} \frac{d^2 A(x)}{dx^2} \quad (4.48)$$

enquanto a equação correspondente à (3.17) é, tendo em conta o fato de que $\partial A/\partial t = 0$,

$$A(x) \frac{d^2 S(x)}{dx^2} + 2 \frac{dA(x)}{dx} \frac{dS(x)}{dx} = 0. \quad (4.49)$$

Se, no entanto, a dinâmica do sistema se dá no regime semi-clássico, o último termo de (4.48) é muito menor que o termo precedente, isto é

$$\hbar^2 \frac{1}{A(x)} \frac{d^2 A(x)}{dx^2} \ll 2M[E - V(x)] \quad \text{ou seja} \quad \frac{\hbar^2}{2M[E - V(x)]} \frac{1}{A(x)} \frac{d^2 A(x)}{dx^2} \ll 1.$$

O primeiro fator desta última relação é na realidade o inverso do quadrado do vetor de onda *local*, $1/k^2(x) \equiv (\lambda(x)/2\pi)^2$, de modo que a última condição implica em “comprimentos de onda locais” $\lambda(x)$ pequenos e/ou amplitudes $A(x)$ de pequena curvatura, no sentido que

$$\frac{\lambda^2(x)}{A(x)} \frac{d^2 A(x)}{dx^2} \ll (2\pi)^2.$$

É imediato que essa condição é fatalmente violada nas proximidades de pontos de inversão clássicos do movimento x_I , nos quais $V(x_I) = E$. Numa região em que a condição esteja satisfeita é possível, porém, usar a aproximação

$$\frac{dS}{dx} \simeq \pm \sqrt{2M[E - V(x)]} = \pm \frac{2\pi\hbar}{\lambda(x)}$$

da qual resulta

$$S(x) \simeq S(x_0) \pm \int_{x_0}^x dx' \sqrt{2M[E - V(x')]} = S(x_0) \pm 2\pi\hbar \int_{x_0}^x \frac{dx'}{\lambda(x')}.$$

A equação (4.49), por outro lado, pode ser resolvida para $A(x)$ em termos de dS/dx como

$$A(x) = C \left(\frac{dS(x)}{dx} \right)^{-1/2}$$

onde C é uma constante relacionada com a normalização da função de onda. Note que esta última relação é *exata*. Usando a expressão aproximada obtida no regime semi-clássico para $S(x)$ ela dá para a forma geral da função de onda

$$\psi(x, t) \simeq C'_+ \sqrt{\frac{\lambda(x)}{2\pi}} e^{-\frac{i}{\hbar}Et + 2\pi i \int_{x_0}^x \frac{dx'}{\lambda(x')}} + C'_- \sqrt{\frac{\lambda(x)}{2\pi}} e^{-\frac{i}{\hbar}Et - 2\pi i \int_{x_0}^x \frac{dx'}{\lambda(x')}} \quad (4.50)$$

onde a dependência com $S(x_0)$ foi incluída nas constantes C'_\pm . A determinação dessas constantes depende de condições adicionais apropriadas impostas sobre $\psi(x, t)$. Elas são fixadas, por exemplo, impondo o valor da função de onda e de sua derivada em algum ponto do domínio de validade da aproximação.

Essa expressão tem propriedades intuitivamente razoáveis: tendo em vista que $\lambda(x)$ é inversamente proporcional à raiz quadrada da energia cinética local, a densidade de probabilidade resulta inversamente proporcional à velocidade local. Por outro lado, $S(x)$, em unidades de \hbar , é a fase acumulada medindo distâncias em unidades do comprimento de onda *local* $\lambda(x)$. Essa fase pode ser vista como uma extensão da fase da função de onda de uma partícula livre $e^{\pm ikx}$, tendo em conta que o vetor de onda k varia lentamente com x , o que corresponde à substituição

$$kx \longrightarrow \int^x dx' k(x').$$

A implicação dessa substituição é a de que o efeito dominante do potencial $V(x)$ é o de alterar o momento da partícula, sem outras complicações ondulatórias (reflexões, interferências, etc.).

Uma restrição séria sobre a aplicabilidade desses resultados é a de que $A(x)$ e $S(x)$ sejam reais, o que implica, em particular, que eles só possam ser usados em regiões *classicamente permitidas*, nas quais $E > V(x)$. Essa restrição pode, no entanto, ser eliminada introduzindo modificações apropriadas no procedimento de dedução. Essas modificações, que serão discutidas a seguir, podem ser feitas de modo a permitir também sistematizar o esquema de aproximação, identificando uma hierarquia de correções à aproximação mais simples, que corresponde essencialmente ao esquema descrito acima.

4.3.1 A aproximação de Wentzel, Kramers e Brillouin (WKB)⁵.

O ponto de partida para a obtenção da aproximação semi-clássica (4.50) acima foi a representação (4.47) de $\psi(x, t)$ em termos das duas funções $A(x)$ e $S(x)$, o que leva à substituição da equação de Schrödinger pelo par de equações (4.48) e (4.49). No entanto, essas mesmas equações podem ser obtidas através de um critério para a determinação de $A(x)$ e $S(x)$ que não impõe a realidade dessas duas funções. O novo critério pode ser motivado observando que a aproximação semi-clássica para a equação (4.48) resulta independente de \hbar , enquanto correções devidas ao último termo são proporcionais a \hbar^2 , pelo menos. Isso sugere substituir a decomposição (4.47) por

$$\psi(x, t) = e^{\frac{i}{\hbar}W(x, t)} \quad \text{com} \quad W(x, t) \equiv S(x, t) - i\hbar T(x, t) \quad (4.51)$$

onde as funções $S(x, t)$ e $T(x, t)$ são definidas de modo a conter apenas potências *pares* de \hbar . Isso corresponde à decomposição da função complexa $W(x, t)$ em duas partes, que envolvem respectivamente potências pares, $S(x, t)$ e ímpares, $-i\hbar T(x, t)$, de \hbar . Em particular, não há agora qualquer restrição de realidade sobre essas duas funções. Um tratamento que se limite aos termos de ordem mais baixa em \hbar deve assim coincidir com o tratamento anterior, nas regiões em que ele seja aplicável, com uma amplitude $A(x, t)$ dada por

$$A(x, t) = e^{T(x, t)}. \quad (4.52)$$

Para implementar esse esquema basta agora substituir o *ansatz* (4.51) na equação de movimento

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2M} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t)$$

e escrever as duas equações que resultam da separação de potências pares e ímpares de \hbar . Este último passo na realidade substitui a separação de termos reais e imaginários envolvida no tratamento anterior, e leva ao par de equações

$$\begin{aligned} -\frac{\hbar^2}{2M} \frac{\partial^2 A}{\partial x^2} + \frac{A}{2M} \left(\frac{\partial S}{\partial x} \right)^2 + V(x)A + A \frac{\partial S}{\partial t} &= 0 \quad \leftrightarrow \quad \text{potências pares} \\ \frac{1}{M} \frac{\partial A}{\partial x} \frac{\partial S}{\partial x} + \frac{A}{2M} \frac{\partial^2 S}{\partial x^2} + \frac{\partial A}{\partial t} &= 0 \quad \leftrightarrow \quad \text{potências ímpares} \end{aligned} \quad (4.53)$$

⁵G. Wentzel, *Zeit. Phys.* **38**, 518 (1926); H. A. Kramers, *Zeit. Phys.* **39**, 828 (1926); L. Brillouin, *C. R. Acad. Sci. Paris* **183**, 24 (1926).

onde a amplitude $A(x, t)$ corresponde à expressão (4.52). Essas equações são na realidade idênticas às (3.16) e (3.17) obtidas no capítulo anterior para a amplitude e fase *reais* da função de onda. A primeira delas pode ser re-escrita sob a forma ref.(3.19), e a particularização para um estado estacionário de energia E corresponde a tomar $S(x, t)$ e $A(x, t)$ como sendo da forma

$$S(x, t) \longrightarrow -Et + S_E(x) \quad \text{e} \quad A(x, t) \longrightarrow A_E(x)$$

com $S_E(x)$ e $A_E(x)$ independentes do tempo mas já não necessariamente reais. Nesse caso a função de onda tem a forma

$$\psi(x, t) \longrightarrow A_E(x) e^{\frac{i}{\hbar} S_E(x)} e^{-\frac{i}{\hbar} Et} \equiv \phi_E(x) e^{-\frac{i}{\hbar} Et}$$

e as equações (4.53) se reduzem a

$$\left(\frac{dS_E}{dx} \right)^2 = 2M[E - V(x)] + \hbar^2 \frac{1}{A_E(x)} \frac{d^2 A_E(x)}{dx^2} \quad (4.54)$$

e

$$A_E(x) \frac{d^2 S_E(x)}{dx^2} + 2 \frac{dA_E(x)}{dx} \frac{dS_E(x)}{dx} = 0 \quad (4.55)$$

que são idênticas às equações (4.48) e (4.49), respectivamente, sem que exista, no entanto, uma condição de realidade imposta sobre as funções $S_E(x)$ e $A_E(x)$. As equações (4.54) e (4.55) são exatamente equivalentes à forma estacionária da equação de Schrödinger. A segunda delas tem por solução

$$A_E(x) = C \left(\frac{dS_E(x)}{dx} \right)^{-1/2}$$

que pode ser usada para eliminar a dependência da primeira com $A_E(x)$, levando dessa forma à equação para $S_E(x)$

$$\begin{aligned} \left(\frac{dS_E(x)}{dx} \right)^2 &= 2M[E - V(x)] + \\ &+ \frac{\hbar^2}{2} \left[\frac{3}{2} \left(\frac{d^2 S_E(x)}{dx^2} \right)^2 \left(\frac{dS_E(x)}{dx} \right)^{-2} - \left(\frac{d^3 S_E(x)}{dx^3} \right) \left(\frac{dS_E(x)}{dx} \right)^{-1} \right], \end{aligned} \quad (4.56)$$

também uma equação exata.

Soluções aproximadas podem ser obtidas através de uma “expansão semi-clássica”, que exprime $S_E(x)$ como uma série de potências em \hbar^2

$$S_E(x) = S_0(x) + \hbar^2 S_2(x) + \hbar^4 S_4(x) + \dots$$

e a aproximação de ordem mais baixa dá novamente

$$\left(\frac{dS_0(x)}{dx} \right)^2 = 2M[E - V(x)].$$

No entanto, não existe agora uma condição de realidade imposta sobre $S_0(x)$ e portanto, embora a sendo essa aproximação inválida nas vizinhanças de pontos de inversão do movimento clássico, ela pode ser utilizada também em regiões classicamente *proibidas*, nas quais $E < V(x)$. Existem portanto duas situações distintas a considerar:

1. $E > V(x)$. Este é o caso já tratado anteriormente, no qual

$$S_0^{(>)}(x) = S_0^{(>)}(x_0) \pm \int_{x_0}^x dx' \sqrt{2M[E - V(x')]} = S^{(>)}(x_0) \pm 2\pi\hbar \int_{x_0}^x \frac{dx'}{\lambda(x')}$$

com a aproximação correspondente para a função de onda

$$\phi_E^{(>)}(x) \simeq C_+^{(>)} \sqrt{\frac{\lambda(x)}{2\pi}} e^{2\pi i \int_{x_0}^x \frac{dx'}{\lambda(x')}} + C_-^{(>)} \sqrt{\frac{\lambda(x)}{2\pi}} e^{-2\pi i \int_{x_0}^x \frac{dx'}{\lambda(x')}}. \quad (4.57)$$

2. $E < V(x)$. Este caso corresponde a regiões classicamente proibidas, e leva a

$$S_0^{(<)}(x) = S_0^{(<)}(x_0) \pm i \int_{x_0}^x dx' \sqrt{2M[V(x') - E]} = S^{(<)}(x_0) \pm 2\pi i \hbar \int_{x_0}^x \frac{dx'}{l(x')}$$

onde

$$l(x) = \frac{2\pi\hbar}{\sqrt{2M[V(x) - E]}}, \quad V(x) > E$$

e a aproximação correspondente para a função de onda é em geral uma combinação linear de exponenciais reais, crescentes e decrescentes

$$\phi_E^{(<)}(x) \simeq C_+^{(<)} \sqrt{\frac{l(x)}{2\pi}} e^{2\pi \int_{x_0}^x \frac{dx'}{l(x')}} + C_-^{(<)} \sqrt{\frac{l(x)}{2\pi}} e^{-2\pi \int_{x_0}^x \frac{dx'}{l(x')}}. \quad (4.58)$$

As expressões (4.57) e (4.58) são as mais comumente usadas no contexto da aproximação WKB. Correções da ordem \hbar^2 podem ser obtidas voltando à equação (4.56) e usando o resultado obtido para a aproximação de ordem zero. A equação resultante para $S_2(x)$ é

$$\frac{dS_2}{dx} = \frac{3}{8} \left(\frac{d^2 S_0(x)}{dx^2} \right)^2 \left(\frac{dS_0(x)}{dx} \right)^{-3} - \frac{1}{4} \left(\frac{d^3 S_0(x)}{dx^3} \right) \left(\frac{dS_0(x)}{dx} \right)^{-2}.$$

O segundo termo dessa equação pode ser calculado sem dificuldade usando as relações de ordem zero

$$\frac{dS_0^{(>)}}{dx} = \pm \frac{2\pi\hbar}{\lambda(x)} \quad \text{e} \quad \frac{dS_0^{(>)}}{dx} = \pm \frac{2\pi i\hbar}{l(x)}$$

o que leva aos resultados

$$\frac{dS_2^{(>)}}{dx} = \pm \frac{1}{8\pi\hbar} \left[\frac{d^2 \lambda(x)}{dx^2} - \frac{1}{2\lambda(x)} \left(\frac{d\lambda(x)}{dx} \right)^2 \right]$$

e

$$\frac{dS_2^{(<)}}{dx} = \mp \frac{i}{8\pi\hbar} \left[\frac{d^2 l(x)}{dx^2} - \frac{1}{2l(x)} \left(\frac{dl(x)}{dx} \right)^2 \right]$$

ou ainda, fazendo uma integração em x ,

$$S_2(x) = \pm \frac{1}{8\pi\hbar} \left[\frac{dL(x)}{dx} - \frac{1}{2} \int^x dx' \frac{1}{L(x')} \left(\frac{dL(x')}{dx'} \right)^2 \right] \quad (4.59)$$

onde os dois casos $E > V(x)$ e $E < V(x)$ correspondem às definições de $L(x)$

$$L(x) = \lambda(x), \quad E > V(x) \quad \text{e} \quad L(x) = -il(x), \quad E < V(x).$$

A rigor, a expressão para $S_2(x)$ deve conter uma constante de integração, que pode no entanto ser incorporada aos fatores multiplicativos que aparecem nas amplitudes da função de onda. As amplitudes corrigidas pelos efeitos de ordem \hbar^2 , de fato, têm a forma

$$\begin{aligned} A_E(x) &\simeq C \left(\frac{dS_0(x)}{dx} + \hbar^2 \frac{dS_2(x)}{dx} \right)^{-1/2} \\ &\simeq C \left(\frac{dS_0(x)}{dx} \right)^{-1/2} \left[1 - \frac{\hbar^2}{2} \left(\frac{dS_2(x)}{dx} \right) \left(\frac{dS_0(x)}{dx} \right)^{-1} \right] \\ &= A_0(x) \left[1 - \frac{\hbar^2}{2} \left(\frac{dS_2(x)}{dx} \right) \left(\frac{dS_0(x)}{dx} \right)^{-1} \right] \end{aligned} \quad (4.60)$$

que representa na realidade quatro diferentes tipos de amplitude, correspondentes respectivamente às soluções com cada um dos dois sinais de $\pm\sqrt{|E - V(x)|}$, para os casos em que $E > V(x)$ e $E < V(x)$.

Uma utilidade imediata das expressões (4.59) e (4.60) para as correções de ordem \hbar^2 é o estabelecimento de critérios de validade para a aproximação de ordem mais baixa. De fato, a validade da aproximação de ordem zero

$$\phi_E(x) \simeq A_0(x)e^{\frac{i}{\hbar}S_0(x)}$$

depende de estarem satisfeitas as condições

$$|\hbar S_2(x)| \ll 1 \quad \text{e} \quad \left| \frac{\hbar^2}{2} \left(\frac{dS_2(x)}{dx} \right) \left(\frac{dS_0(x)}{dx} \right)^{-1} \right| \ll 1,$$

que exprimem a irrelevância das correções de segunda ordem respectivamente para o fator exponencial e para a amplitude da função de onda aproximada.

Em vista da expressão (4.59) para S_2 , é claro que a primeira dessas duas condições estará satisfeita sempre que o módulo da derivada $dL(x)/dx$ for suficientemente pequeno em toda a região que esteja sendo considerada para a variável x . Esse critério pode ser expresso em termos do potencial e da energia E do estado estacionário como

$$\left| \frac{dL(x)}{dx} \right| \ll 1 \quad \longrightarrow \quad \frac{2\pi M\hbar \left| \frac{dV(x)}{dx} \right|}{2M[E - V(x)]^{3/2}} = \frac{1}{2} \frac{\left| L(x) \frac{dV(x)}{dx} \right|}{|E - V(x)|} \ll 1 \quad (4.61)$$

mostrando mais uma vez o colapso da aproximação nas vizinhanças dos pontos de inversão do movimento clássico. A segunda forma da expressão tem ainda uma interpretação clara: a estimativa linear para a variação do potencial ao longo de um comprimento de onda local (ou, no caso $E < V(x)$, ao longo de uma distância igual ao módulo do comprimento de onda local imaginário) deve ser muito menor que a energia cinética local (ou que o seu módulo, no caso $E < V(x)$ em que ela é negativa).

A condição referente à amplitude pode ser também ser escrita em termos de $L(x)$ usando as expressões obtidas para dS_0/dx e para dS_2/dx . Disso resulta a condição

$$\frac{1}{32\pi^2} \left| L(x) \frac{d^2 L(x)}{dx^2} - \frac{1}{2} \left(\frac{dL(x)}{dx} \right)^2 \right| \ll 1$$

que envolve explicitamente também a *segunda derivada* do comprimento de onda local (real ou imaginário para $E > V(x)$ ou $E < V(x)$ respectivamente). Isto introduz a restrição adicional

à expressa por (4.61), a saber, que a variação com x de dL/dx seja suficientemente suave. Mesmo quando o valor absoluto desta derivada seja pequeno em toda a região examinada, variações descontínuas, ou quase descontínuas, desta derivada acarretam valores muito grandes de d^2L/dx^2 . As propriedades do potencial $V(x)$ que levam a situações desse tipo podem ser identificadas calculando

$$\left| L(x) \frac{d^2L(x)}{dx^2} \right| = |L(x)|^2 \left| \frac{3}{4} \frac{\left(\frac{dV(x)}{dx} \right)^2}{[E - V(x)]^2} + \frac{1}{2} \frac{\frac{d^2V(x)}{dx^2}}{E - V} \right|$$

que aponta como ingrediente perigoso uma descontinuidade ou quase descontinuidade na derivada do potencial (isto é, uma força clássica descontínua ou quase-descontínua), o que faz com que a segunda derivada assuma valores absolutos muito grandes.

4.3.2 Tratamento de regiões perigosas: fórmulas de conexão.

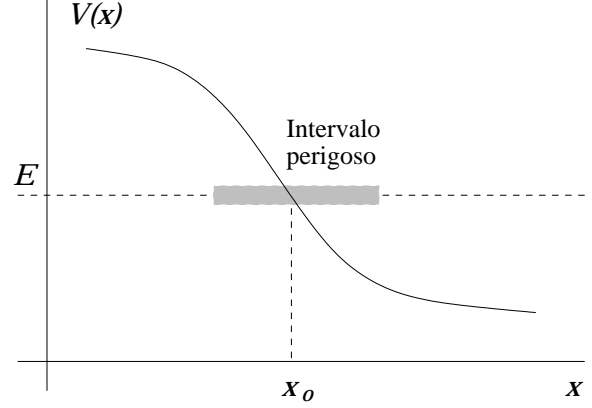
A quebra da validade da aproximação WKB de ordem mais baixa em determinados intervalos da variável x (notadamente nas vizinhanças dos pontos de inversão do movimento clássico) é uma limitação séria do ponto de vista de suas aplicações. As dificuldades decorrentes disso podem, no entanto, ser contornadas através de um tratamento especial da função de onda nos intervalos perigosos. A idéia central consiste em utilizar, em cada um desses intervalos, seja a solução exata, seja outra aproximação válida para a função de onda, e em seguida ligar continuamente essa solução especial às dadas pela aproximação WKB nas fronteiras do intervalo perigoso considerado. Em vários casos de interesse o resultado desse tipo de procedimento pode ser expresso sob a forma de prescrições que relacionam as funções de onda da aproximação WKB nos dois extremos do intervalo. Tais prescrições são conhecidas como “fórmulas de conexão”.

1. Ponto de inversão linear. A situação perigosa mais comum é a que envolve um “ponto de inversão linear”, que é um ponto de inversão x_0 do movimento clássico no qual o potencial $V(x)$ pode ser bem aproximado por uma função linear de x , isto é

$$V(x) \simeq E + (x - x_0) \left. \frac{dV}{dx} \right|_{x=x_0} \equiv E - (x - x_0)f_0.$$

Essa situação corresponde à ilustrada na figura 4.3 (note que neste caso $f_0 > 0$), sendo a aproximação WKB supostamente válida tanto na região classicamente permitida à direita do intervalo sombreado como na região classicamente proibida à esquerda desse intervalo. A constante f_0 é claramente a força que age classicamente sobre a partícula no ponto de inversão x_0 . Neste caso, a equação de Schrödinger estacionária pode ser escrita nas vizinhanças de $x = x_0$ como

Figura 4.3: Potencial $V(x)$ com ponto de inversão linear no ponto x_0 à energia E . O intervalo em que a aproximação WKB deixa de ser válida corresponde à região sombreada no gráfico.



$$\frac{d^2 \phi_E}{dx^2} + \frac{2M f_0}{\hbar^2} (x - x_0) \phi_E = 0$$

ou, introduzindo a variável adimensional $y = \left(\frac{2M f_0}{\hbar^2}\right)^{1/3} (x - x_0)$,

$$\frac{d^2 \phi_E}{dy^2} + y \phi_E = 0, \quad (4.62)$$

onde a função ϕ_E é tratada como uma função de y . Essa equação pode ser resolvida exatamente. As suas soluções gerais podem ser expressas em termos das chamadas *funções de Airy*, que podem por sua vez ser escritas em termos de funções de Bessel cilíndricas de ordem $\pm 1/3$, $J_{\pm 1/3}$. Desse modo a solução geral da equação (4.62) pode ser escrita como

$$\phi_E(y) = \sqrt{y} \left[c_1 J_{\frac{1}{3}}\left(\frac{2}{3}y^{\frac{3}{2}}\right) + c_2 J_{-\frac{1}{3}}\left(\frac{2}{3}y^{\frac{3}{2}}\right) \right]. \quad (4.63)$$

As constantes arbitrárias c_1 e c_2 podem ser ajustadas de modo que a solução satisfaça condições de contorno apropriadas.

Fórmulas de conexão apropriadas para este caso podem ser obtidas comparando essa solução com aquela obtida através da aproximação WKB onde ambas sejam válidas. Por exemplo, os coeficientes c_1 e c_2 de (4.63) podem ser escolhidos de forma que

$$\phi_E(y \rightarrow -\infty) \longrightarrow \frac{c}{2} (-y)^{-\frac{1}{4}} e^{-\frac{2}{3}(-y)^{\frac{3}{2}}} + \mathcal{O}(-y^{-\frac{7}{4}})$$

e

$$\phi_E(y \rightarrow +\infty) \longrightarrow c y^{-\frac{1}{4}} \sin\left(\frac{2}{3}y^{\frac{3}{2}} + \frac{\pi}{4}\right) + \mathcal{O}(y^{-\frac{7}{4}})$$

onde c é um fator multiplicativo global, a ser determinado por um critério de normalização. Utilizando a aproximação linear para o potencial,

$$\lambda(x) = \frac{2\pi\hbar}{\sqrt{2M[E - V(x)]}} \simeq \frac{2\pi\hbar}{\sqrt{2M(x - x_0)f_0}} = 2\pi \left(\frac{\hbar^2}{2Mf_0} \right)^{1/3} \frac{1}{\sqrt{y}}$$

de modo que

$$2\pi \int_{x_0}^x \frac{dx}{\lambda(x)} = \int_0^{x-x_0} dy \sqrt{y} = \frac{2}{3} y^{3/2}$$

e o termo dominante da expressão assintótica para $y \rightarrow +\infty$ se reduz a

$$\phi_E(x \gg x_0) \longrightarrow c' \sqrt{\frac{\lambda(x)}{2\pi}} \sin \left(2\pi \int_{x_0}^x \frac{dx}{\lambda(x)} + \frac{\pi}{4} \right)$$

com $c' = c \left(\frac{\hbar^2}{2Mf_0} \right)^{-1/6}$. Um procedimento inteiramente análogo para a expressão assintótica válida para $y \rightarrow -\infty$ dá, por outro lado,

$$\phi_E(x \ll x_0) \longrightarrow \frac{c'}{2} \sqrt{\frac{l(x)}{2\pi}} e^{2\pi \int_{x_0}^x \frac{dx}{l(x)}}.$$

Esta última expressão pode ser identificada com a solução (4.58) da aproximação WKB, na região classicamente proibida⁶, para $C_-^{(<)} = 0$ e $C_+^{(<)} = c'/2$, ao passo que a expressão válida para $x \gg x_0$ corresponde à particular escolha dos coeficientes $C_+^{(>)}$ e $C_-^{(>)}$ da expressão WKB (4.57), válida na região classicamente permitida,

$$C_+^{(>)} = -C_-^{(>)} = \frac{c'}{2i}.$$

Desse modo, o prolongamento para a região classicamente permitida da função de onda da aproximação WKB que envolve a exponencial decrescente na região classicamente proibida (ou vice versa) deve ser feito segundo a prescrição

⁶Note que a integral $\int_{x_0}^x \frac{dx}{l(x)}$ é *negativa* nas condições representadas na figura 4.3, em que o ponto x está à esquerda do ponto x_0 .

$$\begin{array}{ccc}
x \ll x_0 & & x \gg x_0 \\
(\text{região proibida}) & & (\text{região permitida}) \\
\hline
\phi_E(x) = \frac{c}{2} \sqrt{\frac{l(x)}{2\pi}} e^{-2\pi \int_x^{x_0} \frac{dx'}{l(x')}} & \longleftrightarrow & \phi_E(x) = c \sqrt{\frac{\lambda(x)}{2\pi}} \operatorname{sen} \left(2\pi \int_{x_0}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \\
\hline
\end{array} \tag{4.64}$$

que é uma das fórmulas de conexão relevantes para o caso de um ponto de inversão linear como o representado na figura 4.3.

Como indicado na equação (4.58), a forma geral da aproximação WKB para a função de onda na região classicamente proibida consiste porém numa superposição de contribuições com dependências exponenciais crescentes e decrescentes. Além da fórmula de conexão para a componente exponencial decrescente, é preciso portanto estudar também o prolongamento para a região permitida da componente crescente. Ele pode ser obtido repetindo o procedimento acima com uma escolha diferente dos coeficientes c_1 e c_2 na equação (4.63), e leva a

$$\begin{array}{ccc}
x \ll x_0 & & x \gg x_0 \\
(\text{região proibida}) & & (\text{região permitida}) \\
\hline
\phi_E(x) = c \sqrt{\frac{l(x)}{2\pi}} e^{2\pi \int_x^{x_0} \frac{dx'}{l(x')}} & \longleftrightarrow & \phi_E(x) = c \sqrt{\frac{\lambda(x)}{2\pi}} \cos \left(2\pi \int_{x_0}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \\
\hline
\end{array} \tag{4.65}$$

em que deve ser notada a ausência de um fator $1/2$ na expressão que corresponde à região proibida, em comparação com o caso (4.64).

Um ponto que deve ser observado com relação a esse procedimento para obter as fórmulas de conexão é o de que as funções de onda na aproximação WKB foram comparadas às formas *assintóticas* da solução da equação (4.62), que envolve a aproximação linear para o potencial nas vizinhanças do ponto de inversão. A implicação disso é que a validade das fórmulas obtidas depende de que os parâmetros do problema sejam tais que a aproximação WKB seja válida

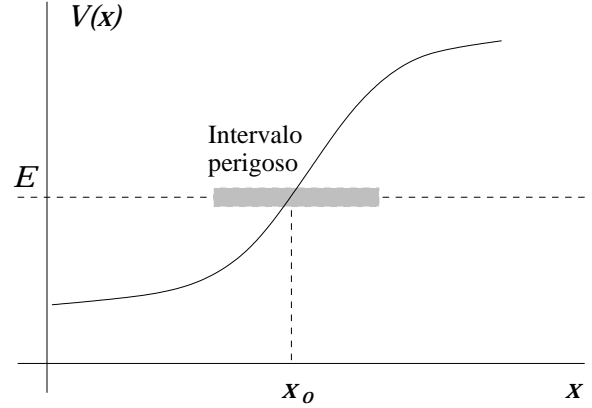
numa região em que tanto a forma assintótica da solução para o potencial linear quanto e própria forma linearizada do potencial sejam válidas. Para que isso seja possível, é preciso que em regiões suficientemente próximas do ponto de inversão para que efeitos de curvatura do potencial sejam desprezíveis, os valores do argumento y se tornem suficientemente grandes para permitir o uso das expressões assintóticas e da aproximação WKB (v. equação (4.61)). Isso ocorrerá se $\lambda(x)$ e $l(x)$ se tornarem suficientemente pequenos dentro da região em que a aproximação linear para o potencial possa ser usada.

As fórmulas de conexão (4.64) e (4.65) foram obtidas para um ponto de inversão linear x_0 com as características da figura 4.3, em que a região classicamente proibida se situa em $x < x_0$. No caso em que essa região esteja à direita do ponto de inversão (v. figura 4.4), $f_0 = -dV/dx|_{x=x_0} < 0$, mas a mesma equação (4.62) é obtida para a função de onda ϕ_E quando considerada função da nova variável $y = -\left(\frac{2M|f_0|}{\hbar^2}\right)^{1/3} (x - x_0)$. Isso permite adaptar facilmente os resultados (4.64) e (4.65) a essa situação. As soluções de (4.62) relevantes para as fórmulas de conexão são, em termos da variável x , simplesmente formas refletidas especularmente em $x = x_0$ das soluções utilizadas no caso ilustrado na figura 4.3. Desse modo as fórmulas de conexão num caso como o ilustrado na figura 4.4 são dadas no quadro abaixo.

$x \ll x_0$	$x \gg x_0$	
(região permitida)	(região proibida)	
$\phi_E(x) = c\sqrt{\frac{\lambda(x)}{2\pi}} \operatorname{sen} \left(2\pi \int_x^{x_0} \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \longleftrightarrow \phi_E(x) = \frac{c}{2}\sqrt{\frac{l(x)}{2\pi}} e^{-2\pi \int_{x_0}^x \frac{dx'}{l(x')}} \quad (4.66)$		
$\phi_E(x) = c\sqrt{\frac{\lambda(x)}{2\pi}} \cos \left(2\pi \int_x^{x_0} \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \longleftrightarrow \phi_E(x) = c\sqrt{\frac{l(x)}{2\pi}} e^{2\pi \int_{x_0}^x \frac{dx'}{l(x')}} \quad (4.66)$		

2. Pontos em que a força clássica é descontínua. Como mencionado no final da seção 4.3.1, os critérios de validade para a aproximação WKB de ordem mais baixa baseados na avaliação da magnitude das correções de segunda ordem excluem pontos ou regiões em que a força clássica (isto é, a derivada do potencial, $dV(x)/dx$) variam de forma descontínua, ou quase-descontínua, pelo fato de que a correção à amplitude envolve a segunda derivada do potencial. Um exemplo simples de uma situação desse tipo é dado pelo potencial

Figura 4.4: Potencial $V(x)$ semelhante ao da figura 4.3, mas com a região classicamente proibida à direita do ponto de inversão x_0 .



$$V(x) = -f_0|x|, \quad f_0 > 0$$

para o qual, a uma energia $E > 0$, o comprimento de onda local é dado por

$$\lambda(x) = \frac{2\pi\hbar}{\sqrt{2M(E + f_0|x|)}} \simeq \lambda_0 \left(1 - \frac{f_0}{2E}|x| \right), \quad \lambda_0 \equiv \lambda(x=0) = \frac{2\pi\hbar}{\sqrt{2ME}}.$$

Esta situação é ilustrada na figura 4.5. A aproximação usada na última expressão é válida para $f_0|x|/E \ll 1$. Por outro lado, $|d\lambda(x)/dx| = \lambda_0 f_0/2E$ para $x \neq 0$. A condição de que esta última quantidade seja também muito menor que a unidade (ligada à validade da aproximação WKB) é portanto de natureza a permitir que a expansão seja válida numa região extensa de x , quando medida em unidades de λ_0 .

Uma aplicação direta da aproximação WKB a esse problema leva a uma função de onda cuja forma geral é (note que um fator $(2\pi)^{-1/2}$ da amplitude foi absorvido nas constantes C_{\pm})

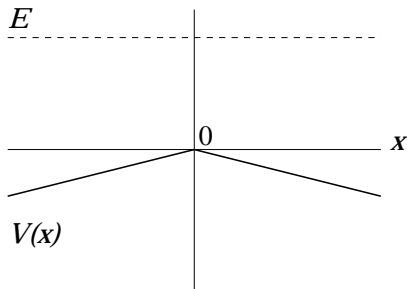


Figura 4.5: Problema estacionário à energia E com potencial $V(x) = -f_0|x|$. A força clássica é descontínua em $x = 0$.

$$\phi_E(x) \simeq \sqrt{\lambda_0 \left(1 - \frac{f_0}{2E}|x|\right)} \left[C_+ e^{2\pi i \int_0^x \frac{dx'}{\lambda_0 \left(1 - \frac{f_0}{2E}|x'|\right)}} + C_- e^{-2\pi i \int_0^x \frac{dx'}{\lambda_0 \left(1 - \frac{f_0}{2E}|x'|\right)}} \right]. \quad (4.67)$$

É imediato verificar, porém, que essa aproximação para a função de onda tem uma derivada descontínua no ponto $x = 0$, no qual a derivada do potencial não é contínua. Essa descontinuidade na derivada da função de onda é, na realidade, inconsistente com a equação de Schrödinger estacionária

$$\frac{d^2 \phi_E(x)}{dx^2} + \frac{2M}{\hbar^2} (E + f_0|x|) \phi_E(x) = 0, \quad (4.68)$$

a qual exige a continuidade da primeira derivada de $\phi_E(x)$ não apenas nesse ponto como também em qualquer ponto x' , como pode ser visto facilmente integrando a equação num pequeno intervalo $x' - \epsilon < x < x' + \epsilon$ e tomando o limite $\epsilon \rightarrow 0$.

Uma forma de contornar o problema da quebra de validade da aproximação WKB em $x = 0$, dada a sua validade tanto na região $x > 0$ como na região $x < 0$, consiste em escrever, em cada uma dessas duas regiões, expressões do tipo (4.67) para a função de onda, relacionando em seguida as constantes associadas às amplitudes de modo a garantir a continuidade da função de onda e também de sua derivada na origem. Este último critério corresponde, de fato, a implementar na função de onda uma condição exigida pela equação de Schrödinger no ponto problemático. Desse modo, a forma geral da função de onda passa a ser tomada como sendo

$$\phi_E(x) \simeq \begin{cases} \sqrt{\lambda_0 \left(1 - \frac{f_0}{2E}x\right)} \left[C_+ e^{2\pi i \int_0^x \frac{dx'}{\lambda_0 \left(1 - \frac{f_0}{2E}x'\right)}} + C_- e^{-2\pi i \int_0^x \frac{dx'}{\lambda_0 \left(1 - \frac{f_0}{2E}x'\right)}} \right] & \text{para } x > 0 \\ \sqrt{\lambda_0 \left(1 + \frac{f_0}{2E}x\right)} \left[D_+ e^{-2\pi i \int_x^0 \frac{dx'}{\lambda_0 \left(1 + \frac{f_0}{2E}x'\right)}} + D_- e^{2\pi i \int_x^0 \frac{dx'}{\lambda_0 \left(1 + \frac{f_0}{2E}x'\right)}} \right] & \text{para } x < 0 \end{cases} \quad (4.69)$$

com as condições

$$\phi_E(0_+) = \phi_E(0_-) \quad \text{e} \quad \frac{d\phi_E(0_+)}{dx} = \frac{d\phi_E(0_-)}{dx}$$

que, de forma explícita, dão

$$C_+ + C_- = D_+ + D_-$$

e

$$-\left(\frac{\lambda_0 f_0}{4E} - 2\pi i\right) C_+ - \left(\frac{\lambda_0 f_0}{4E} + 2\pi i\right) C_- = \left(\frac{\lambda_0 f_0}{4E} + 2\pi i\right) D_+ + \left(\frac{\lambda_0 f_0}{4E} - 2\pi i\right) D_-.$$

O significado das diferentes componentes dessa função de onda está indicado na figura 4.6: as componentes envolvendo C_- e D_+ representam partículas que se movem em direção à origem, respectivamente nas regiões $x > 0$ e $x < 0$, enquanto as demais representam partículas que se afastam da origem.

O fato de haver duas relações a serem satisfeitas por quatro amplitudes reflete a possibilidade de obter soluções aproximadas satisfazendo diferentes condições de contorno nas regiões $x < 0$ e $x > 0$. Tendo em conta que uma das constantes é fixada por um critério de normalização, a liberdade remanescente permite ainda a escolha arbitrária de *uma* das quatro constantes. Por exemplo, uma solução que exclua partículas incidindo sobre a origem pela direita pode ser obtida fazendo $C_- = 0$. Escolhendo $D_+ = 1$ como critério de normalização (que fixa a amplitude da componente que representa partículas incidentes pela esquerda), resulta então

$$C_+ = 1 + D_- \quad \text{e} \quad -\left(\frac{\lambda_0 f_0}{4E} - 2\pi i\right) C_+ = \left(\frac{\lambda_0 f_0}{4E} + 2\pi i\right) + \left(\frac{\lambda_0 f_0}{4E} - 2\pi i\right) D_-$$

ou seja

$$C_+ = \frac{1}{1 + i\frac{\lambda_0 f_0}{8\pi E}} \quad \text{e} \quad D_- = -\frac{i\frac{\lambda_0 f_0}{8\pi E}}{1 + i\frac{\lambda_0 f_0}{8\pi E}}. \quad (4.70)$$

O coeficiente D_- nesse caso tem o significado de um coeficiente de reflexão, pois aparece como o fator multiplicativo da componente de ϕ_E que se afasta da origem na região $x < 0$, enquanto C_+ tem o significado de um coeficiente de transmissão. A remoção da descontinuidade da derivada da aproximação WKB para a função de onda no ponto em que a força clássica tem uma descontinuidade de magnitude $2f_0$ leva, pois, ao aparecimento de uma onda refletida de

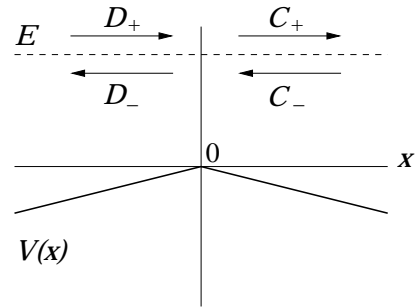


Figura 4.6: Significado das diferentes componentes da aproximação (4.69) para a função de onda à energia E no potencial representado na figura 4.5.

amplitude proporcional a $\lambda_0 f_0 / 8\pi E$, e reduz a amplitude transmitida por um termo da mesma ordem.

Além de prescrever a continuidade da função de onda e de sua derivada em $x = 0$, a equação (4.68) prescreve ainda a continuidade da segunda derivada da função de onda, bem como dá o seu valor em $x = 0$ como sendo

$$\frac{1}{\phi_E(x)} \frac{d^2 \phi_E(x)}{dx^2} \Big|_{x=0} = -\frac{2ME}{\hbar^2} = -\frac{(2\pi)^2}{\lambda_0^2}.$$

Calculando essa mesma quantidade usando a função de onda aproximada, dada por (4.69) e (4.70), o resultado que se obtém é

$$\frac{1}{\phi_E(x)} \frac{d^2 \phi_E(x)}{dx^2} \Big|_{x=0} \simeq -\frac{(2\pi)^2}{\lambda_0^2} \left(1 + \frac{f_0^2 \lambda_0^2}{(8\pi)^2 E^2} \right)$$

que difere do resultado exato apenas por uma contribuição ligada ao último termo da equação (4.54), o qual é pequeno nas condições de aplicabilidade da aproximação WKB e simplesmente truncado no caso da aproximação de ordem mais baixa.

Como observação final, cabe notar que funções de onda *exatas* para o caso particular do potencial representado na figura 4.5 podem ser obtidas em termos das soluções da equação (4.62). A utilidade da aproximação WKB adaptada para garantir a continuidade da derivada da função de onda se limita, nesse caso, à eventual conveniência de sua forma analítica mais explícita. Por outro lado, essa aproximação pode ser imediatamente estendida a casos em que o potencial $V(x)$, tendo uma derivada descontínua em um determinado ponto $x = x_d$, seja de natureza não linear para $x \neq x_d$, embora sendo ainda suficientemente suave para permitir aí o uso da aproximação WKB.

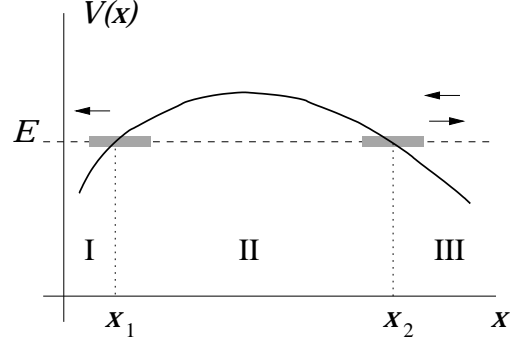
4.3.3 Uso das fórmulas de conexão: penetração de barreiras de potencial e regras de quantização de Bohr.

O uso mais comum das fórmulas de conexão para pontos de inversão lineares se relaciona com o cálculo semi-clássico da penetração de barreiras de potencial. O interesse desse problema remonta a um trabalho clássico de 1928 de G. Gamow⁷, em que ele é relacionado com a forte dependência das vidas médias observadas para o decaimento alfa de núcleos radioativos com a energia das partículas alfa emitidas.

O problema consiste, de uma forma geral, em obter uma aproximação para a função de onda no caso de um potencial como o indicado na figura 4.7. À energia E esse potencial tem

⁷G. Gamow, *Zeitschrift für Physik* **51**, 204 (1928). Esse trabalho está reproduzido em *fac simile* (não traduzido!) na coletânea *Foundations of Nuclear Physics*, Robert T. Beyer (ed.), Dover, N.Y., 1949.

Figura 4.7: Barreira de potencial com dois pontos de inversão lineares x_1 e x_2 à energia E . A aproximação WKB é supostamente válida nas regiões I, II e III, excluídos os “intervalos perigosos” em torno dos pontos de inversão.



dois pontos de inversão clássicos em $x = x_1$ e $x = x_2$, que delimitam a região II, classicamente proibida. As fórmulas de conexão permitem determinar a aproximação WKB para a função de onda na região II em termos da aproximação correspondente na região I, e em seguida a aproximação WKB na região III em termos da correspondente na região II. Isso permite então relacionar as aproximações WKB para as funções de onda nas regiões classicamente permitidas I e III, e com isso estimar a probabilidade de tunelamento através da região classicamente proibida II.

Para obter a probabilidade de tunelamento, é conveniente partir da condição de contorno de que na região I há apenas partículas que se movem para a esquerda, como indicado pela seta na região I da figura 4.7. A solução, na aproximação WKB, que tem essas propriedades na região I pode ser escrita como

$$\phi_E^{(I)}(x) = c \sqrt{\frac{\lambda(x)}{2\pi}} e^{i \left[2\pi \int_{x_1}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right]}, \quad x \ll x_1$$

onde c é uma amplitude arbitrária e a fase $\pi/4$ foi introduzida explicitamente para conveniência posterior, no uso das fórmulas de conexão. A condição $x \ll x_1$ significa que deve ser devida-mente excluído o intervalo perigoso em torno do ponto de inversão linear x_1 . Escrevendo a exponencial em termos de seno e cosseno, e usando as fórmulas de conexão (4.66), resulta que a aproximação WKB para a função de onda na região II tem a forma

$$\begin{aligned} \phi_E^{(II)}(x) &= c \sqrt{\frac{l(x)}{2\pi}} \left[e^{2\pi \int_{x_1}^x \frac{dx'}{l(x')}} + \frac{i}{2} e^{-2\pi \int_{x_1}^x \frac{dx'}{l(x')}} \right] \\ &= c \sqrt{\frac{l(x)}{2\pi}} \left[e^{2\pi \int_{x_1}^{x_2} \frac{dx'}{l(x')}} e^{-2\pi \int_x^{x_2} \frac{dx'}{l(x')}} + \frac{i}{2} e^{-2\pi \int_{x_1}^{x_2} \frac{dx'}{l(x')}} e^{2\pi \int_x^{x_2} \frac{dx'}{l(x')}} \right] \\ &= c \sqrt{\frac{l(x)}{2\pi}} \left[e^{2\pi \int_{x_1}^{x_2} \frac{dx'}{l(x')}} e^{2\pi \int_x^{x_2} \frac{dx'}{l(x')}} + \frac{i}{2} e^{-2\pi \int_{x_1}^{x_2} \frac{dx'}{l(x')}} e^{-2\pi \int_x^{x_2} \frac{dx'}{l(x')}} \right], \quad x_1 \ll x \ll x_2. \end{aligned}$$

A primeira forma dessa expressão resulta diretamente das fórmulas de conexão (4.66). Ela foi em seguida transformada de modo que a dependência com x das integrais de fase tenha como ponto de referência o segundo ponto de inversão x_2 , o que faz com que as amplitudes apareçam multiplicadas por exponenciais das “integrais sob a barreira”

$$B_{\pm} \equiv \exp \left(\pm \int_{x_1}^{x_2} \frac{dx'}{l(x')} \right).$$

A última das três formas permite a utilização direta das fórmulas de conexão (4.64) e (4.65) para obter a forma da aproximação WKB para a função de onda na região III

$$\phi_E^{(\text{III})}(x) = c \sqrt{\frac{\lambda(x)}{2\pi}} \left[2B_+ \sin \left(2\pi \int_{x_2}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) + \frac{i}{2} B_- \cos \left(2\pi \int_{x_2}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \right], \quad x \gg x_2.$$

Essa função de onda pode ser re-escrita em termos de uma componente que representa partículas incidindo pela direita sobre a barreira de potencial

$$\phi_E^{(\text{III})(-)}(x) = ic \sqrt{\frac{\lambda(x)}{2\pi}} \left(B_+ + \frac{B_-}{4} \right) e^{-i \left[2\pi \int_{x_2}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right]}$$

além de uma componente análoga que contém o conjugado complexo do fator exponencial que aparece nesta última expressão, e que representa partículas refletidas pela barreira e se afastando dela

$$\phi_E^{(\text{III})(+)}(x) = ic \sqrt{\frac{\lambda(x)}{2\pi}} \left(-B_+ + \frac{B_-}{4} \right) e^{i \left[2\pi \int_{x_2}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right]}.$$

Nas condições em que a aproximação WKB é útil a integral sob a barreira que aparece no expoente de B_{\pm} é tipicamente muito maior que a unidade, de modo que $B_+ = 1/B_- \gg B_-$ e os termos envolvendo B_- nestas últimas expressões são desprezíveis frente aos que contém B_+ . Uma amplitude de transmissão da barreira de potencial à energia E pode então ser definida através da razão

$$T_E(x_I, x_{\text{III}}) \equiv \left| \frac{\phi_E^{(\text{I})}(x_I \ll x_1)}{\phi_E^{(\text{III})(-)}(x_{\text{III}} \gg x_2)} \right| = \frac{1}{B_+} \sqrt{\frac{\lambda(x_I \ll x_1)}{\lambda(x_{\text{III}} \gg x_2)}} = B_- \sqrt{\frac{\lambda(x_I \ll x_1)}{\lambda(x_{\text{III}} \gg x_2)}}.$$

Essa amplitude depende, em geral dos pontos x_I e x_{III} utilizados para defini-la, a menos que o potencial tenda a valores constantes V_I e V_{III} (independentes de x) nas regiões I e III. Escolhendo então x_I e x_{III} em tais regiões assintóticas, resulta que

$$T_E \rightarrow \left(\frac{E - V_{\text{III}}}{E - V_{\text{I}}} \right)^{1/4} B_-.$$

A dependência dessa amplitude com a energia E é fortemente dominada pelo fator B_- , devido à dependência exponencial desse fator com a integral sob a barreira. Uma afirmação análoga vale para a probabilidade de tunelamento, dada por

$$P_E = |T_E|^2 = \sqrt{\frac{E - V_{\text{III}}}{E - V_{\text{I}}}} B_-^2.$$

Quando a energia da partícula é da ordem da altura da barreira de potencial a ser transposta o tratamento acima deixa de ser válido, pois a distância entre os dois pontos de inversão diminui e conseqüentemente a aproximação WKB para a função de onda deixa de ser aplicável na região classicamente proibida.

Uma aproximação válida nesses casos foi descrita por D. H. Hill e J. A. Wheeler em 1953⁸. Ela consiste em aproximar a barreira de potencial por uma parábola ajustada ao seu topo e fazer a conexão das soluções correspondentes à aproximação WKB “antes” e “depois” da barreira às formas assintóticas da solução exata para a barreira parabólica. O resultado que decorre desse procedimento para a probabilidade de tunelamento é (fazendo o pré-fator relacionado com $\lambda(x_{\text{I}}$ e $\lambda(x_{\text{III}}$ igual a 1)

$$P_E = \left(1 + e^{2\pi \frac{E - E_{\text{max}}}{\hbar \omega_0}} \right)^{-1}$$

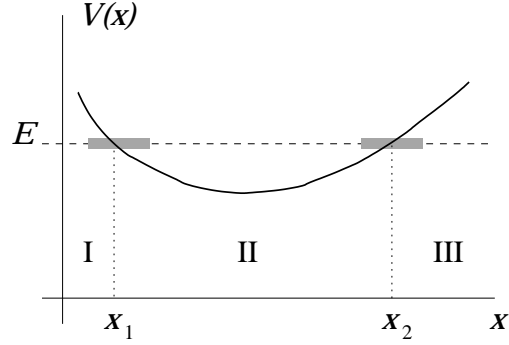
onde E_{max} e ω_0 são parâmetros que caracterizam o potencial parabólico ajustado ao topo da barreira, localizado em $x = x_0$,

$$V(x) \sim V_0(x) = E_{\text{max}} - \frac{M\omega_0^2}{2}(x - x_0)^2.$$

Essa expressão se aplica tanto para valores de E pouco abaixo como também para valores pouco *acima* da altura da barreira, E_{rmax} . Ela dá $P_E = 1/2$ para $E = E_{rmax}$, enquanto uma extrapolação direta do resultado anterior (com $V_{\text{I}} = V_{\text{III}}$) dá $P_E = 1$, nessa energia. Esse resultado mostra que existe uma probabilidade apreciável de reflexão mesmo para energias *pouco superiores* à altura da barreira.

⁸D. H. Hill e J. A. Wheeler, *Phys. Rev.* **89**, pag. 1140, figura 44. Este trabalho clássico de física nuclear tem o formato curioso de apresentar uma grande variedade de resultados em um número correspondente de figuras, com longas legendas.

Figura 4.8: Poço de potencial com dois pontos de inversão lineares x_1 e x_2 à energia E . A aproximação WKB é supostamente válida nas regiões I, II e III, excluídos os “intervalos perigosos” em torno dos pontos de inversão.



Uma segunda aplicação padrão das fórmulas de conexão é a obtenção de aproximações semi-clássicas para as energias de estados estacionários de uma partícula em um poço de potencial. Neste caso, a região classicamente permitida, no interior do poço de potencial, é limitada por dois pontos de inversão. Nas regiões classicamente proibidas, além desses pontos de inversão, a função de onda deve ter um comportamento decrescente, e essas condições de contorno em geral definem as energias em que a solução oscilatória válida na região classicamente permitida pode ser conectada a soluções com essa propriedade nas regiões proibidas. Uma situação desse tipo está representada na figura 4.8, em que a região II é classicamente permitida e as regiões I e III são classicamente proibidas à energia E , para a qual os pontos de inversão são x_1 e x_2 .

Supondo que na região I a função de onda seja puramente decrescente no sentido $x \ll x_1$, isto é

$$\phi_E^{(I)}(x) = \frac{c}{2} \sqrt{\frac{l(x)}{2\pi}} e^{2\pi \int_{x_1}^x \frac{dx'}{l(x')}}, \quad x \ll x_1,$$

então, sendo x_1 um ponto de inversão linear, a sua continuação na região II é dada pela fórmula de conexão (4.64) como

$$\phi_E^{(II)}(x) = c \sqrt{\frac{\lambda(x)}{2\pi}} \sin \left(2\pi \int_{x_1}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right), \quad x_1 \ll x \ll x_2. \quad (4.71)$$

Por outro lado, se a solução na região III é também decrescente, isto é, da forma

$$\phi_E^{(III)}(x) = \frac{c'}{2} \sqrt{\frac{l(x)}{2\pi}} e^{-2\pi \int_{x_2}^x \frac{dx'}{l(x')}}, \quad x \gg x_2,$$

sendo também x_2 um ponto de inversão linear, a sua continuação para a região permitida II é dada pela primeira das fórmulas de conexão (4.66) como

$$\phi_E^{(\text{II})}(x) = c' \sqrt{\frac{\lambda(x)}{2\pi}} \operatorname{sen} \left(2\pi \int_x^{x_2} \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right), \quad x_1 \ll x \ll x_2. \quad (4.72)$$

Para que exista um estado estacionário à energia E é preciso que as expressões (4.71) e 4.72) coincidam. Portanto os valores de E para os quais existem estados estacionários são tais que

$$\begin{aligned} c' \sqrt{\frac{\lambda(x)}{2\pi}} \operatorname{sen} \left(2\pi \int_x^{x_2} \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) &= c \sqrt{\frac{\lambda(x)}{2\pi}} \operatorname{sen} \left(2\pi \int_{x_1}^x \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \\ &= c \sqrt{\frac{\lambda(x)}{2\pi}} \operatorname{sen} \left[2\pi \int_{x_1}^{x_2} \frac{dx'}{\lambda(x')} + \frac{\pi}{2} - \left(2\pi \int_x^{x_2} \frac{dx'}{\lambda(x')} + \frac{\pi}{4} \right) \right] \end{aligned}$$

Essa condição estará satisfeita para valores de E tais que

$$\operatorname{sen} \left(2\pi \int_{x_1}^{x_2} \frac{dx'}{\lambda(x')} + \frac{\pi}{2} \right) = 0 \quad \text{e} \quad -\frac{c}{c'} \cos \left(2\pi \int_{x_1}^{x_2} \frac{dx'}{\lambda(x')} + \frac{\pi}{2} \right) = 1$$

ou seja

$$2\pi \int_{x_1}^{x_2} \frac{dx'}{\lambda(x')} = \left(N + \frac{1}{2} \right) \pi \quad \text{e} \quad \frac{c}{c'} = (-1)^N, \quad N \geq 0, \text{ inteiro.} \quad (4.73)$$

Lembrando que

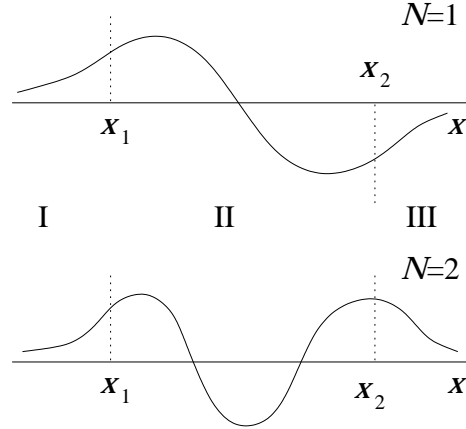
$$\frac{2\pi}{\lambda(x)} = \sqrt{\frac{2M[E - V(x)]}{\hbar^2}} = \frac{p(x)}{\hbar}$$

a primeira dessas duas condições pode ser escrita como

$$\int_{x_1}^{x_2} p(x') dx' = \left(N + \frac{1}{2} \right) \pi \hbar \quad \text{ou ainda como} \quad \oint p(x') dx' = \left(N + \frac{1}{2} \right) h$$

onde a integração se estende sobre um período *completo* do movimento clássico da partícula. Essa condição reproduz a condição de quantização usada por Bohr na velha teoria quântica (cf. eq. (1.4)), incluindo ainda a devida provisão para a energia de ponto zero através do termo $\hbar/2$. O inteiro N pode ser interpretado como o número de zeros da função de onda, e a fase $(-1)^N$ em (4.73) tem em conta o fato de que as caudas decrescentes da função de onda nas regiões classicamente proibidas têm fases opostas quando se ligam a soluções oscilatórias com um número ímpar de zeros na região classicamente permitida (v. figura 4.9).

Figura 4.9: Comportamento qualitativo das funções de onda para estados estacionários com $N = 1$ e $N = 2$ mostrando a diferença de fase nas regiões proibidas I e III nos casos em que N é ímpar.



Da mesma forma que no caso da penetrabilidade de uma barreira de potencial em energias próximas à do topo da barreira, o uso da aproximação WKB para o cálculo de energias de estados estacionários próximos ao mínimo do poço de potencial não é a rigor justificável, devido à diminuição da distância entre os dois pontos de inversão lineares e à conseqüente violação das condições de validade dessa aproximação agora na região classicamente permitida. É fácil verificar, no entanto, que uma aplicação direta da condição (4.73) a um potencial quadrático (oscilador harmônico) leva ao espectro quântico *exato* para esse potencial. Devido a esse fato, sempre que o potencial possa ser bem descrito por uma aproximação quadrática numa região suficientemente extensa em torno de seu mínimo, o espectro obtido através de (4.73) dará uma boa aproximação para o espectro exato também para valores pequenos de N . No entanto, a validade da aproximação WKB na região classicamente permitida a rigor só se justifica para valores grandes de N , o que está relacionado com o “princípio de correspondência” de Bohr.

Problemas.

4.1 - Expanda os autovetores exatos, dados na equação (3.9), para o sistema de dois níveis tratado no capítulo anterior, até segunda ordem em $|\gamma|/\Delta$. Existem termos lineares em $|\gamma|$? Procure entender, em termos do resultado da expansão dos autovetores, a ausência de termos lineares em $|\gamma|$ na expansão dos autovalores associados a esses autovetores.

4.2 - Estude o funcional

$$E[|\varphi\rangle] = \frac{\langle \varphi | H | \varphi \rangle}{\langle \varphi | \varphi \rangle}$$

cujos mínimos absolutos estão associados ao estado fundamental do sistema descrito pelo Hamiltoniano H , no caso de um oscilador harmônico unidimensional, isto é

$$H = \frac{p^2}{2M} + \frac{M\omega_0^2}{2} q^2,$$

utilizando uma família de “estados de prova” caracterizada pelas funções de onda

$$\varphi_a(x) = \begin{cases} \sqrt{\frac{15}{16a}} \left(1 - \frac{x^2}{a^2}\right), & -a \leq x \leq a \\ 0, & x < -a, \ x > a \end{cases}$$

onde a é um parâmetro real. **Sugestão:** verifique que as funções de prova tem norma 1 para qualquer valor de a , utilize a representação de Schrödinger de H e obtenha o valor de a que minimiza o funcional. Determine a partir desse valor o limite superior variacional para a energia do estado fundamental do oscilador e a função de onda da família considerada que melhor representa a autofunção correspondente.

4.3 - Com referência à equação (4.1), supondo que os autovalores de g sejam limitados superiormente por um *maior autovalor* γ_{\max} tal que para qualquer autovalor valha a relação $\gamma \leq \gamma_{\max}$, mostre que vale *também* a relação $G[|\varphi\rangle] \leq \gamma_{\max}$. Mostre em seguida que o funcional $E[\varphi]$ do problema anterior assume valores *arbitrariamente grandes* na família de funções de prova utilizada, e que isto garante que o espectro do Hamiltoniano do oscilador harmônico não tem um limite superior.

4.4 - Obtenha um limite superior variacional para a energia do estado fundamental do oscilador unidimensional perturbado

$$H = \frac{p^2}{2M} + \frac{M\omega_0^2}{2} q^2 - eE_0 q$$

utilizando vetores de prova dependentes do parâmetro variacional c que têm a forma

$$|\varphi(c)\rangle = (1 + c^2)^{-1/2} (|0\rangle + c|1\rangle)$$

onde $|0\rangle$ e $|1\rangle$ são o estado fundamental e o primeiro estado excitado do Hamiltoniano com $E_0 = 0$. Compare o resultado do cálculo variacional com o resultado perturbativo de primeira ordem para o vetor de estado, equação (4.10), e com o de segunda ordem para a energia, equação (4.9).

4.5 - O modelo mais simples para um átomo de hidrogênio pode ser caracterizado por meio do hamiltoniano

$$H = \frac{\vec{p}^2}{2m} - \frac{e^2}{r}$$

onde \vec{p} é o momento (tri-dimensional), m é a massa e $-e$ a carga do elétron. As componentes cartesianas de \vec{p} são canonicamente conjugadas às do vetor de posição do elétron \vec{r} , e $r \equiv |\vec{r}|$.

Obtenha o menor (e portanto o melhor) limite superior variacional para a energia do estado fundamental desse modelo que pode ser obtido usando uma função de prova gaussiana com simetria esférica, da forma

$$\varphi_b(\vec{r}) = N_b e^{-\frac{r^2}{2b^2}},$$

onde b (o parâmetro de largura da gaussiana) é tratado como um parâmetro a ajustar e N_b é uma constante de normalização, que depende do valor de b . Lembre-se de que tudo se passa em três dimensões, de modo que, por exemplo, a integral de normalização de $\varphi(\vec{r})$ é

$$\|\varphi_b\|^2 = |N_b|^2 \int d^3r e^{-\frac{r^2}{b^2}} = 4\pi |N_b|^2 \int_0^\infty dr r^2 e^{-\frac{r^2}{b^2}}$$

e assim por diante. **Sugestão:** o valor médio da energia cinética na função de prova gaussiana pode ser obtido como o valor médio da energia cinética do estado fundamental de um oscilador harmônico de frequência apropriada, a qual por sua vez é facilmente obtida da energia do estado fundamental do oscilador via teorema do virial.

No final compare o resultado variacional com o resultado exato para esse hamiltoniano

$$E_0 = - \left(\frac{e^2}{\hbar c} \right)^2 \frac{mc^2}{2}, \quad \frac{e^2}{\hbar c} \simeq \frac{1}{137}, \quad mc^2 \simeq 511 \text{ KeV}.$$

4.6 - a) Repita o cálculo perturbativo de primeira ordem feito na seção 4.1.2 para o oscilador bidimensional carregado, perturbado por um campo magnético externo perpendicular ao plano das oscilações, considerando porém o espaço degenerado bi-dimensional associado ao autovalor $2\hbar\omega_0$ de H_0 .

b) Trate em seguida o mesmo problema no caso de um oscilador tri-dimensional (isto é, $\vec{q} = \{q_1, q_2, q_3\}$, etc.) com o campo magnético externo na direção 3, de modo que a perturbação pode ainda ser escrita como $gl_3 = g(q_1p_2 - q_2p_1)$. Use como autovetores não perturbados produtos de tres autovetores de oscilador unidimensional, $|\nu_1, \nu_2, \nu_3\rangle$ para os quais os autovalores de H_0 são $\hbar\omega_0(\nu_1 + \nu_2 + \nu_3 + 3/2)$. Trate especificamente o autovalor triplamente degenerado $5\hbar\omega_0/2$.

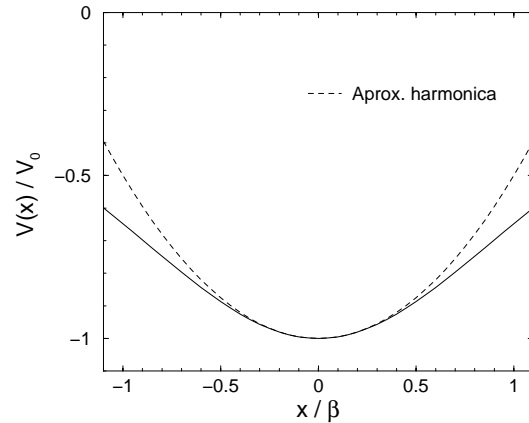
4.7 - Uma partícula de massa M está sujeita ao potencial atrativo anarmônico (em uma dimensão)

$$V(x) = -\frac{V_0}{\cosh \frac{x}{\beta}}, \quad V_0 > 0$$

que pode ser expandido em potências de x como

$$V(x) = -V_0 \left[1 - \frac{1}{2} \left(\frac{x}{\beta} \right)^2 + \frac{5}{24} \left(\frac{x}{\beta} \right)^4 - \dots \right]. \quad (4.74)$$

Figura 4.10: Potencial $V(x)$ e sua aproximação harmônica.



O estado fundamental desse sistema será bem descrito pela aproximação harmônica para o potencial, na qual as potências de x maiores que 2 são ignoradas na expansão (4.74), se $b \ll \beta$, sendo $b = \sqrt{\frac{\hbar}{M\omega_0}}$ o parâmetro do oscilador harmônico correspondente a essa aproximação.

a) Mostre que a condição $b \ll \beta$ se traduz, em termos dos parâmetros do problema, como

$$MV_0\beta^2 \gg \hbar^2$$

e escreva a energia do estado fundamental e do primeiro estado excitado na aproximação harmônica.

b) Calcule a correção perturbativa não nula de mais baixa ordem à energia desses dois estados devida ao termo anarmônico envolvendo x^4 . Exprima o resultado em termos dos parâmetros M , V_0 e β . **Sugestão:** use $x = \frac{b}{\sqrt{2}}(a + a^\dagger)$ nesse cálculo.

4.8 - Um oscilador harmônico unidimensional é perturbado por um campo de forças externo de forma que o Hamiltoniano é

$$H = \frac{p^2}{2M} + \frac{M\omega_0^2}{2}q^2 + \gamma \sin k_0 q$$

onde k_0 é uma constante com dimensões de inverso de comprimento. Obtenha o autovalor e o autovetor (ou a autofunção) para o estado fundamental, corrigidos até a ordem não nula mais baixa em γ . **Sugestão:** trabalhe na representação de Schrödinger e faça uso da função geratriz dos polinômios de Hermite para calcular os elementos de matriz necessários.

Observação: uma forma alternativa de obter o resultado desejado consiste em expandir $\sin k_0 q$ em série de potências e usar o fato de que $q = b(a^\dagger + a)/\sqrt{2}$ para calcular elementos de matriz entre estados não perturbados. Esse procedimento, no entanto, pode resultar bastante trabalhoso devido à necessidade de contabilizar devidamente as contribuições dos diferentes ordenamentos de a^\dagger e a em $(a^\dagger + a)^n$. Existe, no entanto, uma forma simples de fazer isso que consiste em exprimir primeiramente o seno em termos de exponenciais

$$\sin k_0 q = \frac{1}{2i} \left(e^{i\frac{k_0 b}{\sqrt{2}}(a^\dagger + a)} - e^{-i\frac{k_0 b}{\sqrt{2}}(a^\dagger + a)} \right)$$

e em seguida escrever cada uma das exponenciais como um produto de exponenciais envolvendo apenas a^\dagger e a . O fato de que esses dois operadores não comutam invalida o uso de uma relação simples do tipo $e^{A+B} = e^A e^B$. No entanto, a^\dagger e a *comutam com o seu comutador*, que é apenas um número, $[a, a^\dagger] = 1$. Nesse caso a separação pode ser feita utilizando a identidade

$$e^{A+B} = e^{-\frac{[A,B]}{2}} e^A e^B, \quad [A, [A, B]] = [B, [A, B]] = 0 \quad (4.75)$$

que dá, tomando $A = \pm i k_0 b a^\dagger / \sqrt{2}$ e $B = \pm i k_0 b a / \sqrt{2}$,

$$e^{\pm i\frac{k_0 b}{\sqrt{2}}(a^\dagger + a)} = e^{-\frac{k_0^2 b^2}{4}} e^{\pm i\frac{k_0 b}{\sqrt{2}}a^\dagger} e^{\pm i\frac{k_0 b}{\sqrt{2}}a}.$$

Essa identidade, de fato, faz o ordenamento necessário e leva a resultados especialmente simples no cálculo de elementos de matriz envolvendo o estado fundamental, pois

$$e^{\pm i\frac{k_0 b}{\sqrt{2}}a} |0\rangle = |0\rangle$$

devido ao fato de que $a|0\rangle = 0$.

A identidade (4.75) pode ser verificada calculando

$$\begin{aligned} \frac{d}{dx} e^{xA} e^{xB} &= A e^{xA} e^{xB} + e^{xA} B e^{xB} = A e^{xA} e^{xB} + [e^{xA}, B] e^{xB} + B e^{xA} e^{xB} \\ &= (A + B) e^{xA} e^{xB} + [e^{xA}, B] e^{xB}. \end{aligned}$$

O comutador que aparece no último termo pode ser expandido como

$$[e^{xA}, B] = \left[\sum_{n=0}^{\infty} \frac{x^n A^n}{n!}, B \right] = \sum_{n=1}^{\infty} \frac{x^n}{n!} [A^n, B]$$

que, usando o fato de que A e B comutam com o seu comutador, se reduz a

$$[[e^{xA}, B] = x[A, B] \sum_{n=0}^{\infty} \frac{x^n A^n}{n!} = x[A, B] e^{xA}.$$

O objeto $F(x) \equiv e^{xA} e^{xB}$, visto como uma função de x , satisfaz portanto a equação diferencial

$$\frac{dF}{dx} = (A + B)F(x) + [A, B]xF(x)$$

com a condição $F(0) = 1$. Isso leva à solução

$$F(x) \equiv e^{xA} e^{xB} = e^{x(A+B)} e^{\frac{x^2}{2}[A,B]}$$

a qual, para $x = 1$ leva à identidade (4.75).

4.9 - O potencial eletrostático $\Phi(r)$ de uma distribuição de carga esfericamente simétrica dada por $Ze\rho(r)$, com $\int d^3r \rho(r) = 1$, de modo que a carga total é Ze , satisfaz a equação de Poisson

$$\nabla^2 \Phi(r) \equiv \frac{d^2 \Phi(r)}{dr^2} + \frac{2}{r} \frac{d\Phi(r)}{dr} = -4\pi Ze\rho(r).$$

A energia de interação com uma partícula de carga $\pm e$ é portanto $V(r) = \pm e\Phi(r)$, de modo que o espalhamento dessa partícula por esse potencial calculado na aproximação de Born envolve a transformada de Fourier

$$\tilde{V}(\vec{k}) = \pm \frac{e}{(2\pi)^{3/2}} \int d^3r \Phi(r) e^{i\vec{k} \cdot \vec{r}} = \pm \sqrt{\frac{2}{\pi}} \frac{e}{k} \int_0^\infty dr \Phi(r) r \sin kr.$$

a) Tendo verificado esta última expressão, integre-a por partes duas vezes e use a equação de Poisson para exprimir o resultado diretamente em termos da função $\rho(r)$ que dá a distribuição de carga. Para eliminar termos integrados envolvendo $r\Phi(r)$ quando $r \rightarrow \infty$ introduza no integrando um “fator de blindagem” $e^{-\mu r}$ e faça $\mu \rightarrow 0$ no fim dos cálculos. Verifique desse modo que

$$\tilde{V}(\vec{k}) = \pm \sqrt{\frac{2}{\pi}} \frac{1}{k} \frac{4\pi Ze^2}{k^2} \int_0^\infty dr \rho(r) r \sin kr.$$

b) Use esse resultado para calcular a seção de choque de espalhamento por essa distribuição de carga e mostre que o resultado pode ser expresso sob a forma

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Ruth.}} \left| \frac{4\pi}{|\vec{k}_0 - k_0 \hat{k}|} \int_0^\infty dr \rho(r) r \sin(|\vec{k}_0 - k_0 \hat{k}| r) \right|^2$$

onde o primeiro fator no segundo membro é a seção de choque de Rutherford, correspondente ao espalhamento por uma carga puntiforme, e $|\vec{k}_0 - k_0 \hat{k}| = 2k_0 \sin \theta/2$, sendo θ o ângulo de espalhamento. Essa relação é utilizada para determinar distribuições de carga através da medida de seções de choque de espalhamento de elétrons por núcleos⁹.

4.10 - Um neutron (massa $M_n = 939.505$ MeV) decai com uma vida média $\tau \sim 15$ minutos para estados finais que contém um próton (massa $M_p = 938.211$ MeV), um elétron (massa $m_e = 0.511$ MeV) e um (anti-)neutrino (massa zero). Suponha que esse decaimento seja produzido por uma interação H_1 tal que

$$\langle \vec{k}_p, \vec{k}_e, \vec{k}_\nu | H_1 | \vec{k}_n \rangle = \frac{1}{L^3} \langle h_1 \rangle \delta_{\vec{k}_p + \vec{k}_e + \vec{k}_\nu, \vec{k}_n}^{(K)}$$

onde L^3 é o volume de quantização com condições de contorno periódicas, $\delta^{(K)}$ é um delta de Kronecker correspondente à conservação de momento e $\langle h_1 \rangle$ é uma constante com dimensões apropriadas.

a) Use a “regra áurea” de Fermi para escrever a probabilidade de decaimento por unidade de tempo W de um neutron com momento inicial nulo em termos de $\langle h_1 \rangle$, $\delta^{(K)}$ e de uma função delta de Dirac correspondente à conservação de energia.

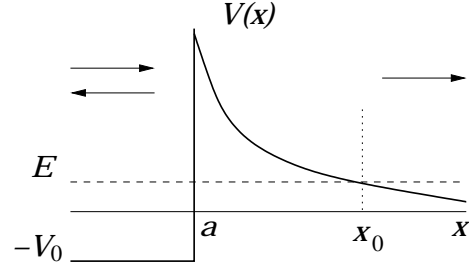
b) Como $M_n - M_p - m_e \ll M_p$, utilize a expressão não relativística $E_p \simeq M_p c^2 + \vec{p}_p^2/2M_p$ para a energia final do próton. A conservação de momento dá $|\vec{p}_p| = |\vec{p}_e + \vec{p}_\nu|$, e como $M_p \gg m_e$, segue que $\vec{p}_p^2/2M_p \ll E_e - m_e c^2$, $p_\nu c$. Desse modo é possível desprezar a energia cinética final do próton e fazer a soma sobre \vec{p}_p usando o delta de Kronecker da conservação de momento. Mostre que essa aproximação de *desprezar a energia de recuo do próton* leva a um fator de espaço de fases com a forma

$$\sum_{\vec{p}_e, \vec{p}_\nu} \delta \left((M_n - M_p) c^2 - \sqrt{m_e^2 c^4 + \vec{p}_e^2 c^2} - |\vec{p}_\nu| c \right)$$

c) Transforme as somas remanescentes em integrais e faça as integrações angulares e também a integração sobre $|\vec{p}_\nu|$. Verifique que desse modo a probabilidade de transição por unidade de tempo pode ser posta sob a forma

⁹v. L. R. B. Elton, *Nuclear Sizes*, Oxford University Press, 1961, pag. 10.

Figura 4.11: Potencial unidimensional que simula a barreira coulombiana a que está sujeita uma partícula alfa em um núcleo alfa-radioativo (problema 4.12). O interior do núcleo corresponde à região $x < a$.



$$W = \int_0^{p_e^{(\max)}} \frac{dW}{dp_e} dp_e$$

com $p_e = |\vec{p}_e|$, e identifique $p_e^{(\max)}$ em termos das massas das partículas, bem como a função dW/dp_e . Essa função dá a distribuição de energia dos elétrons resultantes do decaimento.

d) Para simplificar esta última integral sobre p_e , use a aproximação não relativística $E_e \simeq m_e c^2 + p_e^2/2m_e$ para a energia do elétron. Calcule a integral e, usando $W \sim 1/\tau$, obtenha um valor para $\langle h_1 \rangle$. Escrevendo $\langle h_1 \rangle = \text{Fatores dimensionais} \times g^2$ obtenha o valor de g^2 escrevendo os fatores dimensionais em termos de \hbar , c e $m_\pi = 139.6 \text{ MeV}$.

4.11 - A função de onda na aproximação WKB para um problema unidimensional é escrita como

$$\phi_E(x) = \frac{1}{\sqrt{p(x)}} e^{\frac{i}{\hbar} \int p(x) dx} e^{i\hbar S_2(x)}, \quad p(x) \equiv \sqrt{2M[E - V(x)]}.$$

Mostre que $\hbar S_2(x) \ll 1$ se

$$\frac{\hbar}{\sqrt{2M}} \frac{1}{[p(x)]^3} \left| \frac{dV}{dx} \right| \ll 1.$$

4.12 - Use a aproximação WKB para calcular a probabilidade de tunelamento à energia $E \ll \alpha/a$ do potencial unidimensional

$$V(x) = \begin{cases} -V_0 & \text{se } x < a \\ \frac{\alpha}{x} & \text{se } x \geq a \end{cases} \quad a > 0,$$

por partículas de massa M (v. fig. 4.11). Esse potencial simula a barreira coulombiana a que está sujeita uma partícula alfa em um núcleo alfa-radioativo.

Suponha que a energia seja suficientemente baixa para que a aproximação WKB para a função de onda seja válida na região classicamente proibida $a < x < x_0$ e na região classicamente permitida $x > x_0$, excluindo um “intervalo perigoso” no entorno do ponto de inversão linear x_0 . **Sugestão:** use condições de contorno de partículas se deslocando para a direita na região $x > x_0$, e ondas planas incidente e refletida na região $x < a$, conforme indicado na figura. Use a continuidade da função de onda e de sua derivada em $x = a$ e as fórmulas de conexão apropriadas no ponto de inversão $x = x_0$. Note que $V(x) \rightarrow 0$ para $x \gg x_0$.

4.13 - Mostre que a aplicação das relações (4.73) a um oscilador harmônico unidimensional de frequência ω_0 leva ao espectro quântico *exato* do oscilador.

Referências gerais.

A expansão perturbativa de Rayleigh-Schrödinger é apresentada e discutida em praticamente qualquer livro sobre a Mecânica Quântica, bem como o tratamento de níveis degenerados, e em particular nas referências padrão

★ **A. Messiah**, *op. cit.*, Capítulo XVI, partes I e II.

★ **C. Cohen-Tannoudji, B. Diu e F. Laloë**, *op. cit.*, Vol. II, Capítulo XI.

★ **J. J. Sakurai**, *op. cit.*

A expansão de Brillouin-Wigner é raramente apresentada ou mencionada. Ela não aparece em nenhum dos textos citados acima. Uma excessão é

★ **P. Roman**, *Advanced Quantum Theory*, Addison-Wesley, 1965, páginas 491-492 (exercícios!).

A teoria de perturbações dependente do tempo aparece também, essencialmente na forma apresentada aqui, em qualquer dos textos acima. Uma discussão pormenorizada dos termos de segunda ordem na Regra Áurea é dada por

★ **R. P. Feynman e A. R. Hibbs**, *Quantum Mechanics and Path Integrals*, Mac-Graw-Hill, 1965, páginas 154-158.

O tratamento variacional do estado fundamental é tratado também nos textos padrão citados acima. O tratamento apresentadp da aproximação WKB é essencialmente o que se encontra em

★ **A. Messiah**, *op. cit.*, Capítulo VI.

Um tratamento mais pormenorizado das fórmulas de conexão se encontra em

★ **L. I. Schiff**, *Quantum Mechanics* (2nd. Edition), MacGraw-Hill, 1955, §28.

★ **L. D. Landau e E. M. Lifshitz**, *op. cit.*, Capítulo VII.

A equação e as funções de Airy são descritas, por exemplo, em

★ **M. Abramowitz e I. A. Stegun**, *Handbook of Mathematical Functions*, Dover, 1965, página 446.

Capítulo 5

Momento angular, potenciais centrais.

Na mecânica clássica, o momento angular, visto como variável dinâmica de um sistema Hamiltoniano, gera transformações canônicas infinitesimais que podem ser identificadas com rotações infinitesimais do sistema. Como consequência disso, a invariância rotacional da hamiltoniana que descreve o sistema leva à conservação do seu momento angular total, isto é, acarreta o fato de que o momento angular total é uma constante do movimento. Esses resultados têm uma contrapartida exata na mecânica quântica, como pode ser facilmente verificado no caso particular de um sistema descrito por um operador Hamiltoniano da forma $H(\vec{p}, \vec{q})$ onde \vec{q} e \vec{p} são operadores correspondentes à posição e ao momento no espaço usual de três dimensões, cujas componentes cartesianas satisfazem as relações de comutação canônicas

$$[q_j, p_k] = i\hbar\delta_{jk}, \quad j, k = 1, 2, 3.$$

Para verificar isso é conveniente utilizar a representação de Schrödinger, na qual o vetor de estado que descreve o sistema é dado como uma função de onda de quadrado integrável $\varphi(\vec{r})$ definida no espaço tri-dimensional de posições. Uma rotação de um ângulo infinitesimal $d\theta$ em torno de um eixo dado pelo versor \hat{u} leva o vetor de posição \vec{r} a um vetor \vec{r}' , e a relação entre esses dois vetores pode ser expressa por

$$\vec{r}' = \vec{r} + d\theta \hat{u} \times \vec{r} + \mathcal{O}(d\theta^2) \quad \text{ou} \quad \vec{r} = \vec{r}' - d\theta \hat{u} \times \vec{r}' + \mathcal{O}(d\theta^2)$$

onde o sinal \times indica o produto vetorial. A relação entre a função de onda φ' que descreve o sistema rodado e a função de onda original φ pode ser determinada notando que φ' deve ter no ponto \vec{r}' o valor que φ tem no ponto \vec{r} , isto é

$$\varphi'(\vec{r}') = \varphi(\vec{r}) = \varphi(\vec{r}' - d\theta \hat{u} \times \vec{r}') = \varphi(\vec{r}') - d\theta \hat{u} \times \vec{r}' \cdot \nabla \varphi(\vec{r}') + \mathcal{O}(d\theta^2).$$

O operador diferencial que aparece no segundo termo pode ser re-escrito como

$$d\theta \hat{u} \times \vec{r}' \cdot \nabla = \frac{i}{\hbar} d\theta \hat{u} \cdot \vec{r}' \times \frac{\hbar}{i} \nabla = \frac{i}{\hbar} d\theta \hat{u} \cdot \vec{r}' \times \vec{p} = \frac{i}{\hbar} d\theta \hat{u} \cdot \vec{L}$$

onde $\vec{L} \equiv \vec{r}' \times \vec{p}$ é o momento angular do sistema. Como \vec{r}' pode ser tomado como um valor genérico do argumento da funções de onda, resulta portanto que

$$\varphi'(\vec{r}) = \left(1 - \frac{i}{\hbar} d\theta \hat{u} \cdot \vec{L}\right) \varphi(\vec{r}). \quad (5.1)$$

Essa relação mostra que a variação da função de onda sob uma rotação infinitesimal do sistema físico considerado é proporcional à ação de $\hat{u} \cdot \vec{L}/\hbar$ sobre a função de onda, e por isso esse operador é chamado *gerador* de rotações em torno do eixo \hat{u} . O efeito sobre a função de onda de uma rotação de um ângulo finito θ em torno de \hat{u} , por outro lado, pode ser obtido como o limite, para $N \rightarrow \infty$, de uma sucessão de N rotações de ângulos θ/N cada uma das quais pode ser tratada como uma rotação infinitesimal, de modo que a transformação (5.1) se generaliza nesse caso para

$$\varphi'(\vec{r}) = \lim_{N \rightarrow \infty} \left(1 - \frac{i}{\hbar} \frac{\theta}{N} \hat{u} \cdot \vec{L}\right)^N \varphi(\vec{r}) = e^{-\frac{i}{\hbar} \theta \hat{u} \cdot \vec{L}} \varphi(\vec{r}) \equiv R_{\hat{u}}(\theta) \varphi(\vec{r}). \quad (5.2)$$

O operador $R_{\hat{u}}(\theta)$, definido na equação (5.2) e que tem a forma de uma exponencial do gerador de rotações em torno de \hat{u} , descreve a mudança da função de onda do sistema quando ele é rodado de um ângulo *finito* θ em torno do eixo \hat{u} . É fácil verificar explicitamente a hermiticidade de qualquer componente do momento angular \vec{L} para funções de onda de quadrado integrável, do que decorre, em particular, que este operador de rotação é um operador *unitário*. O efeito de uma rotação sobre uma variável dinâmica g do sistema considerado pode ser obtido notando que

$$(g\varphi)' = R_{\hat{u}}(\theta)g\varphi = R_{\hat{u}}(\theta)gR_{\hat{u}}^{-1}(\theta)R_{\hat{u}}(\theta)\varphi = R_{\hat{u}}(\theta)gR_{\hat{u}}^{-1}(\theta)\varphi' \equiv g'\varphi'$$

que identifica a variável dinâmica para o sistema rodado como $g' = R_{\hat{u}}(\theta)gR_{\hat{u}}^{-1}(\theta)$. No caso em que a variável dinâmica considerada é o Hamiltoniano do sistema, resulta que o Hamiltoniano que descreve o sistema rodado é

$$H' = e^{-\frac{i}{\hbar} \theta \hat{u} \cdot \vec{L}} H e^{\frac{i}{\hbar} \theta \hat{u} \cdot \vec{L}}.$$

O Hamiltoniano é *invariante* sob rotações se $H' = H$ para qualquer eixo de rotação \hat{u} e para qualquer ângulo θ , e nesse caso segue que $[H, \hat{u} \cdot \vec{L}] = 0$ para qualquer eixo de rotação \hat{u} , isto

é, a projeção do momento angular sobre qualquer eixo é uma constante do movimento, o que é usualmente expresso escrevendo simplesmente $[H, \vec{L}] = 0$.

É fácil verificar diretamente, usando as relações canônicas de comutação, que a energia cinética de uma partícula de massa M no espaço de três dimensões, $K = p^2/2M$, é invariante sob rotações, isto é $[K, \vec{L}] = 0$. Isso decorre na realidade de que K é um operador *escalar*, e como tal invariante por rotações espaciais. O mesmo resultado permanece válido se essa partícula está sujeita a um potencial externo central $V(|\vec{q}|)$, e da mesma forma que na mecânica clássica, o uso do momento angular como constante de movimento é extremamente útil no tratamento quântico desse problema.

As componentes cartesianas L_j , $j = 1, 2, 3$ do momento angular $\vec{L} = \vec{q} \times \vec{p}$ podem ser escritas explicitamente como

$$L_1 = q_2 p_3 - p_2 q_3; \quad L_2 = q_3 p_1 - p_3 q_1 \quad \text{e} \quad L_3 = q_1 p_2 - p_1 q_2$$

ou, em uma forma geral mais compacta, como

$$L_j = \sum_{k,l=1}^3 \epsilon_{jkl} q_k p_l$$

onde ϵ_{jkl} é o “símbolo antissimétrico” definido como sendo $+1$ ou -1 quando os índices j, k, l são respectivamente uma permutação par ou ímpar de $1, 2, 3$, e zero se dois ou mais dos três índices forem iguais. Usando as relações de comutação canônicas, é fácil verificar então que essas componentes de \vec{L} por sua vez satisfazem as relações de comutação

$$[L_1, L_2] = i\hbar L_3; \quad [L_2, L_3] = i\hbar L_1 \quad \text{e} \quad [L_3, L_1] = i\hbar L_2$$

ou equivalentemente, em termos do símbolo antissimétrico ϵ_{jkl} ,

$$[L_j, L_k] = i\hbar \sum_l \epsilon_{jkl} L_l. \tag{5.3}$$

Isso mostra que diferentes componentes cartesianas de \vec{L} *não podem ser diagonalizadas simultaneamente* (não são mutuamente compatíveis), e que apenas uma das três componentes pode ser simultaneamente diagonalizada com um hamiltoniano rotacionalmente invariante. Por outro lado, usando as relações de comutação (5.3) é fácil verificar que o operador

$$L^2 \equiv \vec{L} \cdot \vec{L} = L_1^2 + L_2^2 + L_3^2$$

comuta com $\hat{u} \cdot \vec{L}$ para qualquer \hat{u} (e portanto, em particular, com qualquer das três componentes cartesianas L_j), o que novamente decorre do fato de que L^2 é também um operador

escalar, e portanto invariante por rotações espaciais. Desse modo, é possível obter autovetores simultâneos de L^2 e de uma das componentes cartesianas de \vec{L} , usualmente designada como L_3 . Um Hamiltoniano rotacionalmente invariante pode ser também diagonalizado simultaneamente com L^2 , pois de $[H, \vec{L}] = 0$ decorre que $[H, L^2] = 0$, e com L_3 .

Vale a pena lembrar que todos esses fatos têm também uma contrapartida clássica bastante direta. De fato, os parênteses de Poisson das componentes cartesianas do momento angular clássico são

$$\{L_j, L_k\} = \sum_l \epsilon_{jkl} L_l$$

dos quais segue ainda que $\{L^2, L_j\} = 0$. Como os parênteses de Poisson de momentos e coordenadas canônicos devem ser

$$\{q_j, p_k\} = \delta_{jk}, \quad \{q_j, q_k\} = \{p_j, p_k\} = 0$$

resulta que não é possível atribuir a mais de uma das componentes cartesianas do momento angular o papel de momento canônico, embora L^2 e uma das componentes de \vec{L} , por exemplo L_3 , possam desempenhar simultaneamente esse papel.

5.1 Momento angular na dinâmica de uma partícula sujeita a um potencial central.

O problema de autovalores para o Hamiltoniano que descreve uma partícula sujeita a um potencial central corresponde, na representação de Schrödinger, à equação de derivadas parciais

$$-\frac{\hbar^2}{2M} \nabla^2 \phi_E(\vec{r}) + V(r) \phi_E(\vec{r}) = E \phi_E(\vec{r}) \quad (5.4)$$

na qual $\phi_E(\vec{r}) = \langle \vec{r} | \phi_E \rangle$ são autofunções, associadas a autovalores E , que devem satisfazer condições de contorno apropriadas. No caso de autovetores pertencentes ao espectro discreto, as autofunções devem ser de quadrado integrável, o que exige em particular um anulamento suficientemente rápido para $r \rightarrow \infty$, enquanto no caso do espectro contínuo ocorrerão autofunções “impróprias”, de norma infinita, mas “normalizáveis” em termos de funções delta de Dirac (v., por exemplo, a equação (3.26)). A invariância rotacional do Hamiltoniano permite também, nesse caso, que as autofunções $\phi_E(\vec{r})$ sejam escolhidas de forma a serem também autofunções das variáveis dinâmicas compatíveis L^2 e L_3 . As autofunções simultâneas de H e desse dois operadores podem então ser designadas por $\phi_{E\lambda m}(\vec{r})$, onde λ e m se referem aos autovalores de L^2 e de L_3 respectivamente, isto é

$$L^2\phi_{E\lambda m}(\vec{r}) = \hbar^2\lambda\phi_{E\lambda m}(\vec{r}) \quad \text{e} \quad L_3\phi_{E\lambda m}(\vec{r}) = \hbar m\phi_{E\lambda m}(\vec{r}). \quad (5.5)$$

Os fatores \hbar^2 e \hbar foram introduzidos por conveniência, de forma a tornar λ e m adimensionais.

O problema simultâneo de autovalores, envolvendo as equações (5.4) e (5.5), pode ser tratado de forma conveniente explicitando o momento angular na primeira dessas equações. Para isso, um cálculo algo extenso mas sem maiores dificuldades de $L^2 = (-i\hbar\vec{r} \times \nabla)^2$ em coordenadas esféricas $\vec{r} \equiv \{r, \theta, \varphi\}$ leva ao resultado

$$L^2 \equiv -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right]$$

que pode ser reconhecido como sendo proporcional à parte angular do laplaciano expresso também nessas coordenadas, que pode ser escrito como

$$\nabla^2 \equiv \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{L^2}{\hbar^2 r^2} \right].$$

O operador L_3 , por outro lado, é representado simplesmente por

$$L_3 = \frac{\hbar}{i} \frac{\partial}{\partial\varphi}. \quad (5.6)$$

A equação (5.4), portanto, depende das coordenadas angulares apenas através de L^2 . Na realidade, uma dependência explícita de L_3 poderia ter sido excluída “a priori” simplesmente pela invariança rotacional da energia cinética. Devido à forma assumida por essa equação, que permite a separação da variável radial r e das variáveis angulares θ e φ , as autofunções simultâneas de H , L^2 e L_3 podem ser escritas sob a forma de produto

$$\phi_{E\lambda m}(\vec{r}) = R_{E\lambda}(r)Y_{\lambda m}(\theta, \varphi)$$

onde os $Y_{\lambda m}(\theta, \varphi)$ satisfazem

$$L^2Y_{\lambda m}(\theta, \varphi) = \hbar^2\lambda Y_{\lambda m}(\theta, \varphi) \quad \text{e} \quad L_3Y_{\lambda m}(\theta, \varphi) = \hbar mY_{\lambda m}(\theta, \varphi) \quad (5.7)$$

isto é, são autofunções, dependentes apenas das variáveis angulares, de L^2 e L_3 . A equação a ser satisfeita pela função radial $R_{E\lambda}(r)$ é, por outro lado,

$$-\frac{\hbar^2}{2M} \left[\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) - \frac{\lambda}{r^2} \right] R_{E\lambda}(r) + V(r)R_{E\lambda}(r) = ER_{E\lambda}(r). \quad (5.8)$$

Toda a dependência do potencial $V(r)$ está contida nessa equação, de modo que as autofunções angulares podem ser determinadas uma vez por todas, e são aplicáveis a *qualquer* potencial central. Deve ser notado também que a função radial depende do momento angular através do autovalor λ de L^2 . Antes de tratar da solução de (5.8) para diferentes potenciais centrais é preciso, portanto, estudar o problema de autovalores (5.7).

5.1.1 Autovalores e autovetores do momento angular.

Autovalores e autofunções simultâneas de L^2 e L_3 podem ser obtidos explicitamente tratando as equações (5.7) como equações diferenciais nas variáveis angulares θ e φ . O tratamento alternativo apresentado a seguir, de natureza algébrica e baseado apenas nas relações de comutação (5.3), é na realidade mais geral, no sentido de não depende da forma diferencial explícita pela qual os operadores associados ao momento angular $\vec{L} = \vec{r} \times \vec{p}$ são realizados na representação de Schrödinger.

Definindo os operadores não hermiteanos

$$L_{\pm} \equiv L_1 \pm iL_2$$

é imediato obter, a partir de (5.3), as relações de comutação

$$[L_3, L_{\pm}] = \pm \hbar L_{\pm} \quad \text{e} \quad [L_+, L_-] = 2\hbar L_3.$$

Supondo então que $|\lambda m\rangle$ seja um autovetor simultâneo de L^2 e L_3 , isto é,

$$L^2|\lambda m\rangle = \hbar^2 \lambda |\lambda m\rangle \quad \text{e} \quad L_3|\lambda m\rangle = \hbar m |\lambda m\rangle$$

é preciso notar, em primeiro lugar, que λ deve ser *não negativo*, como autovalor de um operador que é uma soma de quadrados de operadores hermiteanos. Além disso, as relações de comutação envolvendo L_{\pm} dão

$$L_3 L_{\pm} |\lambda m\rangle = [L_3, L_{\pm}] |\lambda m\rangle + L_{\pm} L_3 |\lambda m\rangle = \hbar(m \pm 1) L_{\pm} |\lambda m\rangle, \quad (5.9)$$

que mostra que os vetores $L_{\pm} |\lambda m\rangle$ são respectivamente proporcionais a $|\lambda m \pm 1\rangle$. As constantes de proporcionalidade (a menos de um fator de fase) e, como se verá, também a natureza dos autovalores podem ser determinados pela condição de normalização $\langle \lambda m | \lambda m' \rangle = \delta_{mm'}$ juntamente com a condição de positividade da norma. De fato, o quadrado da norma de $L_{\pm} |\lambda m\rangle$, tendo em conta o fato de que $L_+^{\dagger} = L_-$, pode ser calculado usando as definições e as relações de comutação (5.3) como

$$\begin{aligned}
\langle \lambda m | L_{\mp} L_{\pm} | \lambda m \rangle &= \langle \lambda m | L_1^2 + L_2^2 \pm i[L_1, L_2] | \lambda m \rangle \\
&= \langle \lambda m | L^2 - L_3(L_3 \pm \hbar) | \lambda m \rangle = \hbar^2[\lambda - m(m \pm 1)].
\end{aligned} \tag{5.10}$$

Como quadrado da norma de um vetor, essa quantidade deve ser positiva definida. No entanto, a relação (5.9) mostra que a ação do operador L_+ aumenta o valor de m de uma unidade, de modo que $\lambda - m(m + 1)$ se tornará finalmente negativo *a menos que* o autovalor λ seja por sua vez da forma $l(l + 1)$ e que $l \geq 0$ seja um valor possível de m . Nessas condições, de fato, $L_+|\lambda = l(l + 1) \ m = l\rangle$ será o vetor nulo, levando à terminação da seqüência de valores crescentes de m . De modo análogo, a seqüência de valores decrescentes de m gerada pela ação de L_- leva a valores negativos da norma a menos que $-l \leq 0$ também seja um valor possível de m , pois nesse caso $L_-|\lambda = l(l + 1) \ m = l\rangle$ é o vetor nulo que limita inferiormente a seqüência de valores de m . Por outro lado, a passagem de $m = l$ a $m = -l$ se faz através de um número *inteiro* de passos unitários, de modo que os valores possíveis de m são $l, l - 1, l - 2, \dots, l - 2l$. O número de passos é então claramente $2l$, de onde se conclui que $l \geq 0$ deve ser um número *inteiro* (para $2l$ par) ou *semi-inteiro* (para $2l$ ímpar). Com esse resultado é conveniente mudar a notação usada até aqui para os autovetores pondo

$$|\lambda = l(l + 1) \ m\rangle \longrightarrow |lm\rangle$$

com o que a (5.10) dá

$$\langle lm | L_{\mp} L_{\pm} | lm \rangle = \hbar^2[l(l + 1) - m(m \pm 1)]$$

que, com a condição de normalização proposta para os $|lm\rangle$, leva finalmente às relações

$$L_{\pm}|lm\rangle = \hbar\sqrt{l(l + 1) - m(m \pm 1)} |lm \pm 1\rangle. \tag{5.11}$$

A definição explícita dos operadores de momento angular em termos de operadores diferenciais no espaço de configurações $\{r, \theta, \varphi\}$ leva a uma restrição adicional sobre os possíveis valores de l . Uma forma simples de ver isso é notando que as autofunções de L_3 , quando este operador tem a forma (5.6), são exponenciais da forma $e^{im\varphi}$. Como o ângulo φ associado a uma determinada posição é definido apenas a menos de um múltiplo de 2π , deve-se ter $e^{im\varphi} = e^{im(\varphi + 2n\pi)}$ para que essas autofunções sejam univocamente definidas no espaço de configurações, o que *exclui* valores semi-inteiros de m (e portanto também de l), pois nesse caso $2mn$ é ímpar, levando a uma troca de sinal para um acréscimo de 2π no ângulo φ . Realizações alternativas do momento angular, que permitem a ocorrência de valores semi-inteiros de l (em particular o caso $l = 1/2$), são essenciais para a descrição de fenômenos como a estrutura fina e o efeito Zeeman no átomo de hidrogênio, e serão discutidas mais adiante, nesse contexto.

Em termos de operadores diferenciais em coordenadas esféricas, não é difícil verificar que os operadores L_{\pm} são dados por

$$L_{\pm} = e^{\pm i\varphi} \left(\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right). \quad (5.12)$$

Usando as relações (5.11) esses operadores permitem determinar as autofunções simultâneas $Y_{lm}(\theta, \varphi) \equiv Y_{\lambda=l(l+1) m}(\theta, \varphi)$ com $m \neq 0$ a partir de Y_{l0} . Estas últimas são as soluções independentes de φ de

$$L^2 Y_{l0} = \hbar^2 l(l+1) Y_{l0}$$

que são proporcionais aos polinômios de Legendre $P_l(\cos \theta)$, que podem ser calculados a partir da fórmula de Rodrigues

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l.$$

Com a condição de normalização

$$\int_{-1}^1 d(\cos \theta) \int_0^{2\pi} d\varphi |Y_{l0}|^2 = 1$$

e com a escolha mais usual de fases, essas funções são dadas por

$$Y_{l0}(\theta) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos \theta).$$

A escolha usual de fases para as *harmônicas esféricas* $Y_{lm}(\theta, \varphi)$ corresponde a obter essas funções por aplicações sucessivas de (5.11) a partir de Y_{l0} assim definida, usando a forma (5.12) dos operadores L_{\pm} .

5.1.2 Partícula livre com momento angular bem definido.

O caso mais simples em que se pode obter autofunções simultâneas do hamiltoniano e do momento angular é o da partícula livre, em três dimensões espaciais. Nesse caso $V(r) \equiv 0$, e a equação radial (5.8) correspondente ao valor l do momento angular se reduz a

$$\left[\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) - \frac{l(l+1)}{r^2} + \frac{2ME}{\hbar^2} \right] R_{El}(r) = 0.$$

O aparecimento de um termo envolvendo a primeira derivada com relação a r pode ser evitado escrevendo $R_{El}(r)$ sob a forma

$$R_{El}(r) = \frac{u_{El}(r)}{r}$$

com a condição de que $u_{El}(r \rightarrow 0) \rightarrow 0$ de forma a evitar que R_{El} seja singular na origem. Dessa forma a equação que resulta para essa nova função radial é

$$\frac{d^2 u_{El}}{dr^2} + \left(k^2 - \frac{l(l+1)}{r^2} \right) u_{El} = 0$$

onde foi introduzida a notação $k^2 = 2ME/\hbar^2$.

Essa equação pode ser convenientemente escrita em termos da variável adimensional $\rho \equiv kr$ como

$$\left(\frac{l(l+1)}{\rho^2} - \frac{d^2}{d\rho^2} \right) u_l \equiv g_l u_l = u_l \quad (5.13)$$

onde, por simplicidade e sem risco de confusão, o índice E da função radial u_{El} foi omitido. As soluções dessa equação podem ser obtidas convenientemente através do uso de duas fatorizações possíveis do operador g_l , que aparece no primeiro membro dessa equação. Essas fatorizações são dadas por

$$g_l \equiv \left(\frac{l(l+1)}{\rho^2} - \frac{d^2}{d\rho^2} \right) = \left(\frac{l}{\rho} - \frac{d}{d\rho} \right) \left(\frac{l}{\rho} + \frac{d}{d\rho} \right) = \left(\frac{l+1}{\rho} + \frac{d}{d\rho} \right) \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right)$$

e o seu uso permite mostrar que

$$\begin{aligned} g_{l+1} \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) u_l &= \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) \left(\frac{l+1}{\rho} + \frac{d}{d\rho} \right) \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) u_l \\ &= \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) g_l u_l \\ &= \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) u_l \end{aligned} \quad (5.14)$$

donde resulta, tendo em vista a equação (5.13),

$$\left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) u_l \propto u_{l+1}. \quad (5.15)$$

Na primeira passagem de (5.14) foi usada a primeira forma da fatorização para reescrever g_{l+1} , enquanto que, na segunda passagem, foi usada a segunda forma para reconstruir g_l . A relação (5.15) permite portanto obter soluções com valores maiores de l a partir de qualquer solução u_l que seja conhecida. Para completar a solução de (5.13) para qualquer valor de l basta então, tendo em vista o caráter inteiro e não negativo de l , resolver essa equação para $l = 0$. Nesse caso ela se reduz simplesmente a

$$\frac{d^2 u_0}{d\rho^2} + u_0 = 0$$

cujas soluções gerais podem ser escritas como uma combinação linear de

$$u_0^{(1)}(\rho) = \sin \rho \equiv \rho j_0(\rho) \quad \text{e} \quad u_0^{(2)}(\rho) = \cos \rho \equiv \rho n_0(\rho).$$

As funções $j_0(\rho)$ e $n_0(\rho)$ definidas dessa forma são chamadas respectivamente *funções de Bessel e de Neumann esféricas* de ordem zero. No caso da partícula livre com momento angular zero, a função de Neumann não satisfaz a condição de regularidade na origem, e deve portanto ser excluída das soluções admissíveis, de modo que o fator radial $R_{E0}(r)$ da solução é, a menos de uma constante multiplicativa fixada por uma condição de normalização, dado pela função de Bessel esférica $j_0(kr)$, com $k^2 = 2ME/\hbar^2$.

As soluções para $l > 0$ podem agora ser obtidas a partir de j_0 e de n_0 , usando a relação (5.15). Funções de Bessel e de Neumann de ordem $l > 0$ podem ser definidas tomando a constante de proporcionalidade implícita nessa relação como sendo igual a 1, de modo que

$$\left\{ \begin{array}{c} \rho j_{l+1}(\rho) \\ \rho n_{l+1}(\rho) \end{array} \right\} = \left(\frac{l+1}{\rho} - \frac{d}{d\rho} \right) \left\{ \begin{array}{c} \rho j_l(\rho) \\ \rho n_l(\rho) \end{array} \right\} \quad (5.16)$$

ou ainda

$$\left\{ \begin{array}{c} \rho j_l(\rho) \\ \rho n_l(\rho) \end{array} \right\} = \left(\frac{l}{\rho} - \frac{d}{d\rho} \right) \left(\frac{l-1}{\rho} - \frac{d}{d\rho} \right) \dots \left(\frac{1}{\rho} - \frac{d}{d\rho} \right) \left\{ \begin{array}{c} \sin \rho \\ \cos \rho \end{array} \right\}.$$

O comportamento dessas funções para $\rho \rightarrow 0$ e para $\rho \rightarrow \infty$ é dado por¹

$$\begin{aligned} \rho \rightarrow 0 : \quad j_l(\rho) &\rightarrow \frac{\rho^l}{(2l+1)!!} \left(1 + \mathcal{O}(\rho^2) \right), \quad n_l(\rho) \rightarrow \frac{(2l-1)!!}{\rho^{l+1}} \left(1 + \mathcal{O}(\rho^2) \right); \\ \rho \rightarrow \infty : \quad j_l(\rho) &\rightarrow \frac{1}{\rho} \sin \left(\rho - \frac{l\pi}{2} \right), \quad n_l(\rho) \rightarrow \frac{1}{\rho} \cos \left(\rho - \frac{l\pi}{2} \right). \end{aligned} \quad (5.17)$$

¹v. e.g. M. Abramowitz e I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1965, Seção 10.

A singularidade das funções de Neumann na origem mostra que essas funções devem ser excluídas, para qualquer l , como possíveis soluções para as funções radiais da partícula livre com momento angular bem definido. A menos de eventuais fatores de normalização, as autofunções simultâneas de H , L^2 e L_3 são portanto dadas, para esse caso, por

$$\phi_{Elm}(\vec{r}) \propto j_l(kr) Y_{lm}(\theta, \varphi)$$

com $k^2 = 2ME/\hbar^2$, $\vec{r} \equiv \{r, \theta, \varphi\}$. O comportamento de $j_l(kr)$ para r grande mostra que as $\phi_{Elm}(\vec{r})$ não são funções de quadrado integrável, consistentemente com o caráter contínuo do espectro de H , que mostra ainda uma degenerescência infinita mas *discreta*, correspondente aos diferentes valores de l e de m .

Uma questão importante que pode ser colocada aqui diz respeito à relação entre essas autofunções e as ondas planas (3.23) com a normalização (3.25), que são autofunções simultâneas de H e do momento \vec{p} . Como \vec{p} e \vec{L} não são compatíveis, as ondas planas não têm momento angular definido, mas podem ser expandidas nas autofunções $\phi_{Elm}(\vec{r})$, isto é

$$\frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{r}} = \sum_{l=0}^{\infty} \sum_{m=-l}^l C_{\vec{k}lm} j_l(kr) Y_{lm}(\theta, \varphi)$$

e os coeficientes $C_{\vec{k}lm}$, que devem depender explicitamente da direção do vetor $\vec{k} \equiv \{k, \theta_k, \varphi_k\}$ que aparece no primeiro membro, podem de fato ser determinados como sendo

$$C_{\vec{k}lm} = \frac{1}{(2\pi)^{3/2}} 4\pi i^l Y_{lm}^*(\theta_k, \varphi_k).$$

A expressão

$$e^{i\vec{k}\cdot\vec{r}} = 4\pi \sum_{lm} i^l Y_{lm}^*(\theta_k, \varphi_k) j_l(kr) Y_{lm}(\theta, \varphi) \quad (5.18)$$

é usualmente chamada *expansão em ondas parciais* (com momento angular bem definido) da onda plana, e exprime a relação geral existente entre as duas famílias de autofunções (impróprias) da partícula livre. Um caso particular importante da relação (5.18) é aquele em que o eixo z (ou 3) é tomado como sendo $\hat{k} \equiv \vec{k}/k$. Nesse caso, de fato, $\vec{k} \cdot \vec{r} = kr \cos \theta$ e, usando o *teorema da adição* para as harmônicas esféricas

$$\sum_{m=-l}^l Y_{lm}^*(\hat{k}) Y_{lm}(\hat{r}) = \frac{2l+1}{4\pi} P_l(\hat{k} \cdot \hat{r}),$$

onde $P_l(\hat{k} \cdot \hat{r})$ é um polinômio de Legendre, a (5.18) se reduz a

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta). \quad (5.19)$$

5.1.3 Poço de potencial quadrado em três dimensões.

Um problema envolvendo um potencial central não nulo, mas que pode ser resolvido completamente usando os resultados obtidos na seção anterior para a partícula livre é o de uma partícula sujeita a um potencial central com perfil “quadrado”, dado por

$$V(r) = \begin{cases} -V_0, & r < a_0 \\ 0, & r > a_0 \end{cases}.$$

De fato, a equação radial que deve ser resolvida nesse caso pode ser escrita como

$$\begin{aligned} r < a_0 : \quad & \frac{d^2 u_l}{dr^2} + \left(K^2 - \frac{l(l+1)}{r^2} \right) u_l = 0, \quad K^2 = \frac{2M}{\hbar^2}(E + V_0); \\ r > a_0 : \quad & \frac{d^2 u_l}{dr^2} + \left(k^2 - \frac{l(l+1)}{r^2} \right) u_l = 0, \quad k^2 = \frac{2ME}{\hbar^2}. \end{aligned}$$

Nas duas regiões $r < a_0$ e $r > a_0$ essa equação tem a mesma forma que a equação já resolvida, e admite portanto como solução geral uma combinação linear de funções de Bessel e de Neumann esféricas. Na região $r < a_0$ a condição de regularidade na origem exclui as funções de Neumann, mas essa exclusão *não* se aplica agora à região $r > a_0$, já que essa região *exclui* a origem. A função radial $u_l(r)$, correspondente ao valor l do momento angular, deve portanto ter a forma

$$\begin{aligned} r < a_0 : \quad & u_l(r < a_0) = C_l K r j_l(Kr); \\ r > a_0 : \quad & u_l(r > a_0) = kr[A_l j_l(kr) + B_l n_l(kr)]. \end{aligned}$$

É preciso ainda, no entanto, impor sobre $u_l(r)$ condições que garantam que essa função seja solução da equação radial *também no ponto* $r = a_0$. Para determinar essas condições basta usar os mesmos argumentos já utilizados na página 154 com relação à resolução da equação de Schrödinger para problemas unidimensionais. De fato, descontinuidades de $u_l(r)$ ou de $du_l(r)/dr$ em $r = a_0$ dão origem a contribuições singulares para a segunda derivada nesse ponto (proporcionais respectivamente a $d\delta(r - a_0)/dr$ e a $\delta(r - a_0)$). A descontinuidade finita do potencial em $r = a_0$ implica porém apenas uma descontinuidade também finita da *segunda derivada* de $u_l(r)$ nesse ponto, de modo que é preciso impor a continuidade de $u_l(r)$ e de $du_l(r)/dr$ em $r = a_0$. Portanto (cf. eqs. (3.37))

$$\lim_{\epsilon \rightarrow 0} [u_l(a_0 + \epsilon) - u_l(a_0 - \epsilon)] = 0 \quad \text{e} \quad \lim_{\epsilon \rightarrow 0} \left[\frac{du_l}{dr}(a_0 + \epsilon) - \frac{du_l}{dr}(a_0 - \epsilon) \right] = 0. \quad (5.20)$$

1. Espectro contínuo. Para energias positivas, $E > 0$, as funções j_l e n_l têm comportamento oscilatório limitado para $r > a_0$ e podem ser ambas admitidas na solução. Para determinar as três constantes A_l , B_l e C_l existem as duas condições de continuidade (5.20) além de uma terceira condição que consiste num critério de normalização. Isso significa que para todo valor de $E > 0$ existe uma solução que é única a menos de um fator de normalização, de modo que o espectro de H é contínuo e com a mesma degenerescência que o espectro da partícula livre.

Para implementar as condições (5.20) é conveniente reescrever as constantes A_l e B_l sob a forma

$$A_l = c_l \cos \delta_l, \quad B_l = c_l \sin \delta_l$$

com o que as (5.20) dão

$$\begin{aligned} C_l K a_0 j_l(K a_0) &= k a_0 c_l [\cos \delta_l j_l(k a_0) + \sin \delta_l n_l(k a_0)] \\ C_l K^2 a_0 j'_l(K a_0) &= k^2 a_0 c_l [\cos \delta_l j'_l(k a_0) + \sin \delta_l n'_l(k a_0)] \end{aligned}$$

onde j'_l e n'_l indicam derivadas com relação ao argumento das respectivas funções. Dividindo essas duas equações membro a membro resulta a equação para δ_l

$$\frac{j'_l(K a_0)}{j_l(K a_0)} = \frac{k}{K} \frac{\cos \delta_l j'_l(k a_0) + \sin \delta_l n'_l(k a_0)}{\cos \delta_l j_l(k a_0) + \sin \delta_l n_l(k a_0)}.$$

Uma vez determinado o ângulo δ_l , qualquer das duas equações determina C_l em termos de c_l , que por sua vez é determinado por critérios de normalização.

O comportamento de $u_l(r)$ para $r \rightarrow \infty$ pode ser determinado usando as expressões assintóticas (5.17)

$$u_l(r \rightarrow \infty) \rightarrow c_l \left[\cos \delta_l \sin \left(kr - \frac{l\pi}{2} \right) + \sin \delta_l \cos \left(kr - \frac{l\pi}{2} \right) \right] = c_l \sin \left(kr - \frac{l\pi}{2} + \delta_l \right). \quad (5.21)$$

O ângulo δ_l é chamado *defasagem* da onda parcial l e, como mostrado adiante, caracteriza o espalhamento de partículas com momento angular l pelo potencial $V(r)$.

2. Espectro discreto. Para energias negativas ($E < 0$) o número de onda k se torna imaginário na região $r > a_0$, o mesmo acontecendo portanto com o argumento das funções de Bessel e de Neumann. Pondo $k = i\kappa$, com κ real e positivo, o comportamento para $r \rightarrow \infty$ dessas funções é

$$j_l(i\kappa r) \rightarrow \frac{1}{i\kappa r} \operatorname{sen} \left(i\kappa r - \frac{l\pi}{2} \right) = -\frac{1}{2\kappa r} \left(e^{-\kappa r - il\pi/2} - e^{\kappa r + il\pi/2} \right)$$

$$n_l(i\kappa r) \rightarrow \frac{1}{i\kappa r} \cos \left(i\kappa r - \frac{l\pi}{2} \right) = -\frac{i}{2\kappa r} \left(e^{-\kappa r - il\pi/2} + e^{\kappa r + il\pi/2} \right)$$

que nos dois casos é dominado por termos exponencialmente crescentes com κr . As soluções relevantes da equação radial são no entanto as que se anulam no infinito, correspondendo a *estados ligados* da partícula no poço de potencial, e correspondem portanto, na região $r > a_0$, à combinação linear de j_l e n_l que se anula no infinito através do cancelamento dos termos exponencialmente crescentes, que é dada por

$$h_l^{(+)}(i\kappa r) = n_l(i\kappa r) + ij_l(i\kappa r) \rightarrow -\frac{i}{\kappa r} e^{-\kappa r - il\pi/2}.$$

As condições de continuidade (5.20) em $r = a_0$ dão então

$$\frac{j_l'(Ka_0)}{j_l(Ka_0)} = i \frac{\kappa}{K} \frac{h_l^{(+)'}(i\kappa a_0)}{h_l^{(+)}(i\kappa a_0)}, \quad (5.22)$$

e como não há constantes livres que possam ser ajustadas de modo a satisfazer essa relação, ela deve ser interpretada como *uma condição sobre E* para a existência de uma solução com as propriedades requeridas. O espectro é portanto discreto, e os autovalores E_{nl} são determinados pelas soluções da equação transcendente (5.22). Esses autovalores são claramente dependentes de l , e o índice adicional n distingue (quando for o caso) diferentes soluções para um dado valor de l .

No caso particular $l = 0$, a forma simples das funções radiais leva à condição

$$\frac{K}{\kappa} + \tan Ka_0 = 0$$

cujas soluções podem ser obtidas numérica ou graficamente. Não é difícil verificar que essa equação só pode ter soluções para $E > -V_0$, pois do contrário K também se torna imaginário, convertendo a tangente em uma tangente hiperbólica e levando a uma expressão que nunca se anula; e também que para que exista pelo menos uma solução com $-V_0 < E < 0$ é preciso que

$$V_0 a_0^2 \geq \frac{\hbar^2 \pi^2}{8M}.$$

5.1.4 Espectro contínuo de um potencial central de curto alcance. Defasagens.

O resultado (5.21) obtido para o espectro contínuo do poço de potencial quadrado em três dimensões mostra que o efeito do potencial na região assintótica $r \rightarrow \infty$ é introduzir, nessa região, defasagens δ_l relativamente à solução livre nas funções radiais correspondentes a cada uma das ondas parciais. Uma questão importante é delimitar a classe de potenciais centrais para os quais essa propriedade permanece válida.

Para estudar essa questão é conveniente examinar a solução da equação radial para a onda parcial l , envolvendo um potencial $V(r)$ que se anula para $r \rightarrow \infty$,

$$\frac{d^2 u_l}{dr^2} + \left(k^2 - U(r) - \frac{l(l+1)}{r^2} \right) u_l = 0,$$

onde $k^2 = 2ME/\hbar^2$ e $U(r) = 2MV(r)/\hbar^2$, utilizando uma parametrização da solução u_l e de sua derivada du_l/dr em termos de uma amplitude $A(r)$ e uma fase $\varphi(r)$ definidas através das relações

$$u_l(r) = A(r) \sin(kr + \varphi(r)) \quad \text{e} \quad \frac{du_l(r)}{dr} = A(r)k \cos(kr + \varphi(r)).$$

Derivando a segunda dessas relações com relação a r e substituindo o resultado na equação radial resulta

$$\frac{1}{A} \frac{dA}{dr} \cos(kr + \varphi) - \left(\frac{d\varphi}{dr} + \frac{U}{k} + \frac{l(l+1)}{kr^2} \right) \sin(kr + \varphi) = 0.$$

Por outro lado, derivando a parametrização proposta para $u_l(r)$ e comparando o resultado com a parametrização de du_l/dr segue que

$$\frac{1}{A} \frac{dA}{dr} \sin(kr + \varphi) = -\frac{d\varphi}{dr} \cos(kr + \varphi) \tag{5.23}$$

que pode ser usada para eliminar A da equação anterior, que fica então reduzida a

$$\frac{d\varphi}{dr} \cos^2(kr + \varphi) + \left(\frac{d\varphi}{dr} + \frac{U}{k} + \frac{l(l+1)}{kr^2} \right) \sin^2(kr + \varphi) = 0$$

ou ainda, finalmente,

$$\frac{d\varphi}{dr} = -\frac{1}{k} \left(U + \frac{l(l+1)}{r^2} \right) \tan^2(kr + \varphi).$$

Usando esse resultado em (5.23) resulta também que $A(r)$ satisfaz

$$\frac{1}{A} \frac{dA}{dr} = \frac{1}{2k} \left(U + \frac{l(l+1)}{r^2} \right) \sin(2kr + 2\varphi).$$

As duas últimas equações são equivalentes, através da parametrização adotada, à equação radial de partida. A primeira delas dá, fazendo a integração sobre r a partir de um ponto de referência qualquer r_0 ,

$$\varphi(r) = \varphi(r_0) + \int_{r_0}^r dr' \left(U + \frac{l(l+1)}{r'^2} \right) \sin^2(kr' + \varphi(r')).$$

O último fator no integrando é uma função não negativa e menor que 1, e portanto uma condição *suficiente* para que a integral convirja para $r \rightarrow \infty$ é que o potencial $U(r)$ se comporte de modo que

$$\lim_{r \rightarrow \infty} rU(r) = 0.$$

Nesse caso a fase $\varphi(r)$ tende assintoticamente para um valor constante. Potenciais satisfazendo essa condição serão chamados *potenciais de curto alcance*. É importante notar que ela *exclui*, em particular, o potencial coulombiano, para o qual de fato é possível mostrar que o comportamento assintótico de $u_l(r)$ é

$$u_l(r) \rightarrow c_l \sin \left(kr - \eta \ln 2kr - \frac{l\pi}{2} + \sigma_l \right)$$

onde, para o potencial α/r , o parâmetro η é dado por $\eta = \alpha M/\hbar^2 k$, e c_l e σ_l são independentes de r . A equação para $A(r)$ mostra que a condição de que o potencial seja de curto alcance é também suficiente para garantir que a amplitude $A(r)$ também tenda assintoticamente para um valor constante. Como, porém, o integrando contém nesse caso um fator limitado em módulo mas *não positivo*, a constância assintótica de $A(r)$ permanece válida em condições algo mais gerais que as exigidas para a constância assintótica de $\varphi(r)$. Em particular, essa amplitude tende para um valor constante no caso do potencial coulombiano.

Portanto, para potenciais de curto alcance as soluções da equação radial para o espectro contínuo se comportam assintoticamente como

$$u_l(r \rightarrow \infty) \rightarrow c_l \sin \left(kr - \frac{l\pi}{2} + \delta_l \right)$$

onde c_l e δ_l são independentes de r . Como c_l depende de questões de normalização, *toda* a informação relativa aos efeitos do potencial de curto alcance disponível no comportamento

assintótico da função de onda está contido nas defasagens δ_l . Experiências de espalhamento envolvem tipicamente medidas feitas a grandes distâncias do centro espalhador (em termos de suas dimensões características), de modo que o acesso que elas podem ter às propriedades do potencial espalhador deve se dar necessariamente através dessas quantidades.

5.1.5 Condições de contorno, amplitudes e seção de choque de espalhamento.

As autofunções simultâneas do Hamiltoniano rotacionalmente invariante e dos operadores de momento angular L^2 e L_3 , embora importantes do ponto de vista da resolução do problema tri-dimensional de autovalores, não são diretamente interpretáveis em termos experimentais no caso do espectro contínuo. Classicamente, um valor definido do momento angular numa situação de espalhamento significa um valor definido do parâmetro de impacto e do momento da partícula incidente com relação ao centro do potencial. Quanticamente, no entanto, as limitações inerentes às relações de incerteza não permitem a determinação simultânea dessas quantidades. Situações experimentalmente realizáveis envolvem estados que, longe do potencial espalhador, são essencialmente autoestados do momento linear, isto é, ondas planas.

Para uma partícula livre, tomando o eixo 3 do sistema de coordenadas paralelo ao momento $\hbar\vec{k}$, vale a decomposição em ondas parciais

$$\phi_{\vec{k}}^{(0)}(\vec{r}) = e^{ikz} = \sum_{l=0}^{\infty} (2l+1) i^l P_l(\cos\theta) j_l(kr), \quad z = r \cos\theta$$

que, para valores suficientemente grandes de r , se reduz a

$$\phi_{\vec{k}}^{(0)}(\vec{r}) \rightarrow \frac{1}{kr} \sum_{l=0}^{\infty} (2l+1) i^l P_l(\cos\theta) \sin\left(kr - \frac{l\pi}{2}\right).$$

Em presença de um potencial de curto alcance $V(r)$, por outro lado, o comportamento assintótico da onda parcial l foi identificado como sendo

$$u_l(r) \rightarrow c_l \sin\left(kr - \frac{l\pi}{2} + \delta_l\right)$$

que, no caso em que $V(r) = 0$ e portanto $\delta_l = 0$, reproduz o resultado correspondente à onda plana com $c_l = 1$. Uma solução de interesse para a descrição de processos de espalhamento envolve uma superposição de ondas parciais $u_l^{(+)}(r)$, distorcidas pelos efeitos do potencial, com a escolha de amplitudes

$$c_l^{(+)} = e^{i\delta_l}$$

as quais se reduzem novamente a 1 no caso de potencial nulo. A autofunção correspondente do Hamiltoniano tem então o comportamento assintótico

$$\begin{aligned}\phi_k^{(+)}(\vec{r}) &= \frac{1}{kr} \sum_{l=0}^{\infty} (2l+1) i^l P_l(\cos \theta) u_l^{(+)}(r) \\ &\rightarrow \frac{1}{kr} \sum_{l=0}^{\infty} (2l+1) i^l P_l(\cos \theta) e^{i\delta_l} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right).\end{aligned}\quad (5.24)$$

Vale a pena notar que a função $\phi_k^{(+)}(\vec{r})$ está sendo construída como uma combinação linear particular de diferentes autofunções de H com o mesmo autovalor, embora associadas a valores diferentes do momento angular. Os fatores de cada termo que correspondem à forma assintótica da função radial podem ser re-escritos como

$$\begin{aligned}e^{i\delta_l} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right) &= \frac{1}{2i} \left(e^{2i\delta_l} e^{i(kr-l\pi/2)} - e^{-i(kr-l\pi/2)} \right) \\ &= \frac{1}{2i} \left(e^{2i\delta_l} - 1 \right) (-i)^l e^{ikr} + \sin \left(kr - \frac{l\pi}{2} \right).\end{aligned}$$

Substituindo essa forma na expansão assintótica de $\phi_k^{(+)}(\vec{r})$, o último termo reproduz a forma assintótica da onda plana e^{ikz} , de modo que

$$\phi_k^{(+)}(\vec{r}) \rightarrow e^{ikz} + \frac{e^{ikr}}{2ikr} \sum_{l=0}^{\infty} (2l+1) \left(e^{2i\delta_l} - 1 \right) P_l(\cos \theta) \equiv e^{ikz} + f_k^{(+)}(\theta) \frac{e^{ikr}}{r}$$

onde foi introduzida a função de θ

$$f_k^{(+)}(\theta) \equiv \frac{1}{2ik} \sum_{l=0}^{\infty} (2l+1) \left(e^{2i\delta_l} - 1 \right) P_l(\cos \theta) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos \theta). \quad (5.25)$$

Portanto a autofunção $\phi_k^{(+)}(\vec{r})$ de H se comporta assintoticamente como a superposição de uma onda plana com uma onda esférica emergente do centro espalhador, modulada angularmente pela amplitude $f_k^{(+)}(\theta)$.

Outras autofunções do Hamiltoniano, satisfazendo condições de contorno diferentes dessa, podem ser construídas de forma semelhante, superpondo com amplitudes diferentes as soluções correspondentes às várias ondas parciais. Por exemplo, tomando a superposição

$$\phi_k^{(-)}(\vec{r}) = \frac{1}{kr} \sum_{l=0}^{\infty} (2l+1) i^l P_l(\cos \theta) u_l^{(-)}(r) \quad (5.26)$$

onde agora as funções radiais $u_l^{(-)}(r)$ são definidas de forma que

$$u_l^{(-)}(r \rightarrow \infty) \rightarrow e^{-i\delta_l} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right)$$

e repetindo nesse caso o desenvolvimento usado no caso de $\phi_k^{(+)}(\vec{r})$, resulta que a autofunção (5.26) se comporta assintoticamente como a superposição de uma onda plana com uma onda esférica *imergente* para o centro espalhador, modulada angularmente por uma amplitude $f_k^{(-)}(\theta)$, isto é

$$\phi_k^{(-)}(\vec{r}) \rightarrow e^{ikz} + f_k^{(-)}(\theta) \frac{e^{-ikr}}{r}$$

com

$$f_k^{(-)}(\theta) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) (-1)^l e^{-i\delta_l} \sin \delta_l P_l(\cos \theta).$$

As condições de contorno assintóticas que caracterizam as autofunções (5.24) e (5.26)

$$\phi_k^{(\pm)}(\vec{r}) \rightarrow e^{ikz} + f_k^{(\pm)}(\theta) \frac{e^{\pm ikr}}{r}$$

são usualmente chamadas *condições de contorno de espalhamento* com ondas *emergentes* e *imergentes* respectivamente. Essas autofunções do espectro contínuo do Hamiltoniano desempenham um papel central na teoria geral do espalhamento. Suas ondas parciais estão relacionadas por

$$u_l^{(+)}(r) = e^{2i\delta_l} u_l^{(-)}(r)$$

e as fases $S_l \equiv e^{2i\delta_l}$ são na realidade autovalores de uma matriz unitária chamada *matriz de espalhamento* (ou “matriz S ”), que é diagonal na representação em que o momento angular é diagonal, no caso de potenciais centrais.

A forma assintótica da autofunção com ondas emergentes (5.24) está diretamente relacionada com a secção de choque de espalhamento pelo potencial central $V(r)$. Um argumento heurístico que permite explicitar essa relação consiste em interpretar o fluxo da onda esférica emergente, na região assintótica, como o fluxo de partículas espalhadas pelo potencial. Nesse tipo de argumento são ignorados efeitos de interferência entre a onda plana e a onda esférica espalhada,

a pretexto de que a primeira é bloqueada por um sistema de colimação montado de forma que o fluxo incidente não atinja o detector. Isso implica numa distorção da onda plana, mas para dimensões macroscópicas do sistema de colimação essa distorção pode ser desprezada para efeito da descrição do processo de espalhamento.

Dentro desse esquema, a seção de choque diferencial $d\sigma$ deve ser definida como a corrente da onda espalhada que passa através do ângulo sólido $d\Omega$ dividida pelo fluxo da onda plana incidente. A densidade de corrente espalhada na direção \hat{r} é dada por (3.15) como

$$\vec{j} \cdot \hat{r} = \frac{\hbar}{2M} 2 \operatorname{Im} \left[\left(f_k^{(+)}(\theta) \frac{e^{ikr}}{r} \right)^* \nabla \cdot \hat{r} \left(f_k^{(+)}(\theta) \frac{e^{ikr}}{r} \right) \right] = \frac{\hbar k}{M} \frac{1}{r^2} |f_k^{(+)}(\theta)|^2.$$

A corrente dI através do ângulo sólido $d\Omega$ será portanto

$$dI = \vec{j} \cdot \hat{r} r^2 d\Omega = \frac{\hbar k}{M} |f_k^{(+)}(\theta)|^2.$$

Por outro lado, como o fluxo incidente da onda plana com amplitude 1 é dado por $\hbar k/M$, resulta que a secção de choque diferencial é dada simplesmente por

$$\frac{d\sigma}{d\Omega} = |f_k^{(+)}(\theta)|^2. \quad (5.27)$$

Como a secção de choque é dada pelo módulo quadrado de $f_k^{(+)}(\theta)$, essa função é chamada *amplitude de espalhamento* associada ao potencial $V(r)$. Como foi visto, ela é completamente definida pelas defasagens δ_l para as diferentes ondas parciais, o mesmo acontecendo portanto para a seção de choque diferencial, que tem a forma de uma soma dupla

$$\frac{d\sigma}{d\Omega} = \sum_{l,l'=0}^{\infty} \frac{(2l+1)(2l'+1)}{k^2} e^{i(\delta_l - \delta_{l'})} \sin \delta_l \sin \delta_{l'} P_l(\cos \theta) P_{l'}(\cos \theta)$$

onde é importante notar a existência de termos de interferência entre diferentes ondas parciais. A seção de choque total é obtida integrando essa expressão sobre o ângulo sólido $d\Omega$, o que pode ser feito usando a propriedade de ortogonalidade dos polinômios de Legendre

$$\int d\Omega P_l(\cos \theta) P_{l'}(\cos \theta) = \frac{4\pi}{2l+1} \delta_{ll'}$$

que exclui os termos de interferência levando ao resultado

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l.$$

Essa expressão dá a secção de choque total como uma soma de contribuições de cada uma das ondas parciais

$$\sigma_l = \frac{4\pi}{k^2} (2l + 1) \sin^2 \delta_l$$

que tem portanto como limite superior $4\pi(2l + 1)/k^2$. Esse limite está relacionado com a conservação da probabilidade total pela equação de movimento quântica e também com a unitariedade da matriz S , o que lhe vale a denominação de “limite de unitariedade”.

Existe uma relação extremamente curiosa e importante entre a secção de choque total e a amplitude de espalhamento na direção frontal, $\theta = 0$. De fato, como $P_l(1) = 1$,

$$f_k^{(+)}(\theta = 0) = \frac{1}{k} \sum_{l=0}^{\infty} (2l + 1) e^{i\delta_l} \sin \delta_l = \frac{1}{k} \sum_{l=0}^{\infty} (2l + 1) (\cos \delta_l \sin \delta_l + i \sin^2 \delta_l)$$

e portanto

$$\sigma = \frac{4\pi}{k} \operatorname{Im} f_k^{(+)}(\theta = 0).$$

Essa é uma relação bastante peculiar, no sentido que relaciona *linearmente* uma “probabilidade”, σ , com uma “amplitude”, $f_k^{(+)}(\theta = 0)$. Ela é conhecida como *teorema óptico*, e está ligada, mais uma vez, à conservação da probabilidade total no problema de espalhamento. O papel especial que esse resultado atribui à direção frontal $\theta = 0$ está ligado ao fato de que o argumento de “não interferência” da onda plana com a onda esférica espalhada certamente de torna inválido nessa direção; a interferência dessas duas ondas em $\theta = 0$ deve então cancelar uma parte da corrente incidente suficiente para compensar toda a corrente espalhada, medida pela secção de choque total.

Finalmente, uma questão de considerável importância prática é a do número de ondas parciais que contribuem apreciavelmente para a amplitude de espalhamento. Do ponto de vista clássico, uma partícula com momento angular $\hbar l$ e momento linear de módulo $\hbar k$ na região assintótica é caracterizada por um parâmetro de impacto $b = l/k$. Isso significa que, para um dado valor de $\hbar k$, o parâmetro de impacto cresce linearmente com o momento angular $\hbar l$, de modo que, para um potencial cujo alcance pode ser estimado em termos de uma dada distância R , haverá um valor máximo do momento angular, que pode ser estimado como $l_{\max} \sim kR$, além do qual os efeitos do potencial sobre o movimento da partícula serão desprezíveis.

Uma situação semelhante pode ser identificada também no contexto da mecânica quântica. De fato, a equação radial para a onda parcial l contém os efeitos efetivamente repulsivos do potencial centrífugo $l(l + 1)/r^2$, que se tornam mais importantes à medida que o valor de l aumenta. Esses efeitos tendem a reduzir a amplitude radial $u_l(r)$ na região em que $V(r)$

atua, e se o potencial tem um alcance que pode ser estimado em termos de uma distância R , haverá um valor l_{\max} do momento angular a partir do qual os efeitos do potencial sobre $u_l(r)$ serão desprezíveis. Isso significa que as amplitudes radiais para $l > l_{\max}$ serão essencialmente amplitudes radiais livres, com $\delta_l \simeq 0$. Esse argumento indica que, na decomposição em ondas parciais da amplitude de espalhamento, apenas os termos com $l < l_{\max}$ serão relevantes. A expressão clássica $l_{\max} \sim kR$ para o momento angular máximo é freqüentemente útil para uma estimativa do número de ondas parciais relevantes.

5.1.6 Espalhamento a muito baixas energias. Expansão de alcance efetivo.

No limite de energias muito baixas, para um potencial de curto alcance cuja extensão espacial é caracterizada pelo comprimento R , $kR \ll 1$, e apenas o termo $l = 0$ contribui apreciavelmente para a expansão em ondas parciais da amplitude de espalhamento, que se reduz então a

$$f_k^{(+)}(\theta) \rightarrow \frac{1}{k} e^{i\delta_0} \sin \delta_0 \quad kR \ll 1.$$

Como $P_{l=0}(\theta) = 1$ a seção de choque diferencial se torna isotrópica (independente de θ), sendo dada por

$$\frac{d\sigma}{d\Omega} \rightarrow \frac{1}{k^2} \sin^2 \delta_0.$$

Embora baseado na estimativa clássica, esse resultado é válido para uma classe muito grande de potenciais. É preciso apenas excluir potenciais que, embora sendo de curto alcance, tenham caudas atrativas do tipo $r^{-(1+\epsilon)}$ com $0 < \epsilon \leq 1$, pois esses potenciais podem compensar os efeitos do potencial centrífugo pelo menos até algum valor de l .

No limite $k \rightarrow 0$, a seção de choque pode permanecer finita. Nesse caso a defasagem δ_0 deve se anular linearmente com k , o que torna inconveniente a parametrização da amplitude de espalhamento em termos de δ_0 . A quantidade relevante para caracterizar a seção de choque em energias muito próximas de zero é então

$$a \equiv - \lim_{k \rightarrow 0} \frac{\sin \delta_0(k)}{k}$$

que tem dimensões de comprimento e é por isso chamada *comprimento de espalhamento*, em termos do qual a seção de choque diferencial é dada por

$$\frac{d\sigma}{d\Omega} = a^2.$$

O sinal na definição de a é convencional. Uma interpretação “geométrica” do comprimento de espalhamento que justifica, em particular, a convenção de sinal adotada pode ser obtida da seguinte forma. Seja $v_0(r)$ uma solução da equação radial *livre*, para $l = 0$,

$$\frac{d^2 v_0}{dr^2} + k^2 v_0 = 0$$

construída de forma a coincidir, na região assintótica, com a solução $u_0(r)$ da equação radial que inclui o potencial

$$\frac{d^2 u_0}{dr^2} + [k^2 - U(r)]u_0 = 0, \quad u_0(r = 0) = 0,$$

isto é

$$v_0(r) = c \sin(kr + \delta_0) \quad \text{para} \quad u_0(r \rightarrow \infty) \rightarrow c \sin(kr + \delta_0). \quad (5.28)$$

É claro que, se $\delta_0 \neq 0$, a função $v_0(r)$ não se anula na origem. Nesse caso é possível fixar uma normalização conveniente para essa função impondo a condição $v_0(r = 0) = 1$, que corresponde a $c = 1/\sin \delta_0$ e leva ao resultado

$$v_0(r) = \frac{1}{\sin \delta_0} \sin(kr + \delta_0) = \cos(kr) + \sin(kr) \cot \delta_0.$$

Tomando agora o limite $k \rightarrow 0$ resulta

$$\lim_{k \rightarrow 0} v_0(r) = 1 + \lim_{k \rightarrow 0} kr \cot \delta_0 = 1 + \left(\lim_{k \rightarrow 0} k \cot \delta_0 \right) r \equiv 1 - \frac{r}{a}.$$

A quantidade a introduzida na expressão final coincide de fato com o comprimento de espalhamento definido anteriormente quando o limite entre parênteses for finito, pois nesse caso δ_0 deve tender a zero linearmente com k e

$$\lim_{k \rightarrow 0} k \cot \delta_0 = \lim_{k \rightarrow 0} k \frac{1}{\sin \delta_0} = -\frac{1}{a}. \quad (5.29)$$

No limite $k \rightarrow 0$ a função $v_0(r)$ se reduz, portanto, a uma reta que tem o valor 1 para $r = 0$ e que corta o eixo das abscissas no ponto $r = a$. As figuras 5.1 e 5.2 ilustram essa construção geométrica para o comprimento de espalhamento no caso de potenciais simples que são respectivamente puramente atrativos e puramente repulsivos. A inspeção dessas figuras revela que existe, embora *apenas até certo ponto*, uma relação entre o sinal do comprimento de espalhamento a e o caráter repulsivo ou atrativo do potencial correspondente. De fato, $a > 0$ para potenciais puramente repulsivos, e $a < 0$ para potenciais puramente atrativos *quando não haja nenhum estado ligado*

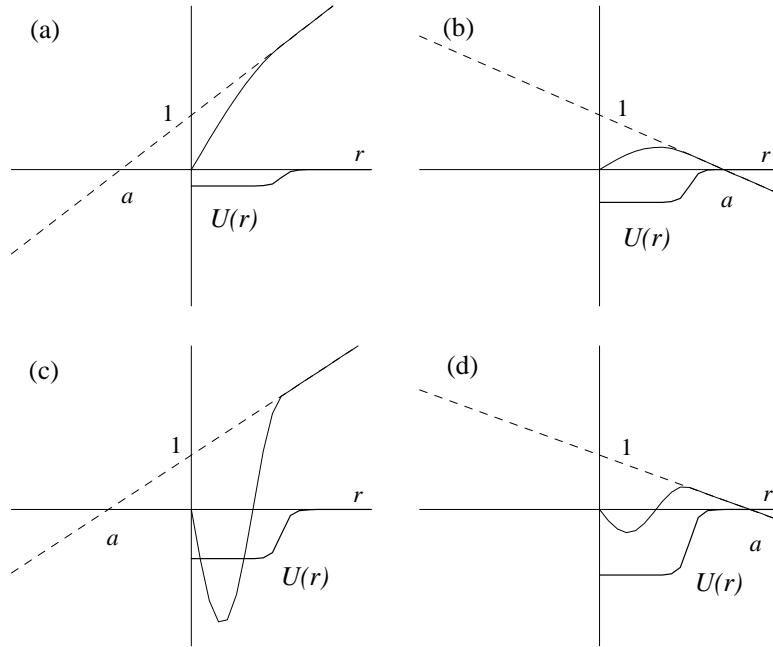
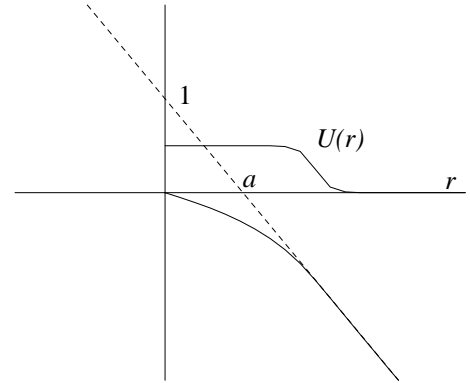


Figura 5.1: Funções $u_0(r)$ (linhas cheias) e $v_0(r)$ (linhas tracejadas) para quatro potenciais *atrativos* $U(r)$ de mesmo alcance mas com diferentes profundidades (representadas apenas qualitativamente, sem escala) e $k \rightarrow 0$. Essas funções são normalizadas de forma que $v_0(0) = 1$, e $u_0(r) = v_0(r)$ para valores de r para os quais $U(r) \simeq 0$. Em (a) o potencial não tem nenhum estado ligado. Em (b) e (c) há apenas um estado ligado, enquanto em (d) há dois estados ligados com $l = 0$.

Figura 5.2: Funções $u_0(r)$ (linha cheia) e $v_0(r)$ (linha tracejada) para um potencial *repulsivo* $U(r)$ e $k \rightarrow 0$, cf. figura 5.1



com $l = 0$. Esta última restrição é indispensável pois, como pode ser visto nos casos (b), (c) e (d) da figura 5.1, quando o potencial atrativo admitir estados ligados com $l = 0$ não existe qualquer restrição geral sobre o sinal de a .

A expressão (5.29) pode ser vista como relacionando o comprimento de espalhamento com o primeiro termo de uma expansão de $k \cot \delta_0$ em potências de k . Uma forma de refinar a descrição do espalhamento em energias próximas de zero consiste em obter termos adicionais dessa expansão, conhecida como *expansão de alcance efetivo*

$$k \cot \delta_0 = -\frac{1}{a} + k \frac{d}{dk} (k \cot \delta_0)_{k=0} + \frac{k^2}{2} \frac{d^2}{dk^2} (k \cot \delta_0)_{k=0} + \dots \quad (5.30)$$

para o que é necessário determinar a dependência de δ_0 com k para valores próximos de zero dessa variável. Uma forma de fazer isso consiste em considerar *um par* de soluções $v_0^{(i)}$ e $u_0^{(i)}$, $i = 1, 2$ construídas exatamente como em (5.28) mas para dois valores diferentes k_1 e k_2 do módulo do vetor de onda. Multiplicando a equação radial para $u_0^{(1)}$ por $u_0^{(2)}$ e vice versa, e em seguida subtraindo as expressões resultantes, se obtém

$$u_0^{(2)} \frac{d^2 u_0^{(1)}}{dr^2} - u_0^{(1)} \frac{d^2 u_0^{(2)}}{dr^2} = \frac{d}{dr} \left(u_0^{(2)} \frac{du_0^{(1)}}{dr} - u_0^{(1)} \frac{du_0^{(2)}}{dr} \right) = (k_2^2 - k_1^2) u_0^{(1)} u_0^{(2)}. \quad (5.31)$$

Esse procedimento elimina a intervenção explícita do potencial no resultado. Procedendo da mesma forma com relação às funções $v_0^{(i)}$ se obtém também

$$\frac{d}{dr} \left(v_0^{(2)} \frac{dv_0^{(1)}}{dr} - v_0^{(1)} \frac{dv_0^{(2)}}{dr} \right) = (k_2^2 - k_1^2) v_0^{(1)} v_0^{(2)}. \quad (5.32)$$

As constantes de normalização de $u_0^{(i)}$ e $v_0^{(i)}$ são escolhidas da mesma forma que no tratamento anterior, isto é,

$$u_0^{(i)}(r \rightarrow \infty) \rightarrow \frac{1}{\sin \delta_0^{(i)}} \sin(k_i r + \delta_0^{(i)}) = v_0^{(i)}(r), \quad v_0^{(i)}(0) = 1.$$

Subtraindo as relações (5.31) e (5.32), e integrando o resultado sobre r usando o fato de que $u_0^{(i)}(0) = 0$ e $v_0^{(i)}(0) = 1$

$$\left(\frac{dv_0^{(2)}}{dr} \right)_{r=0} - \left(\frac{dv_0^{(1)}}{dr} \right)_{r=0} = (k_2^2 - k_1^2) \int_0^\infty dr [v_0^{(1)} v_0^{(2)} - u_0^{(1)} u_0^{(2)}]$$

em que as derivadas que aparecem no primeiro membro são

$$\left(\frac{dv_0^{(i)}}{dr}\right)_{r=0} = \left(\frac{d}{dr} \frac{\text{sen}(k_i r + \delta_0^{(i)})}{\text{sen} \delta_0^{(i)}}\right)_{r=0} = k_i \cot \delta_0^{(i)}.$$

Fazendo então $k_1 = 0$ e $k_2 = k$ e usando a definição do comprimento de espalhamento resulta finalmente

$$\begin{aligned} k \cot \delta_0(k) &= -\frac{1}{a} + \frac{k^2}{2} \int_0^\infty dr \, 2 \left[v_0^{(k)} v_0^{(0)} - u_0^{(k)} u_0^{(0)} \right] = \\ &= -\frac{1}{a} + \frac{k^2}{2} \int_0^\infty dr \, 2 \left[v_0^{(0)2} - u_0^{(0)2} \right] \\ &\quad + \frac{k^2}{2} \int_0^\infty dr \, 2 \left[(v_0^{(k)} - v_0^{(0)}) v_0^{(0)} - (u_0^{(k)} - u_0^{(0)}) u_0^{(0)} \right]. \end{aligned}$$

A integral que aparece no último termo se anula para $k \rightarrow 0$, de modo que esse termo é de ordem superior a k^2 . Desse modo o termo linear em k na expansão de alcance efetivo (5.30) é nulo, e o coeficiente do termo quadrático pode ser expresso como

$$\frac{d^2}{dk^2} (k \cot \delta_0)_{k=0} \equiv r_0 = \int_0^\infty dr \, 2 \left[v_0^{(0)2} - u_0^{(0)2} \right].$$

O parâmetro r_0 , que tem dimensões de comprimento, é chamado *alcance efetivo* do potencial $V(r)$.

5.2 Problema de dois corpos com fôrças centrais. O átomo de hidrogênio.

Um sistema de dois corpos, de massas M_1 e M_2 , que interagem mutuamente através de fôrças centrais derivadas de um potencial, é descrito pelo hamiltoniano

$$H = \frac{p_1^2}{2M_1} + \frac{p_2^2}{2M_2} + V(|\vec{r}_1 - \vec{r}_2|)$$

onde as componentes cartesianas dos momentos \vec{p}_i são canonicamente conjugadas às componentes cartesianas respectivas de \vec{r}_i . Esse Hamiltoniano atua num espaço de estados que pode ser implementado como o produto dos espaços de estados de cada um dos corpos considerados isoladamente. Explicitamente, se $\{|\phi_n^{(1)}\rangle\}$ e $\{|\phi_n^{(2)}\rangle\}$ são bases (por conveniência ortonormais)

dos espaços de um corpo, então uma base do espaço do sistema de dois corpos pode ser escrita em termos dos produtos $\{|\phi_m^{(1)}\rangle \otimes |\phi_n^{(2)}\rangle\} \equiv \{|\phi_m^{(1)}\phi_n^{(2)}\rangle\}$, o que significa que um vetor de estado geral $|\varphi\rangle$ do sistema de dois corpos pode ser escrito como

$$|\varphi\rangle = \sum_{m,n} c_{mn} |\phi_m^{(1)}\phi_n^{(2)}\rangle.$$

Variáveis dinâmicas de um dos dois corpos, como por exemplo \vec{p}_1 , são estendidas ao espaço de estados de dois corpos definindo a sua ação sobre os vetores desse espaço de forma que atuem como o operador identidade no fator correspondente ao outro corpo, isto é

$$\vec{p}_1|\varphi\rangle = \sum_{m,n} c_{mn} (\vec{p}_1|\phi_m^{(1)}\rangle) \otimes (\hat{\mathbf{1}}_2|\phi_n^{(2)}\rangle).$$

Essa extensão é às vezes explicitada escrevendo o operador \vec{p}_1 , visto como operador no espaço produto, como $\vec{p}_1 \otimes \hat{\mathbf{1}}_2$. Com essa notação explícita, o problema de autovalores para o Hamiltoniano de dois corpos se escreve como

$$\left(\frac{p_1^2}{2M_1} \otimes \hat{\mathbf{1}}_2 + \hat{\mathbf{1}}_1 \otimes \frac{p_2^2}{2M_2} + V(|\vec{r}_1 - \vec{r}_2|) \right) |\phi_E\rangle = E|\phi_E\rangle.$$

Ao contrário dos termos de energia cinética, o termo de energia potencial presente nessa equação introduz correlações entre as variáveis dinâmicas dos dois corpos, o que torna difícil a solução do problema, colocado nessa forma.

De modo inteiramente análogo ao utilizado no contexto da mecânica clássica, é possível no entanto mudar a formulação do problema de dois corpos de modo a contornar esse tipo de dificuldade. De fato, introduzindo variáveis dinâmicas do centro de massa

$$\vec{R} \equiv \frac{M_1\vec{r}_1 + M_2\vec{r}_2}{M_1 + M_2}, \quad \vec{P} \equiv \vec{p}_1 + \vec{p}_2$$

bem como variáveis dinâmicas referentes ao movimento relativo dos dois corpos

$$\vec{r} \equiv \vec{r}_2 - \vec{r}_1, \quad \vec{p} \equiv \frac{M_1\vec{p}_2 - M_2\vec{p}_1}{M_1 + M_2}$$

o Hamiltoniano de dois corpos assume a forma

$$H = \frac{P^2}{2M} + \frac{p^2}{2\mu} + V(r) \equiv H_R + H_r, \quad M \equiv M_1 + M_2, \quad \mu \equiv \frac{M_1M_2}{M_1 + M_2}$$

onde $\{\vec{R}, \vec{P}\}$ e $\{\vec{r}, \vec{p}\}$ também correspondem a pares de variáveis canonicamente conjugadas, e portanto $[H_R, H_r] = 0$. Com essa reformulação, portanto, o problema de dois corpos com

forças centrais se reduz a um problema efetivo de dois corpos *sem interação mútua*. Um dos corpos do problema efetivo corresponde de fato ao centro de massa de M_1 e M_2 , associado a um Hamiltoniano *livre* de um corpo, enquanto o segundo corpo, correspondente à dinâmica relativa, é associado a um Hamiltoniano de uma partícula com massa igual à massa reduzida μ e submetida ao potencial central $V(r)$. Nessas condições, a forma conveniente de implementar o espaço de estados consiste em toma-lo como o espaço produto dos estados do centro de massa (com operador identidade $\hat{\mathbf{1}}_R$) pelo espaço dos estados do movimento relativo (com operador identidade $\hat{\mathbf{1}}_r$). O problema de autovalores fica, portanto

$$\left[\frac{P^2}{2M} \otimes \hat{\mathbf{1}}_r + \hat{\mathbf{1}}_R \otimes \left(\frac{p^2}{2\mu} + V(r) \right) \right] |\phi_e\rangle = E |\phi_e\rangle.$$

Nesse caso a dinâmica não envolve correlações entre os dois conjuntos de variáveis, e é possível procurar autovetores simultâneos de H_R e de H_r sob a forma fatorada

$$|\phi_E\rangle = |\phi_R\rangle \otimes |\phi_r\rangle$$

onde

$$H_R |\phi_R\rangle = \frac{P^2}{2M} |\phi_R\rangle = E_R |\phi_R\rangle \quad \text{e} \quad H_r |\phi_r\rangle = \left(\frac{p^2}{2\mu} + V(r) \right) |\phi_r\rangle = E_r |\phi_r\rangle$$

com $E_R + E_r = E$.

Estas últimas equações já foram tratadas nas seções anteriores. Usando a representação de Schrödinger, as autofunções do problema referente ao centro de massa, tomadas como autofunções simultâneas do momento \vec{P} , são ondas planas

$$\phi_R(\vec{R}) \equiv \langle \vec{R} | \phi_R \rangle = \frac{1}{(2\pi)^{2/3}} e^{i\vec{K} \cdot \vec{R}}, \quad E_R = \frac{\hbar^2 K^2}{2M}$$

e as autofunções do problema referente à dinâmica relativa, tomadas como autofunções simultâneas de L^2 e L_3 com $\vec{L} = \vec{r} \times \vec{p}$, são

$$\phi_r(\vec{r}) \equiv \langle \vec{r} | \phi_r \rangle = \frac{u_{E_r, l}(r)}{r} Y_{lm}(\theta, \varphi)$$

onde $u_{E_r, l}(r)$ satisfaz a equação radial

$$\frac{d^2 u_{E_r, l}}{dr^2} + \left[\frac{2\mu E_r}{\hbar^2} - \frac{l(l+1)}{r^2} - \frac{2\mu V(r)}{\hbar^2} \right] u_{E_r, l}(r) = 0.$$

As autofunções resultantes para o problema de dois corpos são portanto

$$\phi_E(\vec{R}, \vec{r}) \equiv \langle \vec{R} | \phi_R \rangle \langle \vec{r} | \phi_r \rangle = \frac{1}{(2\pi)^{2/3}} e^{i\vec{K} \cdot \vec{R}} \phi_r(\vec{r}) = \frac{1}{(2\pi)^{2/3}} e^{i\vec{K} \cdot \frac{M_1 \vec{r}_1 + M_2 \vec{r}_2}{M_1 + M_2}} \phi_r(\vec{r}_2 - \vec{r}_2) \quad (5.33)$$

onde as correlações dinâmicas entre os corpos M_1 e M_2 aparecem sob a forma da não fatorizabilidade desse resultado nas variáveis \vec{r}_1 e \vec{r}_2 .

Esses resultados permitem tratar um modelo simples para “átomos hidrogenóides”, que são sistemas consistindo de um núcleo puntiforme de massa M_1 e carga $+Ze$ interagindo eletrostaticamente com um elétron, também puntiforme, de massa M_2 e carga $-e$. Nesse caso

$$V(r) = -\frac{Ze^2}{r}$$

e a equação radial a ser resolvida para a dinâmica relativa é

$$\frac{d^2 u_{E_r, l}}{dr^2} + \left[\frac{2\mu E_r}{\hbar^2} - \frac{l(l+1)}{r^2} + \frac{2\mu Ze^2}{\hbar^2 r} \right] u_{E_r, l}(r) = 0.$$

O potencial coulombiano se anula para $r \rightarrow \infty$, mas não é um potencial de curto alcance no sentido da seção 5.1.4. Para valores positivos de E_r o espectro é contínuo e o comportamento assintótico das amplitudes radiais $u_{E_r, l}(r)$ é

$$u_{E_r, l}(r \rightarrow \infty) \rightarrow c_l \sin \left(kr - \eta \ln 2kr - \frac{l\pi}{2} + \sigma_l \right)$$

onde $k = \sqrt{2\mu E_r / \hbar^2}$, $\eta = Ze^2 \mu / \hbar^2 k$ é o chamado “parâmetro coulombiano” e a “defasagem coulombiana” σ_l é dada em termos da função gama por

$$\sigma_l = \arg \Gamma(l+1+i\eta).$$

Para valores negativos de E_r o espectro é discreto com um número infinito de estados ligados para cada valor do momento angular l . Para obter a forma desse espectro é conveniente introduzir a variável adimensional

$$\rho \equiv 2\sqrt{\frac{2\mu |E_r|}{\hbar^2}} r$$

bem como re-escrever o potencial coulombiano em termos do parâmetro, também adimensional

$$\lambda(E_r) \equiv \sqrt{\frac{\mu}{2|E_r|}} \frac{Ze^2}{\hbar} \quad (5.34)$$

com o que a equação radial adquire a forma

$$\frac{d^2 u_{\lambda,l}}{d\rho^2} - \left[\frac{1}{4} + \frac{l(l+1)}{\rho^2} - \frac{\lambda}{\rho} \right] u_{\lambda,l} = 0.$$

As soluções dessa equação que tem comportamento suficientemente regular na origem podem ser escritas sob a forma

$$u_{\lambda,l}(\rho) = e^{-\rho/2} \rho^s F(\rho)$$

com $F(\rho=0) = A_0 \neq 0$ e $s \geq 1$. Substituindo essa forma na equação diferencial se obtém

$$\rho^2 \frac{d^2 F}{d\rho^2} + \rho(2s - \rho) \frac{dF}{d\rho} + [s(s-1) - l(l+1) + (\lambda - s)\rho] F = 0$$

que, para $\rho=0$, dá $s = l+1$. Com essa escolha para s a equação para F adquire a forma

$$\rho^2 \frac{d^2 F}{d\rho^2} + \rho(2l+2 - \rho) \frac{dF}{d\rho} + \rho(\lambda - l - 1) F = 0.$$

Expandindo agora $F(\rho)$ em série de potências

$$F(\rho) = \sum_{k \geq 0} A_k \rho^k$$

e substituindo na última equação, resulta que os coeficientes A_k devem satisfazer a relação de recorrência

$$A_{k+1} = \frac{k+l+1-\lambda}{(k+1)(k+2l+2)} A_k.$$

Para valores não inteiros de λ a razão dos termos sucessivos da série indica um comportamento exponencialmente crescente do tipo e^ρ , de modo que para obter soluções limitadas é preciso garantir a terminação da série escolhendo $\lambda = n \geq l+1$, com n inteiro. Nesse caso $A_k = 0$ para $k > n-l-1$ e as soluções regulares da equação radial se reduzem a

$$u_{\lambda,l}(\rho) \rightarrow u_{nl}(\rho) = e^{-\rho/2} \rho^{l+1} \sum_{k=0}^{n-l-1} A_k \rho^k, \quad l \leq n-1. \quad (5.35)$$

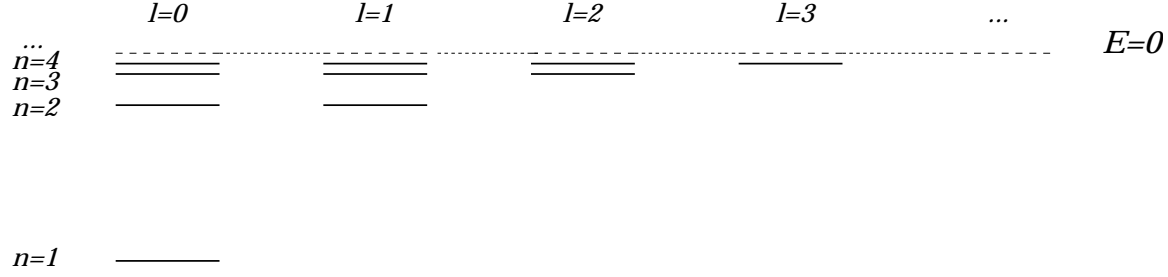


Figura 5.3: Diagrama ilustrativo das energias relativas dos estados ligados do potencial coulombiano, classificados segundo os valores do número quântico n e do momento angular l .

Os polinômios que aparecem nessas soluções são *polinômios de Laguerre associados* e os autovalores discretos do Hamiltoniano resultam da relação (5.34), através das restrições impostas sobre λ . De fato, usando $\lambda = n > l$ nessa relação, e lembrando que $E_r < 0$, segue que

$$E_r(n, l) = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} \equiv -\frac{\mu Z^2}{m_e n^2} R, \quad l < n, \quad n \geq 1 \quad (5.36)$$

onde m_e é a massa do elétron e $R = m_e e^4 / 2\hbar^2$ é a chamada *constante de Rydberg*, cujo valor é $m_e c^2 \times (e^2 / \hbar c)^2 \times 1/2 = 511 \times 10^3 \text{ eV} \times (1/137)^2 \times 1/2 = 13.61 \text{ eV}$. Segundo esse resultado, para cada valor de n existem estados ligados com valores de l dados por $l = 0, 1, \dots, n-1$ cuja energia depende apenas de n , sendo proporcional a n^{-2} (v. figura 5.3). A cada valor de l correspondem ainda os $2l + 1$ autovalores possíveis de L_3 , $-l \leq m \leq +l$. As autofunções que descrevem os estados estacionários da dinâmica relativa do átomo hidrogenóide são portanto

$$\phi_{nlm}(\vec{r}) = c_{nl} \frac{u_{nl}(\rho)}{r} Y_{lm}(\theta, \varphi) \quad (5.37)$$

onde θ e φ são as coordenadas angulares associadas ao vetor relativo \vec{r} , a amplitude radial $u_{nl}(\rho)$ é dada por (5.35) e a constante de normalização c_{nl} é fixada pela condição

$$c_{nl}^2 \int_0^\infty dr u_{nl}^2(\rho) = 1.$$

Essas autofunções devem ainda ser substituídas como último fator na expressão (5.33) para que se obtenha as autofunções do átomo hidrogenóide, visto como sistema de dois corpos, correspondentes aos seus estados estacionários.

5.2.1 Correlações de dois corpos e raios de estados ligados.

Uma propriedade simples desses estados estacionários que pode ser obtida dos resultados acima é o valor médio do módulo da coordenada relativa \vec{r} , que serve para caracterizar o “tamanho” do átomo no estado estacionário considerado. Essa quantidade é, de fato

$$\langle r \rangle_{nl} \equiv \int d\vec{r} \phi_{nlm}^*(\vec{r}) r \phi_{nlm}(\vec{r}) = \frac{\hbar^2}{2\mu Ze^2} [3n^2 - l(l+1)].$$

Ela resulta independente do autovalor $\hbar m$ de L_3 devido ao seu caráter rotacionalmente invariante, como média *do módulo* de \vec{r} . Para o estado fundamental (estado de menor energia, $n = 1$, $l = 0$) ela vale

$$\langle r \rangle_{10} = \frac{3\hbar^2}{2\mu Ze^2} \equiv \frac{3m_e}{2\mu Z} a_0, \quad a_0 \equiv \frac{\hbar^2}{m_e e^2} = 5.2917706(\pm 44) \times 10^{-9} \text{ cm}$$

onde a_0 é chamado *primeiro raio de Bohr*. Para estados excitados, por outro lado, $\langle r \rangle_{nl}$ aumenta rapidamente com n para todos os valores possíveis de l , e $\lim_{n \rightarrow \infty} \langle r \rangle_{nl} = \infty$. Isso significa que existem, em princípio, estados ligados do potencial coulombiano com raio arbitrariamente grande.

Finalmente, é útil identificar essa medida do tamanho do átomo hidrogenóide no contexto da função de onda de dois corpos construída por exemplo, como no caso da (5.33), como autofunção simultânea do momento do centro de massa \vec{P} . Para evitar os problemas técnicos ligados à norma infinita do fator de centro de massa com a normalização do contínuo, é conveniente re-escrever a função de onda de dois corpos utilizando para esse fator condições periódicas de contorno num volume L_3 suficientemente grande. Assim

$$\phi_{E_R, nlm}(\vec{r}_1, \vec{r}_2) = \frac{1}{L^{3/2}} e^{i\vec{K} \cdot \frac{M_1 \vec{r}_1 + M_2 \vec{r}_2}{M_1 + M_2}} \phi_{nlm}(\vec{r}_2 - \vec{r}_1)$$

onde o último fator é obtido re-escrevendo a (5.37) em termos das variáveis apropriadas. O valor médio $\langle r \rangle_{nl}$ é agora dado por

$$\langle r \rangle_{nl} = \frac{1}{L^3} \int d\vec{r}_1 \int d\vec{r}_2 \phi_{E_R, nlm}^*(\vec{r}_1, \vec{r}_2) |\vec{r}_1 - \vec{r}_2| \phi_{E_R, nlm}(\vec{r}_1, \vec{r}_2),$$

como pode ser facilmente verificado mudando as variáveis de integração para as variáveis de centro de massa \vec{R} e relativa \vec{r} , e aparece claramente como uma propriedade *de correlação* entre os dois corpos no estado estacionário $\phi_{E_R, nlm}(\vec{r}_1, \vec{r}_2)$, no qual, em particular, o centro de massa se distribui com densidade de probabilidade uniforme sobre todo o volume de quantização L_3 . Em outras palavras, embora a simetria translacional implementada através da escolha de um estado

estacionário que é autofunção também do momento total tenha o efeito de deslocalizar cada um dos dois corpos considerado separadamente, o tamanho do átomo aparece ainda nesse estado como uma escala de comprimento associada à probabilidade de encontrar um dos corpos em \vec{r}_1 condicionada à presença do segundo corpo em \vec{r}_2 . Probabilidades condicionais desse mesmo tipo são as que devem ser consideradas para identificar propriedades cruciais, por exemplo, para propriedades químicas de sistemas moleculares mais complexos (como a propriedade de uma determinada “forma molecular”) em estados quânticos em que elas parecem ter sido de certa forma dissolvidas pela implementação de simetrias globais do Hamiltoniano.

Os efeitos do tamanho atômico, associados a correlações entre as suas partículas constituintes, se manifestam também numa caracterização *completa*, dentro do contexto quântico, do estado de *uma delas*. Devido à existência de correlações, tal estado não pode ser caracterizado em termos de um vetor de estado, ou de uma função de onda. Para tal fim é introduzida uma *matriz densidade* (ou operador densidade) definida no espaço de configurações da partícula escolhida (e.g. a partícula associada ao índice 1) por

$$\rho(\vec{r}_1, \vec{r}'_1) \equiv \int d\vec{r}_2 \phi_{E_R, nlm}^*(\vec{r}_1, \vec{r}_2) \phi_{E_R, nlm}(\vec{r}'_1, \vec{r}_2).$$

Como pode ser notado, os elementos “diagonais” $\rho(\vec{r}_1, \vec{r}_1)$ dão a densidade de probabilidade correspondente à partícula 1 integrada sobre todo o espaço de configurações da partícula 2. Esta densidade de probabilidade é insuficiente para o estudo de efeitos, vistos através da partícula 1, associados a correlações entre as partículas 1 e 2, mas tais efeitos podem ser estudados a partir da matriz densidade completa, incluindo seus elementos não diagonais. Uma discussão mais completa dessas questões pode ser encontrada em P. R. I. Tommasini, E. Timmermans e A. F. R. de Toledo Piza, *Am. J. Phys.* **66**, 881 (1998).

5.2.2 Propriedades empíricas do espectro do átomo de hidrogênio e insuficiência do modelo simples.

O modelo simples utilizado até aqui para o átomo de hidrogênio, que consiste em descreve-lo como sistema constituído de um próton e um elétron vistos como partículas puntiformes e sem estrutura, ligadas pela interação coulombiana, não é suficientemente rico para dar conta, sequer qualitativamente, das propriedades mais finas do seu espectro experimental. Essa descrição reproduz corretamente propriedades desse espectro apenas na escala da constante de Rydberg

$$R = \frac{m_e e^4}{2\hbar^2} = \frac{1}{2} \left(\frac{e^2}{\hbar c} \right)^2 m_e c^2 \equiv \frac{1}{2} \alpha^2 m_e c^2 = \frac{1}{2} \frac{1}{137^2} \times 510976 \text{ eV} = 13.61 \text{ eV}$$

como é, por exemplo, a série de Balmer, já reproduzida por Bohr no contexto da velha teoria quântica. O estudo experimental mais preciso do espectro do hidrogênio revela porém outras propriedades. Por exemplo, o nível $n = 2$, $l = 1$ ou, em termos da nomenclatura espectroscópica, o nível $2p$, que no modelo simples é exatamente degenerado com o nível $2s$ ($n = 2$, $l = 0$), aparece na realidade como *um par* de níveis cuja separação é da ordem de $\alpha^2 R$, o nível inferior sendo essencialmente degenerado com o nível $2s$. Essa “estrutura fina” (notando que $\alpha^2 \sim 5 \times 10^{-5}$) aparece também no caso de outros níveis com $l \neq 0$. Numa escala de energia que é cerca de 10^{-3} a escala dessa estrutura fina é possível observar que o próprio estado fundamental $1s$, não degenerado no modelo simples, aparece na realidade também como um dubleto (“estrutura hiperfina”), do qual um dos níveis se separa ainda em três níveis diferentes sob a ação de um campo magnético externo.

A descrição desses efeitos, particularmente a presença de vários níveis “hiperfinos” onde o modelo simples prevê um único estado não degenerado, exige a introdução de outros graus de liberdade, além da posição do próton e do elétron, num modelo mais realístico do átomo de hidrogênio. Tanto a estrutura fina como a estrutura hiperfina do espectro podem, de fato, ser descritas em termos relativamente simples admitindo a presença de graus de liberdade adicionais tanto no elétron como no próton com propriedades análogas às do momento angular e com valores semi-inteiros em unidades de \hbar (spin). Isso será feito no capítulo seguinte.

Em 1947 Lamb e Retherford (*Phys. Rev.* **72**, 241 (1947)) verificaram que a componente inferior do dubleto de estrutura fina $2p$ se encontra na realidade a uma energia $h\nu$, com $\nu \sim 10^9 \text{ s}^{-1} \simeq 3 \times 10^{-7} R$, *abaixo* do estado $2s$. Esse “deslocamento de Lamb” requer, para o seu tratamento, a consideração de um número infinito de graus de liberdade associados às propriedades eletromagnéticas do vácuo.

Problemas.

5.1 - Seja \vec{L} um operador de momento angular e $|lm\rangle$ os autovetores normalizados simultâneos de L^2 e de L_z , com autovalores $\hbar^2 l(l+1)$ e $\hbar m$ respectivamente. Calcule

a) $\langle l'm'|L_x|lm\rangle$

b) $\langle l'm'|L_x^2|lm\rangle$

c) $\langle l'm'|(L_x^2 + L_y^2)|lm\rangle$

d) $\langle l'm'|(L_x^2 - L_y^2)|lm\rangle$.

5.2 - Sejam a e b dois operadores não hermiteanos, satisfazendo com os seus respectivos adjuntos a^\dagger e b^\dagger as relações de comutação

$$[a, a^\dagger] = [b, b^\dagger] = 1; \quad [a, b] = [a, b^\dagger] = [a^\dagger, b] = [a^\dagger, b^\dagger] = 0.$$

Mostre então que os operadores

$$J_+ = a^\dagger b; \quad J_- = J_+^\dagger = b^\dagger a \quad \text{e} \quad J_3 = \frac{1}{2} (a^\dagger a - b^\dagger b)$$

satisfazem relações de comutação de momento angular. Construa o operador correspondente a J^2 também em termos de a , b e de seus adjuntos e verifique que ele comuta com J_\pm e com J_3 .

Esses operadores foram utilizados por Schwinger para o tratamento do momento angular quântico (J. Schwinger, *Quantum Theory of Angular Momentum*, L. Biedenharn e H. Van Dam, Eds., Academic Press, N.Y. 1965.)

5.3 - a) Determine, para o caso de um poço de potencial quadrado, em três dimensões

$$V(r) = \begin{cases} -V_0, & r < R \\ 0, & r > R \end{cases}$$

as condições sobre os parâmetros (raio R e profundidade V_0 do potencial) para que

a) existam n estados ligados de momento angular $l = 0$. Relacione n com o número de zeros no intervalo $0 < r < R$ da função de onda radial calculada para energia zero;

b) exista pelo menos um estado ligado de momento angular $l = 1$.

c) Quantos estados ligados de momento angular $l = 0$ haverá quando houver *apenas um* estado ligado de momento angular $l = 1$?

d) Se a profundidade V_0 do potencial aumenta à medida que o raio R diminui, de forma que a integral de volume do potencial $-4\pi V_0 R^3/3$ permanece constante, o número de estados ligados com $l = 0$ aumenta, não se altera ou diminui à medida que o raio do potencial diminui? Qual o limite para a energia do estado fundamental quando $R \rightarrow 0$? Qual é, nesse caso, o limite para a função de onda do estado fundamental?

5.4 - a) Mostre que o Hamiltoniano de oscilador harmônico tri-dimensional e isotrópico

$$H = \frac{p^2}{2M} + \frac{M\omega_0^2}{2} r^2$$

comuta com as componentes do momento angular $\vec{l} = \vec{r} \times \vec{p}$.

b) Utilizando a representação de Schrödinger e coordenadas cartesianas mostre que esse Hamiltoniano pode ser escrito como a soma de três Hamiltonianos de oscilador unidimensionais, $H = H_1 + H_2 + H_3$ que são compatíveis e portanto podem ser diagonalizados simultaneamente.

c) Mostre que cada um dos hamiltonianos unidimensionais H_i do item anterior, correspondentes às componentes cartesianas das variáveis dinâmicas, comuta com a componente respectiva L_i do momento angular mas não comuta com as outras duas componentes, e portanto não comuta também com L^2 .

d) Trate o problema da determinação das autofunções simultâneas de H , de L^2 e da componente L_3 do momento angular. Para isso considere o problema de autovalores na representação de Schrödinger em coordenadas esféricas, determine a forma das autofunções simultâneas dos três observáveis e obtenha a equação radial a ser resolvida nesse caso.

e) Mostre que L^2 e L_3 têm elementos de matriz nulos entre autofunções de H com energias diferentes, e que portanto a degenerescência de qualquer autovalor de H deve ser um número que pode ser escrito como $2l + 1$ ou como uma soma de números dessa forma.

f) Mostre, a partir da solução em coordenadas cartesianas, que o estado fundamental não é degenerado, e que isso implica que a autofunção de H correspondente a esse estado deve ser também uma autofunção do momento angular. Qual é o valor do momento angular para esse estado?

g) Considere o oscilador *anisotrópico*, mas com simetria axial, descrito por

$$H = \frac{p^2}{2M} + \frac{M}{2} [\omega_{\perp}^2 x_3^2 + \omega_{\parallel}^2 (x_1^2 + x_2^2)], \quad \omega_{\perp} \neq \omega_{\parallel}.$$

Mostre que nesse caso L_3 é uma constante do movimento, mas o mesmo não ocorre com L_1 e L_2 . O operador L^2 é uma constante do movimento? Desenvolva um esquema para calcular as autofunções simultâneas de H e L_3 usando a representação de Schrödinger em coordenadas cilíndricas, $\vec{r} \equiv \{\rho, \varphi, x_3\}$.

5.5 - a) Deduza uma expressão para a defasagem δ_l da onda parcial l para o espalhamento por uma esfera dura, isto é, um potencial repulsivo $V_0 \rightarrow \infty$ de perfil quadrado e raio R :

$$V(r) = \begin{cases} V_0 \rightarrow \infty, & r < R \\ 0 & r > R \end{cases}.$$

b) Ache a forma assintótica de δ_l no limite em que a energia da partícula espalhada se aproxima de zero.

c) Expandir $k \cot \delta_0$ em potências de k , e identifique o comprimento de espalhamento e o alcance efetivo em termos do raio R da esfera dura.

5.6 - Mostre que a seção de choque total para o espalhamento por uma esfera dura de raio R tende a $2\pi R^2$ para $kR \gg 1$. Explique a diferença entre esse resultado e a seção de choque clássica πR^2 . **Sugestões:** **i)** as ondas parciais que contribuem apreciavelmente para a seção de choque são as correspondentes aos valores $l \leq kR$ do momento angular; **ii)** para $kR \gg 1$ é possível usar as expressões assintóticas para argumento grande das funções de Bessel e de Neumann esféricas:

$$j_l(\rho) \rightarrow -\frac{1}{\rho} \sin\left(\rho - \frac{l\pi}{2}\right), \quad n_l(\rho) \rightarrow \frac{1}{\rho} \cos\left(\rho - \frac{l\pi}{2}\right).$$

iii) termos proporcionais a kR são desprezíveis frente a termos proporcionais a $k^2 R^2$.

5.7 - a) Um elétron está ligado ao campo coulombiano de um núcleo de carga positiva Ze . Calcule $\sqrt{\langle v^2/c^2 \rangle}$ para o estado fundamental, sendo v a velocidade do elétron. A função de onda normalizada para o estado fundamental desse sistema é

$$\phi_{1s}(\vec{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0}$$

onde $a_0 = \hbar^2/\mu e^2$, sendo $\mu = m_e M/(m_e + M)$ a massa reduzida do sistema elétron-núcleo.

b) Supondo que o núcleo tenha um raio finito R e que a distribuição de carga nuclear seja uniforme, calcule o efeito do raio finito da distribuição de carga sobre a energia do estado fundamental usando teoria de perturbação de primeira ordem.

c) Repita os cálculos dos dois itens anteriores para o caso de um átomo de muônico, no qual o elétron é substituído por um muon, também de carga $-e$ mas com massa $m_\mu = 207 m_e$. Qual é, em termos da energia de ligação do átomo hidrogenóide, a energia de ligação do átomo muônico no estado fundamental? Qual é o tamanho do átomo muônico nesse estado?

d) Considere em seguida as mesmas questões do item **c)** para o caso do positrônio, visto como sistema constituído de um pósitron (massa m_e , carga $+e$) e de um elétron.

5.8 - a) Mostre que, para os estados ligados de um átomo hidrogenóide com $n = l + 1$ (que são os estados de menor energia para o valor considerado do momento angular ou, alternativamente, os estados de maior momento angular para a energia de ligação considerada), os polinômios que aparecem nas funções radiais (5.35) se reduzem a uma constante A_0 .

b) Determine o valor dessa constante normalizando as funções de onda $\phi_{n=l+1,lm}(\vec{r})$ de modo que

$$\int d\vec{r} |\phi_{n=l+1,lm}(\vec{r})|^2 = 1.$$

c) Ache o valor de r em que a densidade de probabilidade integrada sobre os ângulos

$$\rho(r) = \int_{-1}^{+1} d \cos \theta \int_0^{2\pi} |\phi_{n=l+1,lm}(\vec{r})|^2$$

é máxima.

d) Compare o resultado de **c)** com os raios das órbitas circulares do átomo *clássico*.

Capítulo 6

Estrutura fina e campos externos no átomo de hidrogênio.

A estrutura fina e a estrutura hiperfina observada nos estados estacionários do átomo de hidrogênio exigem uma ampliação da cinemática utilizada no modelo simples, baseado em analogias clássicas, tratado no capítulo anterior. Isso é particularmente claro no caso da estrutura hiperfina do estado fundamental. No modelo simples, esse estado é um estado $1s$, não degenerado, e a existência de uma estrutura envolvendo *quatro* estados exige a introdução de graus de liberdade adicionais. Tratando-se de um sistema de dois corpos, é possível conjecturar que tanto o próton quanto o elétron sejam objetos que, além dos graus de liberdade associados à posição, têm propriedades intrínsecas associadas a variáveis dinâmicas que agem num espaço de estados com *duas* dimensões.

Essa idéia pode ser implementada em termos de graus de liberdade do tipo do momento angular, como se pode ver utilizando as propriedades gerais obtidas na seção **5.1.1**, que decorrem diretamente das relações de comutação satisfeitas por suas componentes. Se o momento angular a ser considerado é associado a um operador vetorial $\vec{S} \equiv \{S_1, S_2, S_3\}$ cujas componentes satisfazem as relações de comutação características desse tipo de variável dinâmica

$$[S_j, S_k] = i\hbar\epsilon_{jkl}S_l \quad (6.1)$$

onde ϵ_{jkl} é o símbolo antissimétrico, então os autovalores de S^2 são da forma $\hbar^2 s(s+1)$ onde s é inteiro ou *semi-inteiro*. O número de estados correspondentes a um dado valor de s é $2s+1$, que se reduz a dois para $s = 1/2$. Associando então tanto ao próton como ao elétron um *momento angular intrínseco* (ou *spin*) com $s = 1/2$, o espaço de estados do átomo de hidrogênio passa a ser o espaço produto do que corresponde ao modelo simples pelo espaço de quatro dimensões associado a esses novos graus de liberdade. Na medida em que o Hamiltoniano

seja independente dos spins, cada um dos estados estacionários do modelo simples dá origem a um grupo de quatro estados independentes (que é conveniente tomar como sendo ortogonais) e degenerados, que diferem entre si apenas pelo estado dos dois momentos angulares intrínsecos.

Uma vez feita essa opção cinemática, será preciso ainda determinar a forma da dependência do hamiltoniano com as variáveis de spin que possa dar conta das propriedades finas e hiperfinas do espectro observado do átomo de hidrogênio. Uma observação importante, nesse contexto, é a de que o momento angular e momentos de dipolo magnético são quantidades com propriedades de transformação análogas no espaço de três dimensões. Desse modo, a introdução do spin possibilita associar às partículas momentos magnéticos intrínsecos. Os termos dependentes dos spins que devem ser acrescentados ao hamiltoniano são identificáveis como as energias de interação mútua desses momentos magnéticos, a sua energia de interação com os campos eletromagnéticos associadas à estrutura orbital do átomo ou ainda com campos eletromagnéticos externos. Os hamiltonianos resultantes, com os termos de interação dependentes dos spins, não se prestam facilmente a tratamentos exatos, e os efeitos devidos a esses termos são exemplos importantes do uso de métodos perturbativos.

6.1 Realização explícita das variáveis de spin. Álgebra das matrizes da Pauli.

De forma mais específica, a uma partícula com spin $1/2$ estão associadas, além das variáveis dinâmicas com análogo clássico (isto é, posição \vec{q} e momento \vec{p}), também variáveis de spin \vec{S} que devem ser realizadas em termos de operadores hermiteanos agindo em um espaço complexo adicional de duas dimensões, satisfazendo as relações de comutação (6.1). Esta seção discute apenas as propriedades das variáveis de spin e de seus vetores de estado. A inclusão das variáveis dinâmicas com análogo clássico será feita na seção 6.2 abaixo.

Os vetores do espaço bi-dimensional das variáveis de spin podem ser expressos em termos da base formada pelos autovetores simultâneos de S^2 e de S_3

$$S^2|\frac{1}{2}m_s\rangle = \hbar^2\frac{1}{2}(\frac{1}{2} + 1)|\frac{1}{2}m_s\rangle, \quad S_3|\frac{1}{2}m_s\rangle = \hbar m_s|\frac{1}{2}m_s\rangle, \quad \langle\frac{1}{2}m_s|\frac{1}{2}m'_s\rangle = \delta_{m_sm'_s} \quad (6.2)$$

com $m_s, m'_s = \pm 1/2$. Em termos dessa base, um vetor genérico do espaço do spin $1/2$ pode ser escrito, a menos de um fator de fase global, sob a forma da combinação linear normalizada

$$|\alpha, \beta\rangle = \cos \alpha |\frac{1}{2} \frac{1}{2}\rangle + e^{i\beta} \sin \alpha |\frac{1}{2} - \frac{1}{2}\rangle.$$

Alternativamente, o vetor de spin $|\alpha, \beta\rangle$ pode ser representado em termos de suas componentes nessa base pela coluna de duas entradas

$$\begin{pmatrix} \langle \frac{1}{2} \frac{1}{2} | \alpha, \beta \rangle \\ \langle \frac{1}{2} - \frac{1}{2} | \alpha, \beta \rangle \end{pmatrix} = \begin{pmatrix} \cos \alpha \\ e^{i\beta} \sin \alpha \end{pmatrix}. \quad (6.3)$$

Usando esta última notação, um operador linear g qualquer desse espaço é representado pela matriz 2×2

$$g_{m_s m'_s} \equiv \langle \frac{1}{2} m_s | g | \frac{1}{2} m'_s \rangle, \quad m_s, m'_s = \pm \frac{1}{2}.$$

Em particular, é imediato verificar que os operadores S^2 e S_3 correspondem às matrizes

$$S^2 \rightarrow \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{e} \quad S_3 \rightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Para determinar as matrizes que correspondem a S_1 e a S_2 basta considerar os operadores $S_{\pm} = S_1 \pm iS_2$ e as suas propriedades, decorrentes das relações de comutação (6.1),

$$S_+ |\frac{1}{2} \frac{1}{2}\rangle = 0; \quad S_+ |\frac{1}{2} - \frac{1}{2}\rangle = \hbar |\frac{1}{2} \frac{1}{2}\rangle; \quad S_- |\frac{1}{2} \frac{1}{2}\rangle = \hbar |\frac{1}{2} - \frac{1}{2}\rangle; \quad S_- |\frac{1}{2} - \frac{1}{2}\rangle = 0.$$

Tomando o produto escalar dessas relações com cada um dos dois vetores da base,

$$S_+ \rightarrow \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad S_- \rightarrow \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Exprimindo agora S_1 e S_2 em termos dessas matrizes resulta

$$S_1 \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{e} \quad S_2 \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

O operador S^2 é portanto representado como um múltiplo do operador unidade no espaço bidimensional, ao passo que as componentes de \vec{S} são representadas, na base dos autovetores de S_3 , por $\vec{S} = (\hbar/2)\vec{\sigma}$, onde $\vec{\sigma}$ corresponde às matrizes

$$\vec{\sigma} \equiv \{\sigma_1, \sigma_2, \sigma_3\} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}.$$

Essas matrizes são chamadas *matrizes de Pauli*. É fácil verificar diretamente que elas satisfazem as relações de comutação

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l \quad (6.4)$$

que levam diretamente às relações de comutação (6.1) para as componentes de \vec{S} . Para verificar que $S^2 = S_1^2 + S_2^2 + S_3^2$ de fato reproduz o múltiplo da unidade obtido acima, basta notar que cada uma das três matrizes de Pauli tem como autovalores $+1$ e -1 , de modo que

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \hat{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad (6.5)$$

Como $\vec{S} = (\hbar/2)\vec{\sigma}$, segue que $S^2 = (3\hbar^2/4)\hat{1}$.

Uma outra propriedade das matrizes de Pauli que pode ser obtida do resultado (6.5) é a de que quaisquer duas matrizes diferentes *anti-comutam*, isto é, $\sigma_j\sigma_k + \sigma_k\sigma_j = 0$ para $j \neq k$. De fato, como $\sigma_j^2 = \hat{1}$, é claro que $[\sigma_j^2, \sigma_k] = 0$ para qualquer k . Por outro lado, tomando por exemplo $j = 1$ e $k = 2$ e usando as relações de comutação (6.4) resulta

$$0 = [\sigma_1^2, \sigma_2] = \sigma_1[\sigma_1, \sigma_2] + [\sigma_1, \sigma_2]\sigma_1 = 2i(\sigma_1\sigma_3 + \sigma_3\sigma_1)$$

donde se conclui que σ_1 anti-comuta com σ_3 . Outras escolhas de j e k mostram, de forma análoga, que

$$\sigma_1\sigma_2 + \sigma_2\sigma_1 = 0 \quad \text{e} \quad \sigma_2\sigma_3 + \sigma_3\sigma_2 = 0.$$

Essa propriedade pode ser expressa simultaneamente com a (6.5) através da relação única

$$\sigma_j\sigma_k + \sigma_k\sigma_j = 2\delta_{jk}. \quad (6.6)$$

Uma consequência importante desta última relação é que, se \hat{n} é um vetor unitário qualquer, então $(\vec{\sigma} \cdot \hat{n})^2 = \hat{1}$. De fato,

$$\begin{aligned} (\sigma_1 n_1 + \sigma_2 n_2 + \sigma_3 n_3)^2 &= (\sigma_1^2 n_1^2 + \sigma_2^2 n_2^2 + \sigma_3^2 n_3^2) + \\ &\quad (\sigma_1\sigma_2 + \sigma_2\sigma_1) n_1 n_2 + (\sigma_2\sigma_3 + \sigma_3\sigma_2) n_2 n_3 + (\sigma_3\sigma_1 + \sigma_1\sigma_3) n_1 n_3 \\ &= \hat{1} \hat{n} \cdot \hat{n} = \hat{1}. \end{aligned}$$

Esta relação permite calcular facilmente operadores envolvendo potências de $\vec{\sigma} \cdot \hat{n}$, já que potências pares se reduzem simplesmente a $\hat{1}$ e potências ímpares a $\vec{\sigma} \cdot \hat{n}$. Um exemplo importante do uso desse fato é o cálculo de exponenciais de matrizes de Pauli como o operador *unitário*

$$e^{i\theta(\vec{\sigma} \cdot \hat{n})} = \hat{1} \cos \theta + i(\vec{\sigma} \cdot \hat{n}) \sin \theta$$

para θ real. Essa relação pode ser verificada facilmente expandindo a exponencial em série de potências e colecionando separadamente potências pares e ímpares de $\vec{\sigma} \cdot \hat{n}$. A unitariedade desse operador exponencial decorre da hermiticidade do operador $\vec{\sigma} \cdot \hat{n}$.

6.1.1 Direção de polarização do vetor de estado geral de spin 1/2.

O vetor de estado geral (6.3) para um spin 1/2 pode sempre ser visto como autovetor da projeção do operador vetorial \vec{S} numa determinada direção, definida através de um vetor unitário \hat{p} . Isso significa que, para qualquer estado $|\alpha, \beta\rangle$ é sempre possível determinar um vetor unitário \hat{p} tal que

$$\vec{S} \cdot \hat{p} |\alpha, \beta\rangle = \frac{\hbar}{2} |\alpha, \beta\rangle.$$

De fato, escrevendo as componentes cartesianas de \hat{p} em termos de ângulos esféricos θ e φ

$$\hat{p} \equiv \{p_1, p_2, p_3\} = \{\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta\}$$

o operador $\vec{S} \cdot \hat{p}$ pode ser escrito explicitamente como

$$\vec{S} \cdot \hat{p} = \frac{\hbar}{2} \begin{pmatrix} \cos \theta & e^{-i\varphi} \sin \theta \\ e^{i\varphi} \sin \theta & -\cos \theta \end{pmatrix}$$

e a condição de que $|\alpha, \beta\rangle$ seja autovetor desse operador com autovalor $\hbar/2$ é

$$\frac{\hbar}{2} \begin{pmatrix} \cos \theta & e^{-i\varphi} \sin \theta \\ e^{i\varphi} \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} \cos \alpha \\ e^{i\beta} \sin \alpha \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \cos \alpha \\ e^{i\beta} \sin \alpha \end{pmatrix}$$

ou, equivalentemente,

$$\begin{aligned} \cos \theta \cos \alpha + e^{i(\beta-\varphi)} \sin \theta \sin \alpha &= \cos \alpha, \\ e^{-i(\beta-\varphi)} \sin \theta \cos \alpha - \cos \theta \sin \alpha &= \sin \alpha. \end{aligned}$$

Escolhendo $\varphi = \beta$ e dividindo a primeira equação por $\cos \alpha$ e a segunda por $\sin \alpha$

$$\cos \theta + \sin \theta \tan \alpha = 1; \quad \text{e} \quad \sin \theta \cot \alpha - \cos \theta = 1.$$

A soma dessas duas equações permite eliminar $\cos \theta$ e dá

$$\sin \theta (\tan \alpha + \cot \alpha) = 2 \quad \text{ou} \quad \sin \theta = 2 \sin \alpha \cos \alpha = \sin 2\alpha$$

que leva a $\theta = 2\alpha$. O estado (6.3) é portanto autovetor da projeção de \vec{S} sobre a direção definida pelos ângulos esféricos $\theta = 2\alpha$ e $\varphi = \beta$. Essa direção é a *direção de polarização* do spin caracterizado por esse vetor de estado, e é fácil verificar que ela pode também ser obtida

calculando os valores médios das componentes cartesianas do operador de spin. De fato, um cálculo direto mostra que

$$\langle \alpha, \beta | \vec{S} | \alpha, \beta \rangle = \frac{\hbar}{2} \{ \sin 2\alpha \cos \beta, \sin 2\alpha \sin \beta, \cos 2\alpha \} \equiv \frac{\hbar}{2} \hat{p}.$$

O vetor \hat{p} é freqüentemente chamado *vetor de Bloch* associado ao estado de spin (6.3). O argumento acima mostra que, a menos de um fator de fase global que não contribui para o valor médio de \vec{S} , ele caracteriza completamente o estado de spin ao qual está associado.

Uma primeira consequência direta desse resultado é a de que o vetor de estado que representa um spin 1/2 polarizado numa direção agora *dada* pelos ângulos esféricos θ e φ (isto é, o estado de spin cujo vetor de Bloch é dado pelos ângulos esféricos θ e φ) é, usando a representação (6.3),

$$\begin{pmatrix} \cos \theta/2 \\ e^{i\varphi} \sin \theta/2 \end{pmatrix} \equiv |\theta/2, \varphi\rangle$$

cujas componentes, na base dos autovetores de S_3 , são

$$\langle \frac{1}{2} \frac{1}{2} | \theta/2, \varphi \rangle = \cos \frac{\theta}{2} \quad \text{e} \quad \langle \frac{1}{2} - \frac{1}{2} | \theta/2, \varphi \rangle = e^{i\varphi} \sin \frac{\theta}{2}.$$

Portanto, a probabilidade de obter o valor $+\hbar/2$ numa medida de S_3 nesse estado é $\cos^2(\theta/2)$, que somente se anula para $\theta = \pi$, isto é, se a direção de polarização (ou do vetor de Bloch) é *oposta* ao eixo 3; e, de forma semelhante, a probabilidade $\sin^2(\theta/2)$ de obter o valor $-\hbar/2$ somente se anula se $\theta = 0$, isto é, se o spin está polarizado na direção 3 (ou tem o vetor de Bloch na direção 3).

Uma segunda consequência importante é a possibilidade de verificar explicitamente que operadores unitários do tipo

$$R_{1/2}(\gamma) = e^{-\frac{i}{\hbar} \gamma \vec{S} \cdot \hat{n}} = e^{-\frac{i}{2} \gamma \vec{\sigma} \cdot \hat{n}} = \hat{1} \cos \frac{\gamma}{2} - i \vec{\sigma} \cdot \hat{n} \sin \frac{\gamma}{2}$$

aplicados a um vetor de estado de spin dão como resultado o vetor de estado com a direção de polarização rodada de um ângulo γ em torno do eixo definido pelo vetor \hat{n} . De fato, tomando o eixo 3 na direção de \hat{n} e calculando o vetor de Bloch

$$\frac{\hbar}{2} \hat{p}' \equiv \langle \alpha, \beta | R_{1/2}^\dagger(\gamma) \vec{S} R_{1/2}(\gamma) | \alpha, \beta \rangle$$

resulta facilmente

$$\vec{p}' = \{ \sin 2\alpha \cos(\beta + \gamma), \sin 2\alpha \sin(\beta + \gamma), \cos 2\alpha \}$$

que corresponde ao resultado mencionado. Como o vetor de Bloch é insensível a um possível fator de fase global do estado de spin, é interessante examinar explicitamente o efeito desse operador de rotação sobre o próprio vetor de estado que descreve o spin. Nas mesmas condições usadas para o último cálculo o que se obtém para o novo vetor de estado é

$$R_{1/2}(\gamma)|\alpha, \beta\rangle = \begin{pmatrix} e^{-i\frac{\gamma}{2}} & 0 \\ 0 & e^{i\frac{\gamma}{2}} \end{pmatrix} \begin{pmatrix} \cos \alpha \\ e^{i\beta} \sin \alpha \end{pmatrix} = e^{-i\frac{\gamma}{2}} \begin{pmatrix} \cos \alpha \\ e^{i(\beta+\gamma)} \sin \alpha \end{pmatrix}.$$

O operador de rotação introduz, portanto um fator de fase global $e^{-i\frac{\gamma}{2}}$ além de girar a direção de polarização do estado de spin. Esse resultado mostra o efeito peculiar de que uma rotação de um ângulo 2π multiplica o vetor de estado por -1 , o que é claramente uma consequência do valor semi-inteiro do spin. Isso evidentemente não afeta os valores esperados, e portanto não aparece no cálculo do vetor de Bloch. A presença desse fator de fase pode no entanto ser confirmada experimentalmente através dos processos de interferência que produz. Tipicamente, a propagação de um feixe de partículas de spin $1/2$ entre dois pontos A e B por dois caminhos diferentes (como numa experiência de difração) mostrará interferência *destrutiva* em B cada vez que os spins sofrerem uma rotação de 2π ao longo de *um* dos dois caminhos¹.

6.1.2 Matrizes de Pauli como base de operadores no espaço de spin.

Uma observação simples mas tecnicamente importante é a de que *qualquer* matriz 2×2 pode ser expressa como uma combinação linear das tres matrizes de Pauli $\vec{\sigma}$ e da matriz unidade de dimensão 2, $\hat{1}$. A forma mais simples de verificar isso é notar que

$$\begin{aligned} a_{++} &\equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\hat{1} + \sigma_3); \\ a_{+-} &\equiv \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_1 + i\sigma_2); \\ a_{-+} &\equiv \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_1 - i\sigma_2) \quad \text{e} \\ a_{--} &\equiv \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}(\hat{1} - \sigma_3). \end{aligned}$$

¹v. H. Rauch et al., *Phys. Lett.* **54A**, 425 (1975) e também S. A. Werner et al., *Phys. Rev. Lett.* **35**, 1053 (1975).

Como, por outro lado,

$$M \equiv \begin{pmatrix} m_{++} & m_{+-} \\ m_{-+} & m_{--} \end{pmatrix} = m_{++}a_{++} + m_{+-}a_{+-} + m_{-+}a_{-+} + m_{--}a_{--}$$

é claro que

$$\begin{aligned} M &= \frac{1}{2}(m_{+-} + m_{-+})\sigma_1 + \frac{i}{2}(m_{+-} - m_{-+})\sigma_2 + \frac{1}{2}(m_{++} - m_{--})\sigma_3 + \frac{1}{2}(m_{++} + m_{--})\hat{1} \\ &\equiv M_1\sigma_1 + M_2\sigma_2 + M_3\sigma_3 + M_0\hat{1}. \end{aligned}$$

A utilidade desse tipo de decomposição está ligada ao fato de que ela permite controlar facilmente as propriedades de transformação de M por uma mudança do sistema de coordenadas espaciais $\{x_1, x_2, x_3\}$. De fato, as componentes de $\vec{\sigma}$, da mesma forma que as do operador de spin \vec{S} que lhes são proporcionais, têm propriedades de transformação iguais às das componentes do momento angular orbital $\vec{L} = \vec{r} \times \vec{p}$, que são propriedades de transformação de um *vetor axial*, isto é, se transformam como as componentes de um vetor sob translações e rotações do sistema de coordenadas, mas trocam de sinal por uma inversão espacial, $\{x_1, x_2, x_3\} \rightarrow \{-x_1, -x_2, -x_3\}$. Desse modo, para que M seja, por exemplo, um *escalar* é preciso que $\{M_1, M_2, M_3\}$ se comportem também como componentes de um vetor axial e que M_0 seja um escalar; e se $\vec{M} \equiv \{M_1, M_2, M_3\}$ se comportarem como as componentes de um vetor (polar), então $\vec{M} \cdot \vec{\sigma}$ será um pseudo-escalar, que troca de sinal som uma inversão do sistema de coordenadas.

Para obter os coeficientes \vec{M} , M_0 que expandem uma matriz dada M é possível utilizar as relações (6.6) bem como a invariância cíclica do traço (v. Problema 6.1) que levam às relações

$$M_0 = \frac{1}{2}\text{Tr } M, \quad M_j = \frac{1}{2}\text{Tr } (M\sigma_j), \quad j = 1, 2, 3.$$

6.2 Estados de uma partícula com spin 1/2.

Uma partícula quântica com spin 1/2 é um objeto que, além dos graus de liberdade “internos” associados às variáveis de spin, deve ser caracterizado também pelos graus de liberdade associados às variáveis dinâmicas com análogo clássico, posição e momento. O espaço dos vetores de estado de uma tal partícula pode ser construído de forma semelhante à que foi usada na seção 5.2 para o espaço de duas partículas a partir do espaço de estados de uma só partícula. Ela consiste em construir o espaço de vetores de estado da partícula com spin como o produto do espaço que corresponde à partícula *sem spin* pelo espaço bi-dimensional dos estados de spin 1/2, discutido na seção anterior.

A forma mais simples de fazer isso consiste em construir uma base para o espaço produto a partir de bases de cada um dos espaços fatores. Desse modo, se $\{|\phi_n\rangle\}$ é uma base ortonormal do espaço que corresponde a uma partícula sem spin, e utilizando a base (6.2) de autovetores de S_3 para o espaço de spin, uma base ortonormal do espaço produto que funciona como espaço de fases” para a partícula com spin é formada pelos vetores

$$\left\{ \left| n \frac{1}{2} m_s \right\rangle \right\} \equiv \left\{ |\phi_n\rangle \otimes \left| \frac{1}{2} m_s \right\rangle \right\}. \quad (6.7)$$

Cada vetor $|\phi_n\rangle$ dá portanto lugar a *dois* vetores da base produto, que contém respectivamente os dois autovetores possíveis de S_3 , associados aos autovalores $\pm \hbar/2$. Quando se utiliza a notação (6.3) para os vetores de spin, esses vetores de base podem ser escritos como

$$\left| n \frac{1}{2} \frac{1}{2} \right\rangle \rightarrow \begin{pmatrix} |\phi_n\rangle \\ 0 \end{pmatrix} \quad \text{e} \quad \left| n \frac{1}{2} - \frac{1}{2} \right\rangle \rightarrow \begin{pmatrix} 0 \\ |\phi_n\rangle \end{pmatrix}.$$

Um vetor geral $|\psi\rangle$ do espaço produto (e portanto um vetor de estado geral para a partícula com spin) pode ser escrito como uma combinação linear dos vetores de base, o que corresponde a

$$\begin{aligned} |\psi\rangle &= \sum_n \sum_{m_s=-1/2}^{1/2} c_{nm_s} \left| n \frac{1}{2} m_s \right\rangle \rightarrow \begin{pmatrix} |\psi^{(+)}\rangle \\ |\psi^{(-)}\rangle \end{pmatrix} = \sum_n \left[c_{n \ 1/2} \begin{pmatrix} |\phi_n\rangle \\ 0 \end{pmatrix} + c_{n \ -1/2} \begin{pmatrix} 0 \\ |\phi_n\rangle \end{pmatrix} \right] \\ &= \begin{pmatrix} \sum_n c_{n \ 1/2} |\phi_n\rangle \\ \sum_n c_{n \ -1/2} |\phi_n\rangle \end{pmatrix} \end{aligned}$$

Nessa expressão $|\psi^{(\pm)}\rangle$ são vetores do espaço da partícula sem spin, o que mostra que o estado de uma partícula com spin $1/2$ pode ser dado como um par de vetores “sem spin” correspondentes a duas componentes ortogonais no espaço bi-dimensional do spin (no caso as componentes correspondentes aos dois autovetores de S_3). O produto escalar de dois vetores $|\psi_i\rangle$, $i = 1, 2$, com coeficientes de expansão $c_{nm_s}^{(i)}$ na base produto (6.7), é dado por

$$\langle \psi_2 | \psi_1 \rangle = \sum_{nn'} \sum_{m_s m'_s} c_{n'm'_s}^{(2)*} c_{nm_s}^{(1)} \langle n' \frac{1}{2} m'_s | n \frac{1}{2} m_s \rangle = \sum_n \sum_{m_s} c_{nm_s}^{(2)*} c_{nm_s}^{(1)}$$

ou, na notação alternativa,

$$\langle \psi_2 | \psi_1 \rangle = \begin{pmatrix} \langle \psi_2^{(+)} | & \langle \psi_2^{(-)} | \end{pmatrix} \begin{pmatrix} |\psi_1^{(+)}\rangle \\ |\psi_1^{(-)}\rangle \end{pmatrix} = \langle \psi_2^{(+)} | \psi_1^{(+)} \rangle + \langle \psi_2^{(-)} | \psi_1^{(-)} \rangle$$

que leva evidentemente ao mesmo resultado em termos dos coeficientes de expansão $c_{nm_s}^{(i)}$.

Operadores lineares g agindo no espaço produto, por outro lado, são representados na base produto em termos de matrizes construídas da forma usual, isto é $g \rightarrow \{G_{nm_s, n'm'_s}\}$ com

$$G_{nm_s, n'm'_s} = \langle n \frac{1}{2} m_s | g | n' \frac{1}{2} m'_s \rangle.$$

Aqui é também possível explicitar a estrutura matricial correspondente ao espaço de spin escrevendo

$$G = \begin{pmatrix} \{G_{n \frac{1}{2}, n' \frac{1}{2}}\} & \{G_{n \frac{1}{2}, n' -\frac{1}{2}}\} \\ \{G_{n -\frac{1}{2}, n' \frac{1}{2}}\} & \{G_{n -\frac{1}{2}, n' -\frac{1}{2}}\} \end{pmatrix} \equiv \begin{pmatrix} G_{++} & G_{+-} \\ G_{-+} & G_{--} \end{pmatrix}.$$

Alternativamente, é possível ainda definir os elementos de matriz de g no espaço de spin apenas, que são na realidade *operadores* no espaço da partícula sem spin

$$g_{m_s, m'_s} \equiv \langle \frac{1}{2} m_s | g | \frac{1}{2} m'_s \rangle.$$

Em termos desses objetos o operador g é representado pela *matriz de operadores*

$$g \rightarrow \begin{pmatrix} g_{++} & g_{+-} \\ g_{-+} & g_{--} \end{pmatrix}.$$

Em particular, se g é um operador *independente do spin*, a sua extensão ao espaço produto é definida como $g \otimes \hat{1}_S$, onde $\hat{1}_S$ é o operador identidade no espaço de spin, isto é, $g_{++} = g_{--} = g$, $g_{+-} = g_{-+} = 0$. Desse modo, a energia cinética de uma partícula de spin 1/2, por exemplo, pode ser representada como

$$K = \frac{p^2}{2M} \otimes \hat{1}_S \rightarrow \begin{pmatrix} \frac{p^2}{2M} & 0 \\ 0 & \frac{p^2}{2M} \end{pmatrix}.$$

6.2.1 Momento angular de uma partícula com spin 1/2. Adição de momentos angulares.

Para uma partícula com spin 1/2 existem dois operadores vetoriais cujas componentes satisfazem relações de comutação de momento angular, que são o *momento angular orbital* $\vec{L} = (\vec{r} \times \vec{p}) \otimes \hat{1}_S$ (um operador independente de spin estendido ao espaço produto) e o spin $\vec{S} = \hat{1}_r \otimes \vec{S}$ (um operador independente dos graus de liberdade associados à posição, estendido também ao espaço produto através do operador identidade do espaço da partícula sem spin, $\hat{1}_r$). Esses operadores comutam entre si, pois na realidade agem sobre fatores distintos do espaço

produto. As relações de comutação de momento angular permitem então que se construa uma base produto, do tipo considerado na seção anterior, em que os estados $|\phi_n\rangle$ são tomados como autovetores simultâneos de L^2 e de L_3 . Explicitando esse fato com a notação $|\phi_{\lambda lm}\rangle$, onde o índice λ se refere a outros números quânticos, associados a variáveis dinâmicas compatíveis com L^2 e com L^3 , que sejam eventualmente necessários para distinguir vetores diferentes (ortogonais) com os mesmos valores de l e m ,

$$L^2|\phi_{\lambda lm}\rangle = \hbar^2 l(l+1)|\phi_{\lambda lm}\rangle \quad \text{e} \quad L_3|\phi_{\lambda lm}\rangle = \hbar m|\phi_{\lambda lm}\rangle.$$

A base produto formada a partir desses vetores e dos autovetores (6.2) de S_3 , cujos elementos são portanto da forma

$$|\lambda lm \frac{1}{2} m_s\rangle \equiv |\phi_{\lambda lm}\rangle \otimes |\frac{1}{2} m_s\rangle$$

é uma base de autovetores simultâneos de L^2 , L_3 , S^2 e S_3 .

A partir dos operadores \vec{L} e \vec{S} estendidos ao espaço produto, no entanto, é possível definir também o *momento angular total* \vec{J} da partícula com spin como

$$\vec{J} \equiv \vec{L} \otimes \hat{1}_S + \hat{1}_L \otimes \vec{S}$$

ou simplesmente $\vec{J} \equiv \vec{L} + \vec{S}$, usando uma notação abreviada de uso freqüente, que omite a referência explícita à ação trivial dos operadores identidade sobre os respectivos espaços. As componentes de \vec{J} satisfazem também as relações de comutação características de momento angular

$$[J_j, J_k] = i\hbar \epsilon_{jkl} J_l$$

donde segue portanto que $[J^2, \vec{J}] = 0$, de modo que J^2 é compatível com uma qualquer de suas componentes, por exemplo J_3 . É preciso, no entanto, obter ainda as relações de comutação de J^2 e de J_3 com L^2 , L_3 e com S^2 e S_3 , em particular para verificar a compatibilidade ou incompatibilidade dessas várias variáveis dinâmicas. Usando $[\vec{L}, \vec{S}] = 0$, bem como as relações de comutação de momento angular para \vec{L} , \vec{S} e \vec{J} , é fácil ver que

$$[J^2, L^2] = [J^2, S^2] = [J_3, L^2] = [J_3, S^2] = 0 \quad \text{mas} \quad [J^2, L_3] \neq 0 \quad \text{e} \quad [J^2, S_3] \neq 0.$$

Como, por outro lado, $[J^2, J_3] = 0$ e $J_3 = L_3 + S_3$, segue que $[J^2, L_3] = -[J^2, S_3]$. Esses resultados mostram que o momento angular total J^2 *não é compatível* com as componentes L_3 e S_3 do momento angular orbital e do spin, respectivamente. Os vetores $|\lambda m \frac{1}{2} m_s\rangle$ em geral não são, portanto, autovetores de J^2 .

No entanto, como esses vetores são autovetores de L_3 e de S_3 , e lembrando mais uma vez que $J_3 = L_3 + S_3$, eles são também autovetores de J_3 com autovalor $\hbar(m + m_s) \equiv \hbar m_j$. Como L^2 , S^2 e J_3 são compatíveis com J^2 , o fato de que os vetores $|\lambda l m \ 1/2 m_s\rangle$ não sejam em geral autovetores desse operador deve ser imputado ao fato de que existem diferentes combinações de autovalores $\hbar m$ e $\hbar m_s$ de L_3 e S_3 que levam ao *mesmo* autovalor $\hbar(m + m_s) = \hbar m_j$ de J_3 . Isso é suficiente para indicar que os autovetores simultâneos de L^2 , S^2 , J^2 e J_3 , $|\lambda l \ 1/2 j m_j\rangle$, podem ser escritos como combinações lineares dos autovetores de L^2 , L_3 , S^2 e S_3 com os autovalores *desejados* de L^2 e S^2 e com autovalores de L_3 e S_3 tais que a sua *soma* corresponda ao *autovalor desejado* de J_3 . Devido às restrições gerais impostas pelas relações de comutação de momento angular, os autovalores $\hbar j(j + 1)$ de J^2 devem ser tais que $-j \leq m_j \leq j$. Portanto

$$|\lambda l \ 1/2 j m_j\rangle = \sum_{m, m_s} C_{m m_s m_j}^{l \ \frac{1}{2} \ j} |\lambda l m \ 1/2 m_s\rangle \quad (6.8)$$

onde os coeficientes $C_{m m_s m_j}^{l \ \frac{1}{2} \ j}$, chamados *coeficientes de Clebsch-Gordan*, têm claramente o significado das amplitudes de probabilidade

$$C_{m m_s m_j}^{l \ \frac{1}{2} \ j} = \langle \lambda l m \ 1/2 m_s | \lambda l \ 1/2 j m_j \rangle \quad (6.9)$$

que se anulam a menos que $m + m_s = m_j$ e $-j \leq m_j \leq j$ o que implica, em particular, numa restrição sobre os valores de m e m_s na soma em (6.8). O fato de que essas amplitudes são independentes de λ provém de que as variáveis dinâmicas associadas aos números quânticos adicionais designados por esse parâmetro são supostamente compatíveis com os diferentes operadores de momento angular, e portanto não são afetados pela mudança de representação envolvida em (6.8).

A questão de determinar os valores possíveis de j , que define o autovalor do momento angular total J^2 , para um dado valor de l , pode ser respondida facilmente através de uma contagem apropriada dos estados $|\lambda l m \ 1/2 m_s\rangle$. O caso em que $l = 0$ é na realidade trivial, pois nesse caso $m = 0$ e haverá apenas dois estados com $m_j = m_s = \pm 1/2$ que correspondem a $j = 1/2$. Os coeficientes de Clebsch-Gordan que aparecem em (6.8) são portanto

$$C_{0 \pm \frac{1}{2} \ \pm \frac{1}{2}}^{0 \ \frac{1}{2} \ \frac{1}{2}} = 1.$$

Os outros casos, em que $l \geq 1$, podem ser tratados através da construção da fig. 6.1. No primeiro diagrama dessa figura, os $(2l + 1) \times 2$ estados $|\lambda l m \ 1/2 m_s\rangle$ correspondentes aos diferentes valores possíveis de m (para um dado valor de l) e de m_s estão assinalados no plano $m \times m_s$ (onde o valor $l = 3$ é usado como exemplo). Nesse diagrama é fácil ver que existe um único estado com $m_j = m + m_s = l + 1/2$ e também com $m_j = m + m_s = -l - 1/2$, enquanto existem dois estados com valores de m_j compreendidos no intervalo $-l + 1/2 \leq m_j \leq l - 1/2$. Essa contagem de

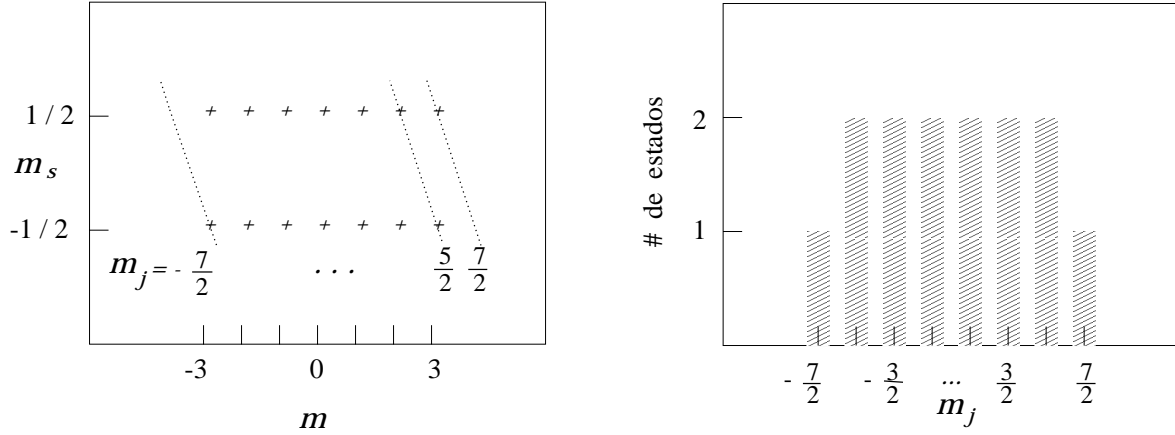


Figura 6.1: Inventário de estados em termos de m e m_s e contagem do número de estados para os diferentes valores possíveis de m_j .

estados classificados pelo valor de m_j é representada no segundo diagrama da Fig. 6.1. O fato de que os estados com $m_j = \pm(l + 1/2)$ são únicos mostra, através da relação (6.8), que eles são na realidade também autovetores de J^2 com autovalor correspondente a $j = l + 1/2$, pois $j \geq |m_j|$. Nesse caso os coeficientes de Clebsch-Gordan são simplesmente

$$C_{l \frac{1}{2} l + \frac{1}{2}}^{l \frac{1}{2} l + \frac{1}{2}} = C_{-l - \frac{1}{2} -l - \frac{1}{2}}^{l \frac{1}{2} l + \frac{1}{2}} = 1.$$

Os autovetores de J^2 correspondentes a $j = l + 1/2$ com valores de m_j tais que $-l + 1/2 \leq m_j \leq l - 1/2$, por outro lado, serão dados como combinações lineares apropriadas dos dois autovetores de L_3 e de S_3 com o valor apropriado de $m + m_s$. Nesses casos, portanto, a soma em (6.8) terá dois termos para cada um desses valores intermediários de m_j . Essa construção deixa ainda sem atribuição de autovalor de J^2 as combinações lineares dos pares de estados com valores intermediários de m_j *ortogonais* às que tenham sido atribuídas a $j = l + 1/2$. Como entre elas ocorrem os autovalores $m_j = \pm(l - 1/2)$ de J_3 , segue que $j = l - 1/2$ corresponde também a um autovalor possível de J^2 dado por $\hbar^2 j(j + 1)$ com $j = l - 1/2$.

A determinação efetiva dos estados $|\lambda l 1/2 j m_j\rangle$ para $-l + 1/2 \leq m_j \leq l - 1/2$ envolve a determinação dos coeficientes de Clebsch-Gordan apropriados. Esse é um processo que não envolve dificuldades maiores que o seu caráter trabalhoso. De fato, o estado $|\lambda l 1/2 j = l + 1/2 m_j = l - 1/2\rangle$ pode ser obtido do estado de projeção máxima $m_j = l + 1/2$ já identificado através do operador $J_- = J_1 - iJ_2$ como

$$\begin{aligned}
J_- |\lambda l \frac{1}{2} \quad l + \frac{1}{2} \quad l + \frac{1}{2} \rangle &= \sqrt{(l + \frac{1}{2})(l + \frac{3}{2}) - (l + \frac{1}{2})(l - \frac{1}{2})} |\lambda l \frac{1}{2} \quad l + \frac{1}{2} \quad l - \frac{1}{2} \rangle \\
&= \sqrt{2l + 1} |\lambda l \frac{1}{2} \quad l + \frac{1}{2} \quad l - \frac{1}{2} \rangle.
\end{aligned}$$

Por outro lado, usando o fato de que o autovetor de projeção máxima é simplesmente o produto dos autovetores também de projeção máxima de L_3 e de S_3 , bem como a ação de L_- e de S_- sobre esses vetores, resulta que

$$J_- |\lambda l \frac{1}{2} \quad l + \frac{1}{2} \quad l + \frac{1}{2} \rangle = (L_- + S_-) |\lambda l l \rangle \otimes |\frac{1}{2} \frac{1}{2} \rangle = \sqrt{2l} |\lambda l l - 1 \rangle \otimes |\frac{1}{2} \frac{1}{2} \rangle + |\lambda l l \rangle \otimes |\frac{1}{2} - \frac{1}{2} \rangle.$$

Comparando essas duas relações resulta então que

$$|\lambda l \frac{1}{2} \quad j = l + \frac{1}{2} \quad m_j = l - \frac{1}{2} \rangle = \sqrt{\frac{2l}{2l + 1}} |\lambda l l - 1 \rangle \otimes |\frac{1}{2} \frac{1}{2} \rangle + \sqrt{\frac{1}{2l + 1}} |\lambda l l \rangle \otimes |\frac{1}{2} - \frac{1}{2} \rangle \quad (6.10)$$

que serve também para identificar os coeficientes de Clebsch-Gordan

$$C_{l-1 \frac{1}{2} \quad l - \frac{1}{2}}^{l \quad \frac{1}{2} \quad l + \frac{1}{2}} = \sqrt{\frac{2l}{2l + 1}} \quad \text{e} \quad C_{l - \frac{1}{2} \quad l - \frac{1}{2}}^{l \quad \frac{1}{2} \quad l + \frac{1}{2}} = \sqrt{\frac{1}{2l + 1}}.$$

Uma vez determinada a decomposição nos autovetores de L_3 e S_3 do estado com $j = l + 1/2$ e $m_j = l - 1/2$ pela relação (6.10), o estado com $j = m_j = l - 1/2$ pode ser escrito como a combinação linear dos mesmos vetores que lhe é ortogonal, isto é

$$|\lambda l \frac{1}{2} \quad j = l - \frac{1}{2} \quad m_j = l - \frac{1}{2} \rangle = -\sqrt{\frac{1}{2l + 1}} |\lambda l l - 1 \rangle \otimes |\frac{1}{2} \frac{1}{2} \rangle + \sqrt{\frac{2l}{2l + 1}} |\lambda l l \rangle \otimes |\frac{1}{2} - \frac{1}{2} \rangle \quad (6.11)$$

que determina os novos coeficientes de Clebsch-Gordan

$$C_{l-1 \frac{1}{2} \quad l - \frac{1}{2}}^{l \quad \frac{1}{2} \quad l - \frac{1}{2}} = -\sqrt{\frac{1}{2l + 1}} \quad \text{e} \quad C_{l - \frac{1}{2} \quad l - \frac{1}{2}}^{l \quad \frac{1}{2} \quad l - \frac{1}{2}} = \sqrt{\frac{2l}{2l + 1}}.$$

Como pode ser facilmente notado, há uma ambiguidade de fase na determinação de (6.11), pois a condição de ortogonalidade fixa apenas o módulo e a *fase relativa* das duas componentes. Para obter a forma mostrada acima essa ambiguidade foi resolvida utilizando uma convenção padrão, segundo a qual os coeficientes de Clebsch-Gordan são reais e o coeficiente $C_{m_l - s \quad m_l - s}^{l \quad s \quad j}$ (que no

caso aparece no segundo termo, que envolve $m_s = -1/2$) é *positivo*. Aplicações sucessivas de J_- aos estados (6.10) e (6.11) permitem determinar de forma análoga os autovetores restantes de J^2 e de J_3 bem como os respectivos coeficientes de Clebsch-Gordan.

A expansão (6.8) dos autovetores de J^2 e J_3 em termos dos autovetores de L_3 e S_3 pode evidentemente ser invertida, o que leva a expressões que dão os vetores de estado $|\lambda m \ 1/2 m_s\rangle$ em termos dos $|\lambda l \ 1/2 j m_j\rangle$. De fato usando a expressão (6.9) para os coeficientes de Clebsch-Gordan e adotando as convenções de fase usuais, que tornam esses coeficientes reais, é imediato que

$$|\lambda m \ 1/2 m_s\rangle = \sum_{j=j_{\min}}^{j_{\max}} |\lambda l \ 1/2 j m_j\rangle \langle \lambda l \ 1/2 j m_j | \lambda m \ 1/2 m_s \rangle = \sum_{j=j_{\min}}^{j_{\max}} C_{m m_s m_j}^{l \ \frac{1}{2} \ j} |\lambda l \ 1/2 j m_j\rangle.$$

Os coeficientes que aparecem nessa expansão são portanto *os mesmos* que aparecem na expansão (6.8), embora a soma seja feita neste caso sobre os diferentes valores possíveis de j . Esta última relação pode ser verificada explicitamente invertendo as relações (6.10) e (6.11) para obter, por exemplo

$$\begin{aligned} |\lambda l \ l - 1 \ \frac{1}{2} \ \frac{1}{2}\rangle &= \sqrt{\frac{2l}{2l+1}} |\lambda l \ \frac{1}{2} \ l + \frac{1}{2} \ l - \frac{1}{2}\rangle - \sqrt{\frac{1}{2l+1}} |\lambda l \ \frac{1}{2} \ l - \frac{1}{2} \ l - \frac{1}{2}\rangle \\ &= C_{l-1 \ \frac{1}{2} \ l - \frac{1}{2}}^{l \ \frac{1}{2} \ l + \frac{1}{2}} |\lambda l \ \frac{1}{2} \ l + \frac{1}{2} \ l - \frac{1}{2}\rangle + C_{l-1 \ \frac{1}{2} \ l - \frac{1}{2}}^{l \ \frac{1}{2} \ l - \frac{1}{2}} |\lambda l \ \frac{1}{2} \ l - \frac{1}{2} \ l - \frac{1}{2}\rangle. \end{aligned}$$

O esquema, baseado na construção da Fig. 6.1, que permite a identificação dos valores possíveis do momento angular total j para valores dados dos momentos angulares constituintes l e s , pode ser estendido de forma imediata para valores quaisquer, inteiros ou semi-inteiros, de l e s . Em qualquer caso deve valer a relação

$$(2l+1)(2s+1) = \sum_{j=j_{\min}}^{j_{\max}} (2j+1)$$

que leva em qualquer caso à limitação

$$j_{\min} = |l - s| \leq j \leq l + s = j_{\max}, \quad \Delta j = \text{inteiro}.$$

Os valores possíveis de j serão inteiros se ambos os momentos angulares constituintes são inteiros ou semi-inteiros, e serão semi-inteiros se um dos constituintes for inteiro enquanto o outro é semi-inteiro. O coeficiente de Clebsch-Gordan para o estado de máximo $j = l + s$ e máximo $m_j = j = l + s$ é sempre igual a 1, e os coeficientes restantes podem ser determinados por aplicações sucessivas do operador J_- a esse estado e através do uso de relações de ortogonalidade, fazendo uso, neste último caso, da convenção padrão para escolha de fases.

6.3 Momentos magnéticos e interações dependentes do spin.

A inclusão de spins $1/2$ como graus de liberdade adicionais do próton e do eléctron que constituem o átomo de hidrogênio tem como único efeito modificar o grau de degenerescência do espectro enquanto não sejam introduzidas interações envolvendo esses novos graus de liberdade. Na ausência de novas interações desse tipo, de fato, o Hamiltoniano age como o operador identidade sobre os estados dos spins que portanto se manifestam apenas tornando cada um dos estados “sem spin” quadruplicamente degenerado. Em particular, os estados estacionários desse Hamiltoniano podem ser escritos sob a forma

$$|nljm\rangle \otimes |\frac{1}{2}m_s^{(e)}\rangle \otimes |\frac{1}{2}m_s^{(p)}\rangle$$

e os autovalores correspondentes E_{nl} são independentes não apenas do autovalor $\hbar m$ de L_3 mas também dos autovalores $\hbar m_s^{(e)}$ e $\hbar m_s^{(p)}$ das componentes $S_3^{(e)}$ e $S_3^{(p)}$ dos spins do eléctron e do próton, respectivamente.

As interações envolvendo os spins que são relevantes no contexto atômico são de natureza eletromagnética, e estão ligadas ao fato experimental de que ao spin do eléctron e ao spin do próton estão associados momentos de dipolo magnético (ou simplesmente “momentos magnéticos”) $\vec{\mu}_e$ e $\vec{\mu}_p$ que podem ser escritos em termos dos respectivos operadores de spin como

$$\vec{\mu}_e = -g_e \frac{e}{2m_e c} \vec{S}^{(e)} \quad \text{e} \quad \vec{\mu}_p = g_p \frac{e}{2m_p c} \vec{S}^{(p)}.$$

A carga eléctrica do eléctron e do próton são $\mp e$ e as respectivas massas são m_e e m_p . As constantes g_e e g_p são chamadas *fatores giromagnéticos* e têm os valores experimentais

$$g_e \simeq 2.00 \quad \text{e} \quad g_p \simeq 5.58.$$

Esses momentos magnéticos interagem entre si, interagem com campos magnéticos associados à dinâmica orbital das partículas carregadas que constituem o átomo e ainda com eventuais campos magnéticos externos a que o sistema atômico seja submetido. Esses tres tipos de interação estão ligados respectivamente à estrutura hiperfina, à estrutura fina e aos vários tipos de efeito Zeeman do espectro atômico.

6.3.1 Interação spin-órbita e estrutura fina.

A energia clássica de interação entre um momento magnético $\vec{\mu}$ e um campo magnético \vec{B} é

$$H_1 = -\vec{\mu} \cdot \vec{B}.$$

No sistema de repouso de um eléctron que se move com velocidade \vec{v} em um campo eléctrico \vec{E} existe um campo magnético que é dado, através de uma transformação de Lorentz, por

$$\vec{B} = -\frac{\vec{v}}{c} \times \vec{E}.$$

Desse modo, a energia de interação do momento magnético $\vec{\mu}_e$ do eléctron com o campo magnético gerado pela sua dinâmica orbital relativa ao próton aparece sob a forma

$$H_1 = \vec{\mu}_e \cdot \frac{\vec{v}}{c} \times \vec{E} = -g_e \frac{e}{2m_e c^2} \vec{S}^{(e)} \cdot \vec{v} \times \vec{E}.$$

O campo eléctrico a ser considerado neste caso é o campo coulombiano do próton

$$\vec{E}(\vec{r}) = \frac{e}{r^3} \vec{r}$$

de modo que, usando $\vec{v} = \vec{p}/\mu$, onde μ é a massa reduzida do sistema eléctron-próton, resulta

$$H_1 = -g_e \frac{e^2}{2m_e \mu c^2} \frac{1}{r^3} \vec{S}^{(e)} \cdot \vec{p} \times \vec{r} = g_e \frac{e^2}{2m_e \mu c^2} \frac{1}{r^3} \vec{S}^{(e)} \cdot \vec{L}$$

onde $\vec{L} = \vec{r} \times \vec{p} = -\vec{p} \times \vec{r}$ é o momento angular orbital do eléctron. A esse resultado deve ser somada uma correção relativística descoberta em 1926 por Thomas² que tem como resultado o cancelamento de metade do seu valor. Como $g_e \simeq 2.00$, isso corresponde essencialmente à substituição $g_e \rightarrow g_{LS} = 1$, e a energia de interação entre o momento magnético do eléctron e a sua dinâmica espacial se exprime através de um termo de acoplamento entre o spin do eléctron e o seu momento angular orbital, dado por

$$H_1 \rightarrow H_{LS} = \frac{e^2}{2m_e \mu c^2} \frac{1}{r^3} \vec{S}^{(e)} \cdot \vec{L} = \frac{1}{2m_e \mu c^2} \frac{1}{r} \frac{dV}{dr} \vec{S}^{(e)} \cdot \vec{L}, \quad V(r) \equiv -\frac{e^2}{r}.$$

Esse termo, que é claramente independente do spin do próton, é chamado *termo de acoplamento spin-órbita*, e dá origem à estrutura fina do átomo de hidrogênio. Termos semelhantes existem também no caso de átomos mais complexos. Com a sua inclusão, o Hamiltoniano que descreve o átomo de hidrogênio passa a ser dado por

²L. H. Thomas, *Nature* **117**, 514 (1926).

$$H = \left(\frac{p^2}{2\mu} - \frac{e^2}{r} \right) \otimes \hat{1}_S + \frac{1}{2m_e\mu c^2} \frac{1}{r} \frac{dV}{dr} \vec{S}^{(e)} \cdot \vec{L} \equiv H_0 + H_{LS}. \quad (6.12)$$

Enquanto o primeiro termo de (6.12) age como o operador unidade no espaço de spin do eléctron, o segundo termo pode ser representado como

$$\frac{1}{2m_e\mu c^2} \frac{1}{r} \frac{dV}{dr} \vec{S}^{(e)} \cdot \vec{L} = \frac{\hbar}{4m_e\mu c^2} \frac{1}{r} \frac{dV}{dr} \begin{pmatrix} L_3 & L_- \\ L_+ & -L_3 \end{pmatrix}$$

que, devido aos termos $L_{\pm} = L_1 \pm iL_2$, é claramente não diagonal em m e m_s na base produto $|\lambda l m 1/2 m_s\rangle$. Em outras palavras, L_3 e $S_3^{(e)}$ deixam de ser constantes do movimento quando o termo de interação spin-órbita é incluído no Hamiltoniano.

Outra forma de ver isso é simplesmente notar que $[L_3, H_{LS}] \neq 0$ e $[S_3, H_{LS}] \neq 0$. Esse mesmo cálculo mostra, no entanto, que $[L_3, H_{LS}] = -[S_3, H_{LS}]$, de modo que $[J_3, H_{LS}] = 0$. Como J^2 também é uma constante do movimento, é claro que a representação mais adequada para tratar os efeitos do termo de interação spin-órbita é aquela em que L^2 , S^2 , J^2 e J_3 são diagonais. De fato, o operador $\vec{S}^{(e)} \cdot \vec{L}$ pode ser escrito sob a forma

$$\vec{S}^{(e)} \cdot \vec{L} = \frac{1}{2} (J^2 - L^2 - S^{(e)2})$$

que mostra imediatamente que H_{LS} é diagonal nos números quânticos de momento angular na base acoplada $|\lambda l 1/2 j m_j\rangle$ construída como na equação (6.8), pois

$$(\vec{S}^{(e)} \cdot \vec{L}) |\lambda l 1/2 j m_j\rangle = \frac{\hbar^2}{2} \left(j(j+1) - l(l+1) - \frac{3}{4} \right) |\lambda l 1/2 j m_j\rangle \quad (6.13)$$

enquanto o termo independente de spin H_0 depende apenas de L^2 , que é também diagonal nessa base. Uma consequência importante da equação (6.13) é a de que não há efeitos da interação spin-órbita para estados com $l = 0$, pois nesse caso $j = 1/2$ e o autovalor de $\vec{S}^{(e)} \cdot \vec{L}$ é zero.

Tratamento perturbativo. Os resultados acima permitem tratar perturbativamente os efeitos da interação spin-órbita de forma bastante simples. Os estados não perturbados são autovalores do termo independente de spin H_0 de (6.12) que, na representação em que L_3 e S_3 são diagonais, são dados por

$$|nlm 1/2 m_s\rangle \rightarrow \frac{u_{nl}(r)}{r} Y_{lm}(\hat{r}) |1/2 m_s\rangle$$

onde \hat{r} indica as coordenadas angulares θ e φ de \vec{r} e Y_{lm} é uma harmônica esférica, autofunção normalizada de L^2 e de L_3 . Os autovalores não perturbados dependem apenas de n e são

$$E_n^{(0)} = -\frac{\mu e^4}{2\hbar^2 n^2} = -\frac{\mu}{m_e n^2} R$$

onde R é a constante de Rydberg. Esses autovalores são portanto degenerados, e contém os diferentes valores possíveis de l , tais que $l \leq n - 1$, além dos $2l + 1$ valores possíveis de m para cada um desses valores de l e dos dois valores possíveis de m_s . Embora o termo de interação spin-órbita não seja diagonal nesse grupo de estados degenerados, ele pode ser diagonalizado pela transformação unitária que diagonaliza o quadrado do momento angular total, J^2 . Essa transformação pode ser escrita em termos dos coeficientes de Clebsch-Gordan apropriados como em (6.8), que superpõe apenas diferentes estados *degenerados* da representação desacoplada $|nlm\ 1/2m_s\rangle$. A matriz que representa o termo de interação spin-órbita no grupo de estados degenerados é portanto diagonal na representação acoplada e tem como elementos, usando a relação (6.13),

$$\begin{aligned} \langle nl1/2\ jm_j | H_{LS} | nl'1/2\ j'm'_j \rangle &= \frac{1}{4m_e\mu c^2} \langle nl1/2\ jm_j | \frac{1}{r} \frac{dV}{dr} (J^2 - L^2 - S^2) | nl'1/2\ j'm'_j \rangle = \\ &= \frac{\hbar^2}{4m_e\mu c^2} \left[j(j+1) - l(l+1) - \frac{3}{4} \right] \delta_{jj'} \delta_{ll'} \delta_{m_j m'_j} \int_0^\infty |u_{nl}(r)|^2 \frac{1}{r} \frac{dV}{dr} dr. \end{aligned}$$

As correções perturbativas (de primeira ordem no termo de interação spin-órbita) para os autovalores da energia dependem portanto explicitamente de l e de j . Os autovalores corrigidos perturbativamente são dados por

$$E_{nlj} = E_n^{(0)} + \frac{\hbar^2}{4m_e\mu c^2} \left[j(j+1) - l(l+1) - \frac{3}{4} \right] R_{nl}$$

onde R_{nl} é a integral radial que aparece nos elementos de matriz do termo de interação spin-órbita. A dependência de l e j é introduzida através do fator

$$j(j+1) - l(l+1) - \frac{3}{4} = \begin{cases} -(l+1) & \text{para } j = l - 1/2, \ l > 0 \\ l & \text{para } j = l + 1/2. \end{cases}$$

Para estados com $l = 0$ (e $j = 1/2$) a correção é nula, e os estados com $j = l \pm 1/2$, para cada valor não nulo de l , sofrem deslocamentos de sinais opostos. Como a integral radial é *positiva* (pois $dV/dr > 0$), os estados com $j = l - 1/2$ se tornam mais ligados, e os estados com $j = l + 1/2$ se tornam menos ligados, que os estados não perturbados de partida. Desse modo,

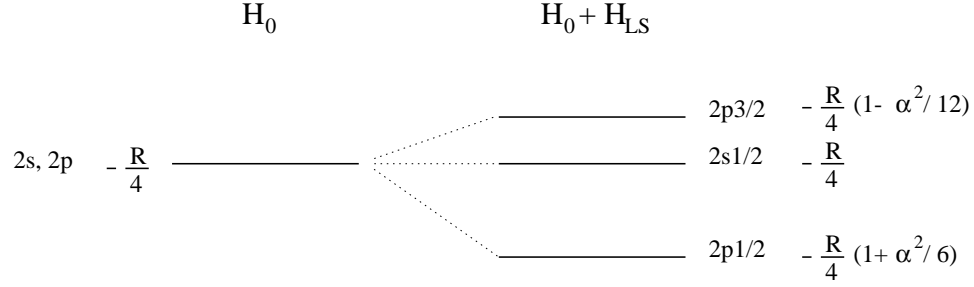


Figura 6.2: Efeito perturbativo da interação spin-órbita sobre os níveis $n = 2$ do átomo de hidrogênio. O zero da escala de energia é a energia de dissociação do hidrogênio, e as energias dos níveis ligados incluídos na figura foram calculadas supondo $\mu/m_e \simeq 1$.

o estado fundamental ($n = 1$) não é afetado pelo termo de interação spin-órbita; o estado com $n = 2$ e $l = 0$ tão pouco é afetado, enquanto que os estados com $n = 2$ e $l = 1$ se quebram em dois níveis, associados respectivamente aos valores $j = 1/2$ e $j = 3/2$ do momento angular total. A integral radial pode ser nesse caso calculada explicitamente como

$$\int_0^\infty |u_{21}(r)|^2 \frac{e^2}{r^3} dr = \frac{\mu^3}{m_e^3} \frac{e^2}{24a_0^3}$$

onde a_0 é o raio de Bohr. Desse modo resulta que

$$E_{21\frac{1}{2}} = -\frac{\mu}{4m_e} \left(1 + \frac{\mu^2}{m_e^2} \frac{\alpha^2}{6} \right) R \quad \text{e} \quad E_{21\frac{3}{2}} = -\frac{\mu}{4m_e} \left(1 - \frac{\mu^2}{m_e^2} \frac{\alpha^2}{12} \right) R \quad (6.14)$$

onde $\alpha = e^2/\hbar c \simeq 1/137$ é a *constante da estrutura fina*. A energia do estado com $n = 2$ e $l = 0$ permanece com o valor não perturbado $E_{20\frac{1}{2}} = E_2^{(0)} = -\mu R/4m_e$. A escala de energias associada aos efeitos da interação spin-órbita é de fato *mais fina* por um fator $\sim 10^{-4}$ que a escala dos espaçamentos dos níveis mais baixos do espectro não perturbado, dada pela constante de Rydberg R . O esquema de níveis resultante é então o que aparece na fig. 6.2.

O tratamento perturbativo do termo de interação spin-órbita na realidade encobre uma dificuldade que se manifesta quando se procura uma solução exata do problema de autovalores associado ao Hamiltoniano (6.12). A parte angular e de spin da equação de autovalores pode ser tratada na representação acoplada (em que J^2 é diagonal) levando, na representação de Schrödinger, a uma equação ordinária que envolve apenas a coordenada radial r . De fato, escrevendo a solução do problema de autovalores, para o valor E_r da energia relativa, sob a forma

$$\frac{u_{E_r}(r)}{r} \langle \hat{r} | l 1/2 j m_j \rangle \equiv \frac{u_{E_r}(r)}{r} \sum_{m_s} C_{m_j - m_s}^l \frac{1}{m_s} \frac{j}{m_j} Y_{l m_j - m_s}(\hat{r}) | \frac{1}{2} m_s \rangle,$$

resulta que a função radial $u_{E_r}(r)$ deve satisfazer a equação

$$\frac{d^2 u_{E_r l j}(r)}{dr^2} + \left[\frac{2\mu E_r}{\hbar^2} - \frac{l(l+1)}{r^2} + \frac{2\mu e^2}{\hbar^2 r} - \frac{e^2}{m_e c^2 r^3} \left(j(j+1) - l(l+1) - \frac{3}{4} \right) \right] u_{E_r l j}(r) = 0$$

cujas soluções $u_{E_r}(r)$ dependem de l e também de j , o que foi indicado explicitamente através dos índices adicionados a essas funções. Como no caso do modelo simples, independente do spin do elétron, a função radial deve satisfazer às condições $u(r=0)=0$ e de ser de quadrado integrável (no caso do espectro discreto) ou limitada (no caso do espectro contínuo).

A dificuldade consiste em que, como visto anteriormente no contexto do tratamento perturbativo, o termo de interação spin-órbita leva, no caso em que $l \neq 0$ e $j = l - 1/2$, a um termo de potencial de caráter *atrativo* e, no caso de uma carga puntiforme, proporcional a $1/r^3$. O cálculo da correção perturbativa de ordem mais baixa devida a esse termo não apresentou qualquer problema, pois o comportamento na vizinhança da origem das funções radiais não perturbadas *para* $l \neq 0$ é suficiente para garantir a convergência das integrais radiais relevantes, enquanto que no caso $l = 0$ o coeficiente do termo em $1/r^3$ é nulo. No entanto, um potencial atrativo com a dependência radial $1/r^3$ nas vizinhanças da origem é suficientemente forte³ para que a energia do estado de menor energia para qualquer valor dado de l colapse para $-\infty$. Isso pode de fato ser verificado usando um argumento variacional para obter um limite superior para o menor autovalor da equação radial, vista como uma equação de autovalores

$$h_r u(r) = \frac{2\mu E_r}{\hbar^2} u(r)$$

para o operador

$$h_r \equiv -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} - \frac{2\mu e^2}{\hbar^2 r} + \frac{e^2}{m_e c^2 r^3} \left(j(j+1) - l(l+1) - \frac{3}{4} \right).$$

De fato, usando a função de prova normalizada, dependente do parâmetro b

$$u(r) = \sqrt{\frac{8}{3\sqrt{\pi} b^5}} r^2 e^{-\frac{r^2}{2b^2}}, \quad \int_0^\infty u^2(r) dr = 1$$

é fácil verificar que

³v. L. D. Landau e E. M. Lifshitz, *op. cit.*, §14 e §35.

$$\int_0^\infty u(r) \frac{d^2 u}{dr^2} dr = \frac{-|C_K|}{b^2} \quad \text{e} \quad \int_0^\infty \frac{u^2(r)}{r^n} dr = \frac{|C_n|}{b^n}, \quad n = 1, 2, 3$$

e portanto se o coeficiente de $1/r^3$ em h_r é *negativo*, como acontece no caso em que $j = l - 1/2$, a quantidade

$$\int_0^\infty u(r) h_r u(r) dr$$

pode ser tornada arbitrariamente *negativa*, tomando b suficientemente pequeno, o que implica num limite superior também arbitrariamente negativo para E_r .

Como essa dificuldade está associada a uma singularidade que decorre em última análise do potencial coulombiano de uma carga puntiforme, parece natural supor que para evita-la seja preciso modificar o potencial central para distâncias pequenas. Isso, no entanto, não é verdade, pois o tratamento relativístico do potencial coulombiano no contexto da equação de Dirac (v. Capítulo 11) evita a “queda para o centro” presente no tratamento não relativístico.

Discrepâncias e consertos. Os resultados para os estados estacionários do átomo de hidrogênio com $n = 2$ contém uma discrepância importante com relação aos níveis observados experimentalmente. De fato, embora a diferença de energia experimental entre os níveis $2s_{1/2}$ e $2p_{1/2}$ seja apenas a do deslocamento de Lamb, o que corresponde a $E_{20\frac{1}{2}}^{(\text{exp})} - E_{21\frac{1}{2}}^{(\text{exp})} \simeq 3 \times 10^{-7} R$, o resultado da inclusão dos efeitos da interação spin-órbita dá $E_{20\frac{1}{2}} - E_{21\frac{1}{2}} = \alpha^2 R/24 \simeq 2 \times 10^{-6} R$, que é praticamente uma ordem de grandeza maior que o valor observado. A origem dessa discrepância pode ser determinada notando, em primeiro lugar, que o termo de interação spin-órbita é uma correção relativística ao modelo mais simples do átomo de hidrogênio que é de ordem v^2/c^2 , tomando para v um valor típico da velocidade do elétron, que pode ser obtido no contexto da teoria quântica como

$$v \sim \frac{\langle p \rangle}{m_e} \sim \frac{1}{m_e} \frac{\hbar}{a_0}.$$

De fato, valores típicos do termo de interação spin-órbita podem ser estimados como

$$H_{LS} \sim \frac{e^2}{2m_e^2 c^2} \frac{1}{a_0^3} \frac{\hbar}{2} m_e v a_0 \sim \frac{1}{4} \frac{e^2}{a_0} \frac{v^2}{c^2}.$$

Em segundo lugar, deve-se notar que existem outras correções relativísticas dessa mesma ordem que estão sendo até aqui ignoradas. A primeira delas, que pode ser obtida de forma simples, é

uma correção ao termo de energia cinética, que resulta da expansão da expressão relativística para a energia cinética

$$\sqrt{m_e^2 c^4 + p^2 c^2} - m_e c^2 = \frac{p^2}{2m_e} - \frac{p^4}{8m_e^3 c^2} + \mathcal{O}\left(\frac{v^4}{c^4}\right)$$

o que leva à identificação da correção

$$H_K = -\frac{p^4}{8m_e^3 c^2}. \quad (6.15)$$

Para obter as contribuições perturbativas para os autovalores da energia devidas a esse novo termo é útil notar que

$$H_K = -\frac{\mu^2}{2m_e^3 c^2} \frac{p^4}{4\mu^2} = -\frac{\mu^2}{2m_e^3 c^2} \left(H_0 + \frac{e^2}{r}\right)^2$$

que, sendo independente do spin e compatível com L^2 , tem como únicos elementos de matriz não nulos nos estados não perturbados com $n = 2$

$$\langle 20 \ 1/2 \ 1/2 m_j | H_K | 20 \ 1/2 \ 1/2 m_j \rangle = -\frac{\mu^3}{m_e^3} \frac{13}{64} \alpha^2 R$$

e

$$\langle 21 \ 1/2 \ j m_j | H_K | 21 \ 1/2 \ j m_j \rangle = -\frac{\mu^3}{m_e^3} \frac{7}{192} \alpha^2 R.$$

Existe ainda uma segunda correção relativística de mesma ordem para o potencial central, conhecida como *termo de Darwin*, que é dada por

$$H_D = \frac{\hbar^2}{8m_e^2 c^2} \nabla^2 V(r) \rightarrow \frac{\pi e^2 \hbar^2}{2m_e^2 c^2} \delta(\vec{r}) \quad (6.16)$$

onde a última passagem envolveu a suposição de um potencial coulombiano $-e^2/r$ e o uso da equação de Poisson. A contribuição perturbativa desse termo se restringe portanto aos estados com $l = 0$, pois apenas estes estados têm função de onda diferente de zero em $\vec{r} = 0$. O elemento de matriz relevante é, no caso dos estados com $n = 2$,

$$\langle 20 \ 1/2 \ 1/2 m_j | H_D | 20 \ 1/2 \ 1/2 m_j \rangle = \frac{\mu^3}{m_e^3} \frac{1}{8} \alpha^2 R.$$

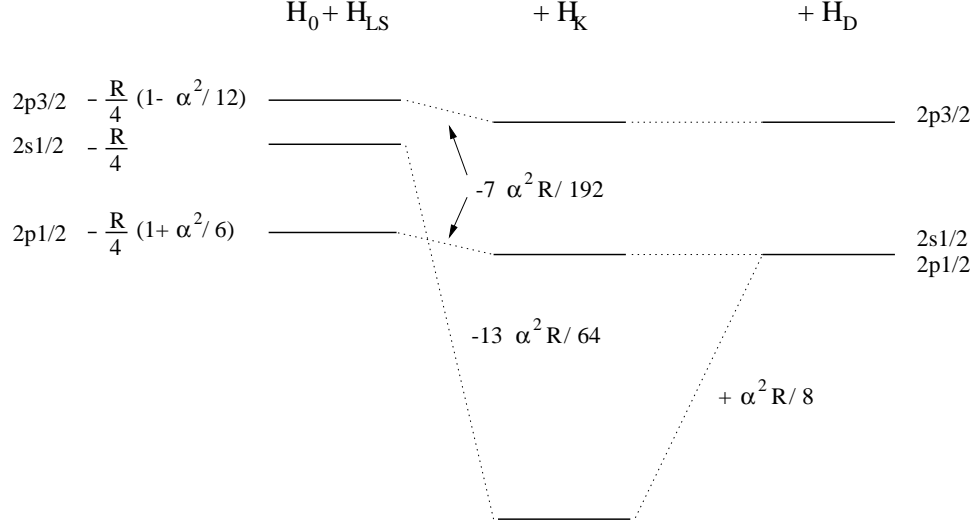


Figura 6.3: Deslocamentos perturbativos dos níveis $n = 2$ do hidrogênio devidos à correção relativística para a energia cinética H_K e ao termo de Darwin H_D . As energias foram calculadas supondo $\mu/m_e \simeq 1$.

Adicionando finalmente as novas correções ao resultado (6.14) o que se obtém para as energias perturbadas dos estados com $n = 2$ é

$$\begin{aligned}
 E_{21\frac{3}{2}} &= -\frac{\mu}{m_e} \frac{R}{4} \left(1 - \frac{\mu^2}{m_e^2} \frac{\alpha^2}{12} \right) - \frac{\mu^3}{m_e^3} \frac{7}{192} \alpha^2 R = -\frac{\mu}{m_e} \frac{R}{4} \left(1 + \frac{\mu^2}{m_e^2} \frac{\alpha^2}{16} \right) \\
 E_{21\frac{1}{2}} &= -\frac{\mu}{m_e} \frac{R}{4} \left(1 + \frac{\mu^2}{m_e^2} \frac{\alpha^2}{6} \right) - \frac{\mu^3}{m_e^3} \frac{7}{192} \alpha^2 R = -\frac{\mu}{m_e} \frac{R}{4} \left(1 + \frac{\mu^2}{m_e^2} \frac{5\alpha^2}{16} \right) \\
 E_{20\frac{1}{2}} &= -\frac{\mu}{m_e} \frac{R}{4} - \frac{\mu^3}{m_e^3} \frac{13}{64} \alpha^2 R + \frac{\mu^3}{m_e^3} \frac{1}{8} \alpha^2 R = -\frac{\mu}{m_e} \frac{R}{4} \left(1 + \frac{\mu^2}{m_e^2} \frac{5\alpha^2}{16} \right).
 \end{aligned}$$

Esses resultados estão representados na figura 6.3. Os estados com $l = 0$ e $l = 1$ cujo momento angular total é $j = 1/2$ aparecem como degenerados após a inclusão da correção relativística para a energia cinética e do termo de Darwin, enquanto a separação entre esses estados e o estado com $j = 3/2$, devida ao termo de acoplamento spin-órbita, se mantém (a menos do fator próximo de 1, envolvendo a massa reduzida) com o valor $\alpha^2 R/16$.

6.3.2 Interação spin-spin e estrutura hiperfina.

Sendo o próton também uma partícula de spin 1/2 dotada de um momento magnético $\vec{\mu}_p$, existe no átomo de hidrogênio uma interação spin-spin, associada à interação dos momentos magnéticos intrínsecos do próton e do elétron, e uma interação do momento magnético intrínseco do próton com o momento magnético orbital associado ao elétron, além da interação do spin do elétron com o campo magnético associado à dinâmica orbital, que dá origem ao termo de interação spin-órbita. Os efeitos destas interações, envolvendo o momento magnético do núcleo, são muito pequenos na escala da estrutura fina e dão origem à chamada *estrutura hiperfina* do espectro atômico.

Quando o momento angular orbital é nulo (como no caso do estado fundamental do modelo simples para o átomo de hidrogênio) não há momento magnético orbital, e a interação hiperfina se reduz à interação entre os momentos magnéticos intrínsecos do elétron e do núcleo. Por simplicidade a discussão que segue se restringe a este caso. A energia associada à interação dos dois momentos magnéticos pode ser escrita como $-\vec{\mu}_e \cdot \vec{B}_{\vec{\mu}_p}(\vec{r})$ onde $\vec{B}_{\vec{\mu}_p}(\vec{r})$ é o campo magnético devido ao momento magnético do próton (suposto como sendo puntiforme) a uma distância \vec{r} do próton. Uma expressão clássica para esse campo é⁴

$$\vec{B}_{\vec{\mu}_p}(\vec{r}) = \nabla \times \frac{\vec{\mu}_p \times \vec{r}}{r^3} = \frac{3\hat{r}(\vec{\mu}_p \cdot \hat{r}) - \vec{\mu}_p}{r^3} + 4\pi(1-d)\vec{\mu}_p\delta(\vec{r}) \quad (6.17)$$

na qual os dois primeiros termos correspondem ao campo *externo* a um momento de dipolo magnético $\vec{\mu}_p$ e o último termo corresponde, no limite em que a distribuição de magnetização responsável por $\vec{\mu}_p$ é tratada como sendo puntiforme, a um “campo interno” a essa distribuição. É bem sabido que o campo magnético \vec{B} interno a uma distribuição finita de magnetização *uniforme* \vec{M} , limitada por uma dada superfície Σ na realidade depende da forma de Σ . O problema de condições de contorno correspondente pode ser resolvido analiticamente para o caso de uma superfície elipsoidal⁵ com o resultado que o campo \vec{B} interno à distribuição de magnetização é dado por

$$\vec{B} = \vec{H} + 4\pi\vec{M} \quad \text{com} \quad \vec{H} = -4\pi d \vec{M}$$

ou seja

⁴v. e.g. J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, N.Y. 1962, p. 146; v. também o tratamento dado a esta questão na segunda edição deste livro (1975, pp. 180-187), comparando especialmente a eq. (5.58) da primeira edição com a eq. (5.64) da segunda.

⁵v. e.g. J. C. Slater e N. H. Frank, *Electromagnetism*, McGraw-Hill Book Co., 1947, pag. 73, ou ainda A. Sommerfeld, *Electrodynamics* (Lectures on Theoretical Physics, Vol III), Academic Press, N.Y., 1952, pag. 95. Em textos mais recentes, como por exemplo J. D. Jackson, *Op. cit.* a referência ao fator de desmagnetização parece ter sido abandonada, embora um fenômeno análogo seja tratado no contexto da eletrostática em termos de um *fator de despolarização*.

$$\vec{B} = 4\pi(1-d)\vec{M}$$

onde d é o chamado *fator de desmagnetização*, que varia de valores próximos de zero para um elipsóide extremamente alongado (isto é, com geometria tipo agulha) até valores próximos de 1, para um elipsóide extremamente achatado (geometria tipo disco). Para o caso intermediário de uma distribuição esférica o fator de desmagnetização tem o valor $1/3$. No que segue esse termo será escrito, como na equação (6.17), em termos de um fator de desmagnetização d genérico, a ser tratado oportunamente como um parâmetro fenomenológico.

Escrevendo os momentos magnéticos como

$$\vec{\mu}_e \equiv \mu_e \frac{\vec{S}^{(e)}}{\hbar} \quad \text{e} \quad \vec{\mu}_p \equiv \mu_p \frac{\vec{S}^{(p)}}{\hbar} \quad (6.18)$$

onde $\vec{S}^{(e,p)} = \hbar\vec{\sigma}_{e,p}/2$ são os operadores de spin para o elétron e para o próton, com $\vec{\sigma} \equiv \{\sigma_1, \sigma_2, \sigma_3\}$ e sendo σ_i as matrizes de Pauli, a energia de interação $-\vec{\mu}_e \cdot \vec{B}_{\vec{\mu}_p}$ adquire a forma de uma interação spin-spin dada por

$$H_{SS} = -\frac{\mu_e \mu_p}{\hbar^2} \left[\frac{3(\vec{S}^{(e)} \cdot \hat{r})(\vec{S}^{(p)} \cdot \hat{r}) - \vec{S}^{(e)} \cdot \vec{S}^{(p)}}{r^3} + 4\pi(1-d)\vec{S}^{(e)} \cdot \vec{S}^{(p)}\delta(\vec{r}) \right].$$

Essa expressão pode ser re-escrita em termos do spin total $\vec{S} \equiv \vec{S}^{(e)} + \vec{S}^{(p)}$ usando as identidades

$$\vec{S}^{(e)} \cdot \vec{S}^{(p)} \equiv \frac{1}{2} (S^2 - S^{(e)2} - S^{(p)2}) = \frac{1}{2} \left(S^2 - \frac{3\hbar^2}{2} \right) \quad (6.19)$$

e

$$(\vec{S}^{(e)} \cdot \hat{r})(\vec{S}^{(p)} \cdot \hat{r}) \equiv \frac{1}{2} [(\vec{S} \cdot \hat{r})^2 - (\vec{S}^{(e)} \cdot \hat{r})^2 - (\vec{S}^{(p)} \cdot \hat{r})^2] \equiv \frac{1}{2} \left[(\vec{S} \cdot \hat{r})^2 - \frac{\hbar^2}{2} \right]. \quad (6.20)$$

A última passagem envolveu o fato de que o quadrado de qualquer projeção de $\vec{\sigma}$ é igual à matriz unidade (cf. Problema **6.2**). Usando as relações (6.19) e (6.20) a interação spin-spin no átomo de hidrogênio pode ser escrita como

$$H_{SS} = -\frac{\mu_e \mu_p}{2\hbar^2} \left[\frac{3(\vec{S} \cdot \hat{r})^2 - S^2}{r^3} + 4\pi(1-d) \left(S^2 - \frac{3\hbar^2}{2} \right) \delta(\vec{r}) \right]. \quad (6.21)$$

Essa parte da interação hiperfina, sem contribuições provenientes do momento angular orbital do elétron, é suficiente para tratar a quebra de degenerescência do estado fundamental do modelo simples para o átomo de hidrogênio, no qual o estado relativo é descrito pela função de onda $nl = 1s$. Incluindo as quatro combinações possíveis de estados de spin para o próton e para o elétron na representação em que o spin total S^2 é diagonal

$$|1s1/2\ 1/2\ SM_S\rangle \rightarrow \left(\frac{\mu^3}{\pi m_e^3 a_0^3}\right)^{\frac{1}{2}} e^{-\frac{\mu r}{m_e a_0}} |1/2\ 1/2\ SM_S\rangle$$

onde $S = 0, 1$ e $-S \leq M_S \leq S$, a_0 é o raio de Bohr e μ , m_e são respectivamente a massa reduzida e a massa do elétron. Esses quatro estados degenerados são independentes dos ângulos $\hat{r} \equiv \{\theta, \varphi\}$ de \vec{r} , de modo que as integrações angulares no caso do termo dipolar de (6.21) no cálculo dos elementos de matriz levam a um resultado nulo. De fato, usando

$$\vec{S} \cdot \hat{r} = S_1 \sin \theta \cos \varphi + S_2 \sin \theta \sin \varphi + S_3 \cos \theta$$

um cálculo simples mostra que

$$\int_{-1}^{+1} d \cos \theta \int_0^{2\pi} d\varphi (\vec{S} \cdot \hat{r})^2 = \frac{4\pi}{3} S^2$$

donde segue

$$\int d^2 \hat{r} [3(\vec{S} \cdot \hat{r})^2 - S^2] = 0.$$

Portanto, as contribuições não nulas para os elementos de matriz se devem neste caso *apenas* ao “termo de contacto”, que contém a função delta de Dirac, associada ao campo magnético “interno” do próton. Como essas contribuições dependem do spin apenas através de S^2 , a matriz da perturbação é diagonal na representação em que S^2 é diagonal, seus autovalores sendo proporcionais ao valor da função de onda na origem. Eles podem ser calculados explicitamente sem dificuldade de forma direta, dando

$$\begin{aligned} & \langle 1s1/2\ 1/2\ S'M'_S | H_{SS} | 1s1/2\ 1/2\ SM_S \rangle = \\ &= -\frac{\mu_e \mu_p}{2\hbar^2} 4\pi(1-d) \frac{\mu^3}{\pi m_e^3 a_0^3} \int d^3 \vec{r} e^{-2\frac{\mu r}{m_e a_0}} \delta(\vec{r}) \hbar^2 \left(S(S+1) - \frac{3}{2} \right) \delta_{SS'} \delta_{M_S M'_S} \\ &= -\mu_e \mu_p 2(1-d) \frac{\mu^3}{m_e^3 a_0^3} \delta_{SS'} \delta_{M_S M'_S} \times \begin{cases} \frac{1}{2} & \text{para } S = 1 \\ -\frac{3}{2} & \text{para } S = 0 \end{cases}. \end{aligned} \quad (6.22)$$

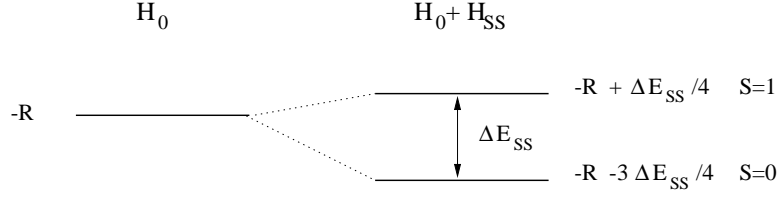


Figura 6.4: Efeito perturbativo da interação hiperfina para os estados $n = 1$ do hidrogênio. O zero da escala de energia é a energia de dissociação do hidrogênio.

Os autovalores assim obtidos dão as correções perturbativas de primeira ordem para as energias dos estados com $S = 1$ e com $S = 0$. O efeito da interação hiperfina entre os momentos magnéticos do elétron e do próton no estado fundamental do hidrogênio levam portanto a uma quebra parcial da degenerescência dos quatro níveis não perturbados, em que um estado com $S = 0$ tem sua energia de ligação aumentada enquanto tres outros estados, correspondentes aos tres valores possíveis de S_3 para $S = 1$, se tornam menos ligados. Esse resultado está representado na figura 6.4. Usando o fator de desmagnetização $d = 1/3$ que corresponde a supor uma distribuição esférica e uniforme de magnetização para o próton, e introduzindo os valores dos momentos magnéticos

$$\mu_e = -\frac{g_e e \hbar}{2m_e c} \quad \text{e} \quad \mu_p = \frac{g_p e \hbar}{2m_p c}$$

onde e é a carga elétrica do próton, é fácil ver que essas correções perturbativas podem ser escritas como

$$\Delta E_{1s}^S = \frac{2}{3} g_e g_p \alpha^2 R \frac{\mu^3}{m_e^2 m_p} \times \begin{cases} \frac{1}{2} & \text{para } S = 1 \\ -\frac{3}{2} & \text{para } S = 0 \end{cases}$$

onde $\alpha \simeq 1/137$ é a constante da estrutura fina e R é a constante de Rydberg. Como $m_p \sim 2000m_e$ e portanto $\mu \simeq m_e$ a escala de energia de separação entre os estados $S = 0$ e $S = 1$ difere por um fator $\simeq m_e/m_p \sim 1/2000$ da escala correspondente à estrutura fina do espectro. A diferença de energia entre os estados $S = 0$ e $S = 1$ é dada por

$$\Delta E_{SS} \equiv \Delta E_{1s}^1 - \Delta E_{1s}^0 = \frac{4g_e g_p}{3} \alpha^2 R \frac{\mu^3}{m_e^2 m_p} \simeq 5.868 \times 10^{-6} eV = h \times 1419 \text{ MHz}. \quad (6.23)$$

Essa frequência corresponde a um comprimento de onda de 21 cm, que tem uma importância considerável para a radioastronomia devido à presença residual de hidrogênio no espaço interestelar.

6.4 Campos eletromagnéticos externos. Efeitos Stark e Zeeman no hidrogênio.

A dinâmica de uma partícula *clássica* de massa M e carga elétrica e num campo eletromagnético externo, descrito em termos de um potencial vetor $\vec{A}(\vec{r}, t)$ e de um potencial escalar $\Phi(\vec{r}, t)$ através das relações usuais

$$\vec{E}(\vec{r}, t) = -\nabla\Phi(\vec{r}, t) - \frac{1}{c} \frac{\partial \vec{A}(\vec{r}, t)}{\partial t} \quad \text{e} \quad \vec{B}(\vec{r}, t) = \nabla \times \vec{A}(\vec{r}, t)$$

pode se definir através da lagrangeana⁶

$$L = \frac{M}{2} \dot{\vec{r}}^2 + \frac{e}{c} \dot{\vec{r}} \cdot \vec{A}(\vec{r}, t) - e\Phi(\vec{r}, t),$$

como pode ser facilmente verificado calculando as equações de Euler-Lagrange correspondentes

$$\frac{d}{dt} (M\dot{\vec{r}}) = \frac{e}{c} \frac{\partial \vec{A}}{\partial t} - e\nabla\Phi + \frac{e}{c} \dot{\vec{r}} \times \nabla \times \vec{A} = e\vec{E}(\vec{r}, t) + \frac{e}{c} \dot{\vec{r}} \times \vec{B}(\vec{r}, t)$$

em que o lado direito corresponde à força de Lorentz que age sobre a partícula. Vale a pena notar explicitamente que, para obter esse resultado, o argumento \vec{r} dos potenciais eletromagnéticos deve ser identificado com o vetor de posição da partícula. O momento canonicamente conjugado a \vec{r} é

$$\vec{p} = M\dot{\vec{r}} + \frac{e}{c} \vec{A}(\vec{r}, t)$$

de modo que a hamiltoniana correspondente à lagrangeana L é

$$H \equiv \vec{p} \cdot \dot{\vec{r}} - L = \frac{1}{2M} \left(\vec{p} - \frac{e}{c} \vec{A}(\vec{r}, t) \right)^2 + e\Phi(\vec{r}, t).$$

O adjetivo *externo* usado para os campos eletromagnéticos significa que em todas as equações precedentes a dependência de \vec{A} e de Φ com \vec{r} e t é dada *a priori*, sendo determinada por dispositivos estranhos ao sistema, que por isso deixa de ser *autônomo*. É claro que situações envolvendo campos externos *independentes do tempo* estão incluídas nesse tratamento como casos particulares em que \vec{A} e Φ não dependem de t .

A quantização canônica desse sistema se faz interpretando \vec{r} e \vec{p} como operadores vetoriais cujas componentes cartesianas satisfazem relações canônicas de comutação. Na representação

⁶v. e.g. E. J. Saletan e A. H. Cromer, *Theoretical Mechanics*, John Wiley & Sons, Inc., N. Y. (1971), pag. 51.

de Schrödinger, as componentes de \vec{r} são realizadas como operadores de multiplicação pelas componentes do vetor de posição enquanto o momento canônico \vec{p} é realizado em termos do operador diferencial $\hbar/i\nabla$, de modo que a equação de movimento quântica para a função de onda de uma partícula sujeita a campos eletromagnéticos externos descritos por \vec{A} e Φ é

$$\left[\frac{1}{2M} \left(\frac{\hbar}{i} \nabla - \frac{e}{c} \vec{A}(\vec{r}, t) \right)^2 + e\Phi(\vec{r}, t) \right] \psi(\vec{r}, t) = i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t}. \quad (6.24)$$

Em correspondência com o fato de que na equação de movimento clássica os potenciais devem ser calculados para \vec{r} correspondente à posição da partícula no instante t considerado, na equação de movimento quântica o argumento \vec{r} dos potenciais deve ser entendido como a variável dinâmica correspondente à posição da partícula. É importante observar que isso faz com que o momento $\vec{p} \rightarrow \hbar\nabla/i$ em geral não comute com $\vec{A}(\vec{r}, t)$. De fato, para uma função $\psi(\vec{r}, t)$ qualquer

$$\vec{p} \cdot \vec{A} \psi = \frac{\hbar}{i} \nabla \cdot (\vec{A} \psi) = \frac{\hbar}{i} (\nabla \cdot \vec{A}) \psi + \vec{A} \cdot \left(\frac{\hbar}{i} \nabla \psi \right) = \frac{\hbar}{i} (\nabla \cdot \vec{A}) \psi + \vec{A} \cdot \vec{p} \psi$$

de modo que $\vec{p} \cdot \vec{A} = \vec{A} \cdot \vec{p}$ apenas se $\nabla \cdot \vec{A} = 0$.

O fato de que os potenciais eletromagnéticos aparecem explicitamente na equação de movimento quântica coloca a questão do efeito de uma transformação de padrão (“gauge”)

$$\vec{A}(\vec{r}, t) \rightarrow \vec{A}' \equiv \vec{A} - \nabla \chi(\vec{r}, t), \quad \Phi(\vec{r}, t) \rightarrow \Phi' \equiv \Phi - \frac{1}{c} \frac{\partial \chi(\vec{r}, t)}{\partial t}, \quad (6.25)$$

onde $\chi(\vec{r}, t)$ é uma função real arbitrária, sobre as soluções da equação (6.24), já que os campos \vec{E} e \vec{B} não são afetados por essa transformação. Não é difícil verificar que as alterações introduzidas na equação de movimento pela transformação de padrão (6.25) podem ser compensadas por transformação unitária cujo efeito sobre as funções de onda é multiplica-las por um fator de fase

$$\psi(\vec{r}, t) \rightarrow \psi'(\vec{r}, t) \equiv e^{-\frac{ie}{\hbar c} \chi(\vec{r}, t)} \psi(\vec{r}, t).$$

De fato, introduzindo o momento transformado

$$\vec{p}' \equiv e^{\frac{ie}{\hbar c} \chi(\vec{r}, t)} \frac{\hbar}{i} \nabla e^{-\frac{ie}{\hbar c} \chi(\vec{r}, t)} = \frac{\hbar}{i} \nabla - \frac{e}{c} (\nabla \chi(\vec{r}, t)),$$

onde o último termo é um operador de multiplicação, e notando que a aplicação dessa mesma transformação a \vec{r} dá $\vec{r}' = \vec{r}$, resulta que, se $\psi(\vec{r}, t)$ é solução de (6.24), então $\psi'(\vec{r}, t)$ é solução da equação

$$\left[\frac{1}{2M} \left(\vec{p}' - \frac{e}{c} \vec{A}'(\vec{r}, t) \right)^2 + e\Phi'(\vec{r}, t) \right] \psi'(\vec{r}, t) = i\hbar \frac{\partial \psi'(\vec{r}, t)}{\partial t}.$$

Os efeitos de campos eletromagnéticos externos sobre sistemas mais complexos, envolvendo mais de uma partícula e forças internas, possivelmente também de origem eletromagnética, podem ser tratados dessa mesma forma, notando que os campos externos agem sobre cada uma das partículas carregadas presentes no sistema. No caso do átomo de hidrogênio, por exemplo, o Hamiltoniano de dois corpos “mínimo” a ser considerado quando em presença de campos eletromagnéticos externos é

$$H = \frac{1}{2m_p} \left(\frac{\hbar}{i} \nabla_p - \frac{e}{c} \vec{A}(\vec{r}_p, t) \right)^2 + \frac{1}{2m_e} \left(\frac{\hbar}{i} \nabla_e + \frac{e}{c} \vec{A}(\vec{r}_e, t) \right)^2 + e (\Phi(\vec{r}_p, t) - \Phi(\vec{r}_e, t)) - \frac{e^2}{|\vec{r}_e - \vec{r}_p|}$$

onde os índices p e e se referem ao próton e ao elétron respectivamente. O último termo provém do campo interno associado à interação coulombiana entre o próton e o elétron. É importante notar que as duas partículas carregadas que constituem o átomo de hidrogênio interagem com os campos externos descritos por \vec{A} e Φ . Esse hamiltoniano é “mínimo” no sentido que, a rigor, devem ser consideradas ainda as interações envolvendo os spins do próton e do elétron. Os termos correspondentes à interação spin-órbita (responsáveis pela estrutura fina do espectro) bem como os responsáveis pela estrutura hiperfina pertencem também à classe de forças internas e podem ser importados diretamente do tratamento anterior. Os termos *externos* envolvendo spins devem ser mais uma vez acrescentados explicitamente, já que esse tipo de grau de liberdade não aparece no procedimento que leva à equação (6.24). O que se suporá no que segue é que o acoplamento dos spins com os campos externos se dá através da interação dos momentos magnéticos associados a eles com o campo magnético externo. Isso pode ser representado por termos do tipo

$$-\vec{\mu} \cdot \vec{B} = -g \left(\frac{e}{2Mc} \right) \vec{s} \cdot \vec{B},$$

com $\vec{B} = \nabla \times \vec{A}$, onde e e M são respectivamente a carga e a massa da partícula associada ao spin, e g é o fator giromagnético apropriado.

6.4.1 Efeito Stark no hidrogênio.

Um primeiro problema que pode ser tratado com os ingredientes expostos acima é a perturbação de um átomo de hidrogênio por um campo *elétrico* externo \vec{E} , espacialmente uniforme e independente do tempo, conhecida como *efeito Stark*. Nesse caso os potenciais eletromagnéticos podem ser escolhidos como sendo

$$\vec{A} \equiv 0 \quad \text{e} \quad \Phi(\vec{r}) = -\vec{E} \cdot \vec{r}$$

de modo que o hamiltoniano a ser considerado é

$$H = H_0 - e\Phi(\vec{r}_p) + e\Phi(\vec{r}_e) = H_0 - e\vec{E} \cdot (\vec{r}_p - \vec{r}_e)$$

onde H_0 é o hamiltoniano que descreve o átomo de hidrogênio na ausência do campo externo. Os sinais dos termos envolvendo o campo externo $\Phi(\vec{r})$ são determinados pelos sinais das cargas do próton e do elétron respectivamente.

A escolha mais simples para H_0 é

$$H_0 = \frac{p_p^2}{2m_p} + \frac{p_e^2}{2m_e} - \frac{e^2}{|\vec{r}_e - \vec{r}_p|}$$

onde \vec{p}_p e \vec{p}_e correspondem respectivamente ao momento do próton e do elétron. Essa particular escolha ignora efeitos das interações internas envolvendo os spins (em particular a interação spin-órbita), e pode portanto não ser suficientemente realística em algumas situações. No entanto, ela é relevante quando os efeitos do campo externo são muito maiores que os efeitos devidos a essas forças, além de ser útil para ilustrar o esquema perturbativo para o tratamento do campo externo independentemente de complicações técnicas não essenciais. Os efeitos de spin à escala da estrutura fina do espectro serão discutidos mais adiante.

O termo perturbativo de H , associado ao campo elétrico externo, depende das variáveis dinâmicas do átomo apenas através do momento de dipolo elétrico

$$\vec{d} = e(\vec{r}_p - \vec{r}_e) \equiv -e\vec{r}$$

sendo $\vec{r} = \vec{r}_e - \vec{r}_p$ o vetor de posição do elétron relativamente ao próton⁷. Devido a essa propriedade, é conveniente utilizar os resultados já obtidos tratando o Hamiltoniano não perturbado H_0 em termos de variáveis relativas e de centro de massa

$$H_0 = \frac{P^2}{2M} + \frac{p^2}{2\mu} - \frac{e^2}{r}, \quad M = m_e + m_p, \quad \mu = \frac{m_e m_p}{m_e + m_p}. \quad (6.26)$$

A parte referente ao centro de massa tem o espectro contínuo de partículas livres de massa M e não é afetada pelo campo elétrico externo, como poderia ter sido antecipado tendo em vista a neutralidade do átomo como um todo. O problema se reduz, portanto, a obter os efeitos perturbativos do termo de interação

⁷O próton está sendo usado como origem para a definição do vetor de posição relativa do sistema próton-elétron na suposição de que a mesma convenção seja usada no tratamento da dinâmica relativa não perturbada (v. eqs. (5.33) e (6.27) abaixo). Desse modo as variáveis angulares que aparecem nas funções de onda relativas não perturbadas serão as mesmas que aparecem em \vec{d} através de \vec{r} .

$$H' = -\vec{E} \cdot \vec{d} = e\vec{E} \cdot \vec{r} \quad (6.27)$$

sobre o espectro discreto (5.36) e (5.37) do Hamiltoniano que descreve a dinâmica relativa do sistema próton-elétron.

1. Perturbação do estado fundamental. O estado fundamental $|\phi_{100}\rangle$ é o único estado não degenerado do espectro da parte de H_0 referente ao movimento relativo. Ele corresponde aos números quânticos $n = 1, l = m = 0$, com a notação usada nas equações (5.36) e (5.37). A correção de primeira ordem à energia desse estado, devida à interação (6.27) é portanto

$$e_{100}^{(1)} = \langle \phi_{100} | H' | \phi_{100} \rangle = \vec{E} \cdot \int d^3r \phi_{100}^*(\vec{r}) e\vec{r} \phi_{100}(\vec{r}) \equiv -\vec{E} \cdot \langle \vec{d} \rangle_{100}.$$

O valor médio $\langle \vec{d} \rangle_{100}$ é chamado *momento de dipolo elétrico* do estado $|\phi_{100}\rangle$, e é fácil ver que o seu valor é zero. De fato, o integrando que aparece na definição desse valor médio troca de sinal com a substituição de \vec{r} por $-\vec{r}$; como a integral é estendida a todo o espaço, isso implica que $\langle \vec{d} \rangle_{100} = -\langle \vec{d} \rangle_{100} = 0$. A consequência imediata desse fato é a de que a correção perturbativa de primeira ordem em H' ao autovalor do estado fundamental de H_0 se anula, de modo que a correção perturbativa não nula de ordem mais baixa é pelo menos de segunda ordem, e portanto pelo menos proporcional a E^2 , não havendo correção que depende linearmente do módulo do campo externo aplicado.

O que deve ser feito nesse caso é, portanto, calcular a correção perturbativa de segunda ordem $e_{100}^{(2)}$. De acordo com as expressões gerais de teoria de perturbações para estados não degenerados essa correção envolve uma soma sobre estados intermediários $|\phi_{nlm}\rangle$, e é dada por

$$e_{100}^{(2)} = \sum_{nlm \neq 100} \frac{|\langle \phi_{100} | H' | \phi_{nlm} \rangle|^2}{E_1 - E_n} = \sum_{nlm \neq 100} \frac{|\vec{E} \cdot \langle \vec{d} \rangle_{100, nlm}|^2}{E_1 - E_n} \quad (6.28)$$

onde aparecem agora os *momentos de dipolo de transição*

$$\langle \vec{d} \rangle_{100, nlm} = - \int d^3r \phi_{100}^*(\vec{r}) e\vec{r} \phi_{nlm}(\vec{r}).$$

Como no caso do momento de dipolo do estado fundamental, os momentos de dipolo de transição serão nulos sempre que o integrando trocar de sinal com a substituição de \vec{r} por $-\vec{r}$. Como $\phi_{100}(-\vec{r}) = \phi_{100}(\vec{r})$, e tendo em conta a presença do fator \vec{r} proveniente do operador de dipolo elétrico, as contribuições não nulas para a correção de segunda ordem só podem provir de estados $\phi_{nlm}(\vec{r})$ tais que $\phi_{nlm}(-\vec{r}) \neq \phi_{nlm}(\vec{r})$. É fácil ver, por inspeção direta das funções de onda (5.37), que tais estados existem. Já considerando os estados mais próximos em energia do estado fundamental (e que são portanto os estados para os quais o denominador de energia em

(6.28) tem o menor valor absoluto), que são os estados com $n = 2$, $|\phi_{200}\rangle$ e $|\phi_{21m}\rangle$, $m = -1, 0, 1$, resulta que

$$\phi_{200}(-\vec{r}) = \phi_{200}(\vec{r})$$

devido ao fato de que essa função de onda depende apenas de $|\vec{r}|$ mas, por outro lado,

$$\phi_{21m}(-\vec{r}) = \frac{u_{21}(\rho)}{r} Y_{1m}(\pi - \theta, \varphi + \pi) = -\frac{u_{21}(\rho)}{r} Y_{1m}(\theta, \varphi) = -\phi_{21m}(\vec{r}), \quad \rho \propto r = |\vec{r}|.$$

A função de onda $\phi_{21m}(\vec{r})$, portanto, troca de sinal com a substituição de \vec{r} por $-\vec{r}$, o que permite que $\langle \vec{d} \rangle_{100,21m}$ possa ser diferente de zero.

O problema de inventariar os estados $|\phi_{nlm}\rangle$ capazes de contribuir para a correção de segunda ordem (6.28) pode, na realidade, ser consideravelmente simplificado através da consideração explícita de certas propriedades de simetria do Hamiltoniano que descreve o movimento relativo do sistema próton-elétron. Seja, de fato, o *operador de inversão espacial* \mathcal{P} definido pela propriedade

$$\langle \vec{r} | \mathcal{P} | \psi \rangle = \langle -\vec{r} | \psi \rangle = \psi(-\vec{r})$$

para qualquer $|\vec{r}\rangle$ e qualquer vetor de estado $|\psi\rangle$ (associado à função de onda $\langle \vec{r} | \psi \rangle$) do espaço de Hilbert da dinâmica relativa. Uma propriedade simples mas importante desse operador é a de que $\mathcal{P}^2 = \hat{1}$, isto é, o quadrado de \mathcal{P} é o operador identidade, o que significa que esse operador é igual ao seu inverso, $\mathcal{P}^{-1} = \mathcal{P}$, e implica em que seus possíveis autovalores são apenas $+1$ e -1 . Para verifica-la basta notar que

$$\langle \vec{r} | \mathcal{P}^2 | \psi \rangle = \langle -\vec{r} | \mathcal{P} | \psi \rangle = \langle \vec{r} | \psi \rangle$$

para quaisquer $|\vec{r}\rangle$ e $|\psi\rangle$. A natureza dos autovetores $|\psi_p\rangle$, $p = \pm 1$ de \mathcal{P} pode ser facilmente reconhecida notando que

$$\begin{aligned} \langle \vec{r} | \mathcal{P} | \psi_p \rangle &= p \langle \vec{r} | \psi_p \rangle \\ &= \langle -\vec{r} | \psi_p \rangle \quad \text{logo} \quad \langle -\vec{r} | \psi_p \rangle = p \langle \vec{r} | \psi_p \rangle \end{aligned}$$

e as autofunções $\langle \vec{r} | \psi_p \rangle$ são funções pares (para $p = +1$) e ímpares (para $p = -1$) com relação à substituição de \vec{r} por $-\vec{r}$. Além disso, o operador de inversão é hermiteano (e portanto, devido à propriedade anterior, também unitário). De fato, utilizando a definição de \mathcal{P} e mudanças simples de variável de integração resulta que

$$\begin{aligned}
\langle \psi' | \mathcal{P}^\dagger | \psi \rangle &= \langle \psi | \mathcal{P} | \psi' \rangle^* = \int d^3 r \left(\langle \psi | \vec{r} \rangle \langle \vec{r} | \mathcal{P} | \psi' \rangle \right)^* = \int d^3 r \left(\langle \psi | \vec{r} \rangle \langle -\vec{r} | \psi' \rangle \right)^* \\
&= \int d^3 r \left(\langle \psi | -\vec{r} \rangle \langle \vec{r} | \psi' \rangle \right)^* = \int d^3 r \langle \psi' | \vec{r} \rangle \langle -\vec{r} | \psi \rangle \\
&= \langle \psi' | \mathcal{P} | \psi \rangle
\end{aligned}$$

para quaisquer vetores de estado $|\psi\rangle$ e $|\psi'\rangle$. O observável associado ao operador \mathcal{P} é chamado *paridade*, e os autovetores de \mathcal{P} são portanto estados *de paridade definida*. Autovetores associados aos autovalores $p = +1$ e $p = -1$ são estados de paridade *positiva* e *negativa* respectivamente.

O efeito do operador de inversão espacial sobre variáveis dinâmicas pode ser obtido através das propriedades de transformação das variáveis canônicas

$$\mathcal{P} \vec{q} \mathcal{P}^\dagger = -\vec{r} \quad \text{e} \quad \mathcal{P} \vec{p} \mathcal{P}^\dagger = -\vec{p}$$

que podem também ser verificadas sem dificuldade com as definições e resultados anteriores. Para a segunda dessas relações, por exemplo,

$$\langle \vec{r} | \mathcal{P} \vec{p} \mathcal{P}^\dagger | \psi \rangle = \langle -\vec{r} | \vec{p} \mathcal{P} | \psi \rangle = -\frac{\hbar}{i} \vec{\nabla} \langle -\vec{r} | \mathcal{P} | \psi \rangle = -\frac{\hbar}{i} \vec{\nabla} \langle \vec{r} | \psi \rangle = \langle \vec{r} | -\vec{p} | \psi \rangle$$

para quaisquer $|\vec{r}\rangle$ e $|\psi\rangle$. Em particular, o efeito de \mathcal{P} sobre o Hamiltoniano relativo é

$$\mathcal{P} \left(\frac{p^2}{2\mu} - \frac{e^2}{r} \right) \mathcal{P}^{-1} = \frac{p^2}{2\mu} - \frac{e^2}{r}$$

ou ainda

$$\left[\mathcal{P}, \frac{p^2}{2\mu} - \frac{e^2}{r} \right] = 0,$$

o que significa que \mathcal{P} pode ser diagonalizado simultaneamente com o Hamiltoniano relativo ou, em outras palavras, que as autofunções ϕ_{E_k} do Hamiltoniano relativo podem ser escolhidas de forma a serem também autofunções de \mathcal{P} , ou seja, de modo que

$$\mathcal{P} \phi_{E_k}(\vec{r}) = \phi_{E_k}(-\vec{r}) = \phi_{E_k}(\vec{r}) \quad \text{ou} \quad \mathcal{P} \phi_{E_k}(\vec{r}) = \phi_{E_k}(-\vec{r}) = -\phi_{E_k}(\vec{r}).$$

A presença de degenerescência no espectro (5.36) e (5.37) do Hamiltoniano relativo não permite, a rigor, identificar imediatamente as autofunções ϕ_{nlm} também como autofunções de \mathcal{P} . No entanto, a forma (5.37) dessas autofunções, juntamente com a propriedade das harmônicas esféricas

$$\langle \hat{r} | \mathcal{P} | lm \rangle = \langle -\hat{r} | lm \rangle = Y_{lm}(\pi - \theta, \varphi + \pi) = (-1)^l Y_{lm}(\theta, \varphi) \quad (6.29)$$

mostra que, de fato,

$$\langle \vec{r} | \mathcal{P} | \phi_{nlm} \rangle = \langle -\vec{r} | \phi_{nlm} \rangle = (-1)^l \phi_{nlm}(\vec{r})$$

isto é, os autovetores $|\phi_{nlm}\rangle$ são também autovetores de \mathcal{P} com autovalor $(-1)^l$. Vale a pena notar explicitamente que os dois ângulos que servem de argumento às harmônicas esféricas em (6.29) são os ângulos polares esféricos de $\hat{r} \equiv \vec{r}/r$, o que justifica o efeito de \mathcal{P} nessa relação, pois os ângulos correspondentes a $-\hat{r}$ são $\pi - \theta$, $\varphi + \pi$.

Esse resultado generaliza para uma autofunção qualquer $\phi_{nlm}(\vec{r})$ as propriedades obtidas explicitamente para ϕ_{200} e para ϕ_{21m} , mostrando que apenas autofunções com l ímpar podem contribuir para a correção de segunda ordem para a energia do estado fundamental (6.28). Além disso, ele permite obter um resultado mais geral com relação a elementos de matriz do tipo $\langle \phi_{n'l'm'} | \mathcal{O}_p | \phi_{nlm} \rangle$, onde \mathcal{O}_p é um operador *de paridade definida*, isto é, tal que

$$\mathcal{P} \mathcal{O}_p \mathcal{P}^\dagger = p \mathcal{O}_p, \quad p = +1 \text{ ou } -1.$$

O operador de dipolo elétrico e o momento \vec{p} são exemplos de operadores de paridade ímpar, e o Hamiltoniano relativo é um exemplo de operador de paridade par. Os resultados anteriores permitem escrever tais elementos de matriz como

$$\langle \phi_{n'l'm'} | \mathcal{O}_p | \phi_{nlm} \rangle = \langle \phi_{n'l'm'} | \mathcal{P}^\dagger \mathcal{P} \mathcal{O}_p \mathcal{P}^\dagger \mathcal{P} | \phi_{nlm} \rangle = p(-1)^{l+l'} \langle \phi_{n'l'm'} | \mathcal{O}_p | \phi_{nlm} \rangle,$$

de modo que $\langle \phi_{n'l'm'} | \mathcal{O}_p | \phi_{nlm} \rangle = 0$ a menos que $p(-1)^{l+l'} = 1$. Isso significa que $l + l'$ deve ser *par* para $p = +1$, e *ímpar* para $p = -1$. Em particular, o valor médio de qualquer operador de paridade ímpar em um estado de paridade bem definida *qualquer* (por exemplo, o valor médio do operador de dipolo elétrico em um estado $|\phi_{nlm}\rangle$ qualquer) é zero.

As restrições devidas a considerações de paridade não são, no entanto, as únicas relevantes para determinar o não anulamento dos elementos de matriz que aparecem na correção de segunda ordem (6.28) para a energia do estado fundamental do hidrogênio. Neste caso particular existem de fato restrições mais fortes provenientes de questões ligadas ao momento angular. Elas podem ser obtidas de forma simples examinando a parte angular da integral que define o momento de dipolo de transição $\langle \vec{d} \rangle_{100, nlm}$. Esta envolve, além da harmônica esférica $Y_{lm}(\theta, \varphi)$ associada ao estado ϕ_{nlm} , a dependência angular implícita no caráter vetorial do operador de dipolo elétrico \vec{d} . Uma forma de explicitar esta última dependência angular é escrever as componentes cartesianas x, y, z de \vec{r} em termos de $|\vec{r}|$ e dos ângulos esféricos θ e φ

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta.$$

Outra forma mais conveniente consiste em exprimir a dependência angular em termos de harmônicas esféricas como

$$rY_{11}(\theta, \varphi) = -\sqrt{\frac{3}{8\pi}}(x + iy), \quad rY_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}}z, \quad rY_{1-1}(\theta, \varphi) = \sqrt{\frac{3}{8\pi}}(x - iy). \quad (6.30)$$

Essas relações mostram que as componentes cartesianas de \vec{r} podem ser escritas como combinações lineares das tres harmônicas esféricas $Y_{1m}(\theta, \varphi)$. Conseqüentemente, as componentes cartesianas de $\langle \vec{d} \rangle_{100, nlm}$ são dadas em termos de integrais angulares que podem ser calculadas simplesmente utilizando as propriedades de ortogonalidade das harmônicas esféricas

$$\begin{aligned} \int_{-1}^{+1} d \cos \theta \int_0^{2\pi} Y_{1m}(\theta, \varphi) Y_{lm'}(\theta, \varphi) &= \int_{-1}^{+1} d \cos \theta \int_0^{2\pi} (-1)^m Y_{1-m}^*(\theta, \varphi) Y_{lm'}(\theta, \varphi) \\ &= (-1)^m \delta_{1l} \delta_{-mm'} \end{aligned}$$

mostrando que *apenas estados com $l = 1$* darão contribuições não nulas. É claro que esse resultado depende crucialmente do caráter vetorial do operador de dipolo elétrico bem como fato de que a função de onda $\phi_{100}(\vec{r})$ é independente das coordenadas angulares, e que ele inclui a restrição proveniente de considerações de paridade. Este último fato se relaciona à relação estreita, indicada na equação (6.29), existente entre a paridade e o valor de l . Formas mais gerais das restrições associadas a questões de momento angular serão discutidas no capítulo 9.

Tendo em conta as restrições devidas a questões de paridade e momento angular, o alcance das somas envolvidas na correção de segunda ordem para a energia do estado fundamental do hidrogênio (6.28) pode, portanto, ser consideravelmente reduzido. De fato, definindo a direção z de forma que ela coincida com a direção do campo externo aplicado \vec{E} , essa correção fica expressa em termos da componente z do momento de dipolo de transição o que, de acordo com (6.30), seleciona apenas os estados $|\phi_{n10}\rangle$ como possíveis estados intermediários. Dessa forma a soma tripla que aparece em (6.28) fica reduzida a uma única soma, sobre o número quântico principal n

$$e_{100}^{(2)} = E^2 \sum_{n>1} \frac{|\langle d_z \rangle_{100, n10}|^2}{E_1 - E_n}$$

com $\langle d_z \rangle_{100, n10}$ dados em termos de integrais radiais (v. equação (5.37))

$$\langle d_z \rangle_{100, n10} = -\frac{e}{\sqrt{3}} \int_0^\infty dr u_{10}^*(\rho) r u_{n1}(\rho), \quad \rho \propto r.$$

O deslocamento perturbativo da energia do estado fundamental é portanto proporcional a E^2 , e é por isso chamado *efeito Stark quadrático*.

2. Perturbação do nível $n = 2$, efeito Stark linear. No modelo simples em que interações envolvendo os spins do próton e do elétron são ignoradas, o nível $n = 2$ do átomo de hidrogênio contém estados degenerados com $l = 0$ e $l = 1$. Os spins podem ser ignorados completamente pois, nessas condições, a única consequência de sua existência é a de produzir quatro réplicas idênticas (correspondentes a cada um dos quatro estados independentes que podem ser construídos com dois spins $1/2$) do problema em que o próton e o elétron são considerados partículas sem spin, bastando portanto obter a solução deste último problema.

De acordo com as prescrições gerais de teoria de perturbações, a presença de degenerescência exige a diagonalização do termo perturbativo H' (eq (6.27)) no sub-espço de estados degenerados, que no caso envolve quatro estados, a saber, o estado $|\phi_{200}\rangle$ e os tres estados $|\phi_{21m}\rangle$, $m = -1, 0, +1$. Os elementos da matriz a ser diagonalizada são portanto

$$\langle \phi_{2l'm'} | -\vec{E} \cdot \vec{d} | \phi_{2lm} \rangle$$

onde l e l' são 0 ou 1, e $-l, l' \leq m, m' \leq l, l'$. As restrições ligadas à paridade, no entanto, implicam no anulamento dos elementos de matriz com $l = l'$, de modo os elementos de matriz diferentes de zero são apenas

$$\langle \phi_{200} | -\vec{E} \cdot \vec{d} | \phi_{21m} \rangle$$

e seus conjugados complexos. O problema secular pode ainda ser simplificado tomando o eixo z na direção do campo elétrico externo \vec{E} , pois nesse caso

$$\vec{E} \cdot \vec{d} \rightarrow E d_z$$

e, em virtude de (6.30), apenas $\langle \phi_{200} | -E d_z | \phi_{210} \rangle$ e seu conjugado complexo são diferentes de zero. Esse elemento de matriz pode ser facilmente calculado levando ao resultado

$$\gamma \equiv -E \langle \phi_{200} | d_z | \phi_{210} \rangle = \frac{Ee}{\sqrt{3}} \int_0^\infty dr u_{20}^*(\rho) r u_{21}(\rho). \quad (6.31)$$

Para evitar complicações técnicas desnecessárias, convém observar que com a escolha usual de fases (cf. equação (5.37)) as funções $u_{nl}(\rho)$ são *reais*⁸, de modo que sem perda de generalidade é possível tomar γ como sendo também real. O problema secular resultante é portanto

⁸A condição de que as funções $u_{nl}(\rho)$ sejam reais deixa ainda uma indefinição de sinal. O sinal de γ dependerá da particular escolha de sinais que tenha sido feita para estas funções. Essa arbitrariedade no sinal de γ é devidamente compensada, no entanto, na solução do problema secular, através dos fatores $\gamma/|\gamma|$ que aparecem nos autovetores perturbados, v. eq. (6.32) abaixo.

$$\det \begin{pmatrix} -\lambda & \gamma & 0 & 0 \\ \gamma & -\lambda & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix} = \lambda^2(\lambda^2 - \gamma^2) = 0$$

em que as linhas e colunas foram ordenadas de modo a corresponder aos estados ϕ_{200} , ϕ_{210} , ϕ_{211} e ϕ_{21-1} respectivamente. A forma dessa matriz depende da escolha feita para o eixo z , que leva ao desacoplamento dos estados correspondentes às duas últimas linhas e colunas. Fisicamente, isso significa que estados com $l = 1$ e $m = \pm 1$ *na direção do campo elétrico externo* não são afetados em primeira ordem por esse campo, o que faz com que a equação secular possa ser fatorada em duas equações quadráticas

$$\lambda^2(\lambda^2 - \gamma^2) = 0 \rightarrow \lambda^2 - \gamma^2 = 0 \quad \text{ou} \quad \lambda^2 = 0$$

com as soluções (que correspondem às correções de primeira ordem para a energia não perturbada E_2)

$$\lambda_1 = |\gamma|, \quad \lambda_2 = -|\gamma|, \quad \lambda_3 = \lambda_4 = 0.$$

Os autovetores normalizados correspondentes a esse autovalores são respectivamente

$$\begin{aligned} |\lambda_1\rangle &= \frac{1}{\sqrt{2}} \left(|\phi_{200}\rangle + \frac{\gamma}{|\gamma|} |\phi_{210}\rangle \right), \quad |\lambda_2\rangle = \frac{1}{\sqrt{2}} \left(|\phi_{200}\rangle - \frac{\gamma}{|\gamma|} |\phi_{210}\rangle \right), \\ |\lambda_3\rangle &= |\phi_{211}\rangle, \quad |\lambda_4\rangle = |\phi_{21-1}\rangle. \end{aligned} \quad (6.32)$$

O esquema de níveis perturbado é portanto aquele representado na figura 6.5. É importante notar que, neste caso, o deslocamento em energia dos níveis afetados pelo campo externo depende *linearmente* do módulo desse campo pois, como se vê na equação (6.31), γ é proporcional a E , um efeito conhecido como *efeito Stark linear*. Ele é uma consequência direta da degenerescência dos estados de paridades diferentes $|\phi_{200}\rangle$ e $|\phi_{210}\rangle$ no espectro não perturbado. Os estados perturbados que sofrem deslocamentos em energia proporcionais a E são combinações lineares desses dois estados, e portanto não tem paridade definida. O valor médio do operador de dipolo elétrico d_z nesses estados pode portanto ser diferente de zero, e um cálculo simples, usando o fato de que o valor médio de d_z em estados de paridade definida é zero, mostra que

$$\frac{1}{2} \left(\langle \phi_{200} | \pm \frac{\gamma}{|\gamma|} \langle \phi_{210} | \right) d_z \left(|\phi_{200}\rangle \pm \frac{\gamma}{|\gamma|} |\phi_{210}\rangle \right) = \mp \frac{|\gamma|}{E}. \quad (6.33)$$

$$\begin{array}{c}
\frac{p^2}{2\mu} + \frac{e^2}{r} \quad \quad \quad + H' \\
\\
\begin{array}{ccc}
& & \begin{array}{l} \frac{E_2 + |\gamma|}{\sqrt{2}} \left(|\phi_{200}\rangle + \frac{\gamma}{|\gamma|} |\phi_{210}\rangle \right) \\ \hline E_2 \\ \hline \frac{E_2 - |\gamma|}{\sqrt{2}} \left(|\phi_{200}\rangle - \frac{\gamma}{|\gamma|} |\phi_{210}\rangle \right) \end{array} \\
|\phi_{200}\rangle, |\phi_{21m}\rangle \xrightarrow{E_2} & & |\phi_{211}\rangle, |\phi_{21-1}\rangle
\end{array}
\end{array}$$

Figura 6.5: Efeito perturbativo de um campo elétrico externo na direção z sobre os estados $n = 2$ do hidrogênio quando interações envolvendo os spins do próton e do elétron são ignorados.

Classicamente, a energia de interação de um dipolo elétrico de momento \vec{d}^{cl} com um campo elétrico uniforme \vec{E} na direção z é dada por $\Delta E = -\vec{E} \cdot \vec{d}^{\text{cl}} = -E d_z^{\text{cl}}$. Essa energia *depende linearmente do módulo do campo elétrico aplicado*, é positiva quando o dipolo tem sentido contrário ao do campo, e negativa quando o dipolo tem o mesmo sentido que o campo. No caso dos estados $|\lambda_1\rangle$ e $|\lambda_2\rangle$, por outro lado, $\Delta E_{1,2} = \pm|\gamma| = -E\langle d_z \rangle_{1,2}$, de acordo com o resultado (6.33), o que mostra que esses estados se comportam como se dotados de dipolos clássicos de momento $|\gamma|/E$ orientados anti-paralela e paralelamente ao campo externo, respectivamente. Estados desse tipo são chamados *estados polares*. De uma forma geral, no caso de Hamiltonianos não perturbados que comutam com o operador de inversão \mathcal{P} , cujos autovetores podem portanto ser escolhidos de forma a serem autovetores simultâneos de \mathcal{P} (em outras palavras, estados de paridade definida), a ocorrência de estados polares depende essencialmente da existência de degenerescência entre estados de paridades diferentes, que se superpõe sob a ação do campo externo, exatamente como no caso dos estados com $n = 2$ do modelo simples do átomo de hidrogênio. Um exemplo disso são as chamadas *moléculas polares*, como é o caso bem conhecido da molécula de água, que se comportam como se dotadas de momentos de dipolo elétrico quando submetidas à ação de campos elétricos externos, embora seus Hamiltonianos sejam invariantes sob a operação de inversão espacial.

3. Inclusão da interação spin-órbita no cálculo do efeito Stark linear dos estados $n = 2$. Mesmo ignorando interações hiperfinas, que envolvem a interação do momento magnético associado ao spin do próton com os momentos magnéticos eletrônicos, as energias dos estados com $n = 2$ do átomo de hidrogênio diferem suficientemente da energia comum E_2 predita pelo modelo simples de forma a alterar de forma importante o efeito Stark associado a campos externos fracos, no sentido que a energia de interação $|\gamma|$, eq. (6.31), seja menor que os deslocamentos típicos da estrutura fina do espectro. Como indicado nas figuras 6.2 e 6.3, uma

descrição minimamente realística desses níveis na escala da estrutura fina exige, além da inclusão dos efeitos da interação spin-órbita, a inclusão também de outras correções relativísticas de mesma ordem em v/c , que são a correção ao termo de energia cinética (6.15) e o termo de Darwin (6.16). Isso leva, de fato à situação representada na figura 6.3, em que os níveis $2s_{1/2}$ e $2p_{1/2}$ são degenerados. Como esses níveis têm paridades definidas mas diferentes, a introdução de um campo elétrico externo produz um deslocamento Stark linear no campo aplicado. Para calculá-lo, no entanto, é preciso levar em conta o acoplamento spin-órbita dos estados envolvidos. Os demais estados com $n = 2$, devido a restrições de paridade, sofrerão apenas deslocamentos em segunda ordem de teoria de perturbações, que dependem quadraticamente do módulo do campo externo aplicado, como no caso do estado fundamental.

As correções perturbativas de ordem mais baixa para os estados degenerados $|\phi_{20\frac{1}{2}m}\rangle$ e $|\phi_{21\frac{1}{2}m}\rangle$ são obtidas diagonalizando a matriz cujos elementos são

$$\langle \phi_{20\frac{1}{2}m} | -\vec{E} \cdot \vec{d} | \phi_{21\frac{1}{2}m'} \rangle \equiv \gamma_{mm'}. \quad (6.34)$$

Essa é também uma matriz 4×4 , tendo em vista que tanto m como m' podem assumir os valores $\pm 1/2$. Os seus elementos $\gamma_{mm'}$ podem ser calculados explicitando a estrutura das funções de onda envolvidas

$$\begin{aligned} \langle \vec{r}, s | \phi_{20\frac{1}{2}m} \rangle &= \frac{u_{20\frac{1}{2}}(r)}{r} Y_{00}(\hat{r}) \chi_{\frac{1}{2}m}(s) \\ \langle \vec{r}, s | \phi_{21\frac{1}{2}m} \rangle &= \frac{u_{21\frac{1}{2}}(r)}{r} \sum_{m_l} C_{m_l m' - m_l}^{1 \frac{1}{2} \frac{1}{2}} Y_{1m_l}(\hat{r}) \chi_{\frac{1}{2}m' - m_l}(s) \end{aligned}$$

onde as $u_{nlj}(r)$ são funções radiais (possivelmente dependentes também de j), $Y_{lm}(\hat{r})$ são harmônicas esféricas, $\chi_{\frac{1}{2}m}(s)$ são funções de onda de spin $1/2$ e os coeficientes C são coeficientes de Clebsch-Gordan que acoplam o momento angular orbital $l = 1$ com o spin $s = 1/2$ dando o momento angular total $j = 1/2$. Analogamente ao que foi escrito no caso em que $l = 1$, no caso $l = 0$ há também que considerar a questão do acoplamento do momento angular orbital com o spin, dando o valor pretendido do momento angular total. Neste caso, porém, é fácil ver que os coeficientes de Clebsch-Gordan envolvidos $C_{0 \frac{1}{2} m'}^{0 \frac{1}{2} \frac{1}{2}}$ são todos iguais a 1. Escolhendo, como anteriormente, o eixo z na direção do campo externo aplicado, $\vec{E} = E\hat{z}$, e substituindo essas expressões em (6.34) se obtém

$$\gamma_{mm'} = Ee\sqrt{\frac{4\pi}{3}} R_{20\frac{1}{2}, 21\frac{1}{2}} \sum_{m_l} C_{m_l m' - m_l}^{1 \frac{1}{2} \frac{1}{2}} \sum_s \chi_{\frac{1}{2}m}^\dagger(s) \chi_{\frac{1}{2}m' - m_l}(s) \int d\hat{r} Y_{00}^*(\hat{r}) Y_{10}(\hat{r}) Y_{1m_l}(\hat{r})$$

onde aparece a integral radial

$$R_{20\frac{1}{2}, 21\frac{1}{2}} = \int_0^\infty dr u_{20\frac{1}{2}}^*(r) r u_{21\frac{1}{2}}(r)$$

e o operador de dipolo foi escrito em termos de $Y_{10}(\hat{r})$ usando (6.30). A soma sobre os valores da variável de spin s pode ser feita imediatamente, usando as propriedades de ortonormalidade das funções de spin, com o resultado

$$\sum_s \chi_{\frac{1}{2}m}^\dagger(s) \chi_{\frac{1}{2}m'-m_l}(s) = \delta_{m, m'-m_l};$$

a integração sobre os ângulos de \vec{r} também pode ser feita neste caso usando apenas propriedades de ortogonalidade das harmônicas esféricas, pois $Y_{00}(\hat{r}) = 1/\sqrt{4\pi}$ é na realidade independente dos ângulos e portanto

$$\int d\hat{r} Y_{00}^*(\hat{r}) Y_{10}(\hat{r}) Y_{1m_l}(\hat{r}) = \frac{1}{\sqrt{4\pi}} \int d\hat{r} Y_{10}^*(\hat{r}) Y_{1m_l}(\hat{r}) = \frac{1}{\sqrt{4\pi}} \delta_{0m_l}.$$

Neste último cálculo a mudança de $Y_{10}(\hat{r})$ em $Y_{10}^*(\hat{r})$ é possível sem qualquer outra alteração em vista da propriedade geral de conjugação complexa das harmônicas esféricas $Y_{1m}^*(\hat{r}) = (-1)^m Y_{1-m}(\hat{r})$. O resultado final para os elementos de matriz é portanto

$$\gamma_{mm'} = \frac{Ee}{\sqrt{3}} R_{20\frac{1}{2}, 21\frac{1}{2}} C_{0\frac{1}{2}m\frac{1}{2}}^1 \delta_{mm'},$$

o que mostra que apenas os elementos diagonais γ_{mm} (onde m é a projeção do momento angular total sobre \hat{z} , definido como tendo a direção e o sentido do campo elétrico aplicado \vec{E} , são diferentes de zero. Os coeficientes de Clebsch-Gordan que aparecem na expressão final são

$$C_{0\frac{1}{2}m\frac{1}{2}}^1 = \begin{cases} -\frac{1}{\sqrt{3}} & \text{para } m = \frac{1}{2} \\ \frac{1}{\sqrt{3}} & \text{para } m = -\frac{1}{2} \end{cases}$$

de modo que

$$\gamma_{\frac{1}{2}-\frac{1}{2}} = \gamma_{-\frac{1}{2}\frac{1}{2}} = 0$$

$$\gamma_{\frac{1}{2}\frac{1}{2}} = -\gamma_{-\frac{1}{2}-\frac{1}{2}} = \frac{Ee}{3} R_{20\frac{1}{2}, 21\frac{1}{2}} \equiv \gamma.$$

No contexto de um tratamento perturbativo de primeira ordem, as funções radiais a utilizar para o cálculo da integral radial são na realidade as funções $u_{nl}(r)$ do modelo simples utilizado na seção anterior. Desse modo, o valor de γ obtido neste caso difere por um fator $1/\sqrt{3}$ do que corresponde ao caso tratado anteriormente. Esse fator provém do coeficiente de Clebsch-Gordan responsável pelo acoplamento do momento angular orbital $l = 1$ com o spin eletrônico, dando o valor $j = 1/2$ para o momento angular total. A intensidade faltante é a que está associada aos estados com momento angular $j = 3/2$, afastados do grupo degenerado pelos efeitos da interação spin-órbita.

Com esses resultados, as correções perturbativas de primeira ordem à energia são obtidas como soluções da equação secular

$$\det \begin{pmatrix} -\lambda & -\gamma & 0 & 0 \\ -\gamma & -\lambda & 0 & 0 \\ 0 & 0 & -\lambda & \gamma \\ 0 & 0 & \gamma & -\lambda \end{pmatrix} = (\lambda^2 - \gamma^2)^2 = 0$$

onde as linhas e colunas correspondem à ordenação de estados $(l, m) = (0, -1/2), (1, -1/2), (0, 1/2)$ e $(1, 1/2)$, escolhida por conveniência de forma que os estados acoplados através de H' ocupem posições contíguas na matriz a ser diagonalizada. Essa equação secular tem duas raízes *duplas* $\lambda_i = \pm|\gamma|$ às quais correspondem os autovetores

$$\begin{aligned} \lambda_1 &= +|\gamma| \rightarrow \begin{cases} |\lambda_1\rangle = \frac{1}{\sqrt{2}} \left[|\phi_{20\frac{1}{2}+\frac{1}{2}}\rangle + \frac{\gamma}{|\gamma|} |\phi_{21\frac{1}{2}+\frac{1}{2}}\rangle \right] \\ |\lambda'_1\rangle = \frac{1}{\sqrt{2}} \left[|\phi_{20\frac{1}{2}-\frac{1}{2}}\rangle - \frac{\gamma}{|\gamma|} |\phi_{21\frac{1}{2}-\frac{1}{2}}\rangle \right] \end{cases} \\ \lambda_2 &= -|\gamma| \rightarrow \begin{cases} |\lambda_2\rangle = \frac{1}{\sqrt{2}} \left[|\phi_{20\frac{1}{2}\frac{1}{2}}\rangle - \frac{\gamma}{|\gamma|} |\phi_{21\frac{1}{2}\frac{1}{2}}\rangle \right] \\ |\lambda'_2\rangle = \frac{1}{\sqrt{2}} \left[|\phi_{20\frac{1}{2}-\frac{1}{2}}\rangle + \frac{\gamma}{|\gamma|} |\phi_{21\frac{1}{2}-\frac{1}{2}}\rangle \right] \end{cases}. \end{aligned}$$

O campo elétrico externo portanto remove apenas parcialmente a degenerescência dos quatro estados, dando lugar a um esquema de dois níveis duplamente degenerados cuja separação depende linearmente da magnitude do campo. Os estados perturbados são estados polares que correspondem a momentos de dipolo elétrico $d^{\text{el}} = |\gamma|/E$ paralelos e anti-paralelos ao campo externo para os estados associados aos autovalores λ_1 e λ_2 respectivamente. Os dois estados de cada tipo diferem entre si pela estrutura de spin.

Deve ser lembrado ainda que esses resultados supõe a degenerescência dos níveis $2s1/2$ e $2p1/2$ e portanto devem ser modificados devido ao deslocamento de Lamb ΔE_{Lamb} quando os

valores do campo externo sejam suficientemente baixos, no sentido de que $|\gamma| < \Delta E_{\text{Lamb}}$ (v. problema 6.8).

4. Efeito Stark e ionização. É bem conhecido o fato de que a aplicação de um campo elétrico externo em um átomo pode levar a um rápido processo de ionização quando o campo for suficientemente intenso. Ele é de fato utilizado até mesmo para a construção de detectores de ionização, que são capazes de discriminar entre estados excitados diferentes do átomo, através dos diferentes tempos de ionização que correspondem a cada um deles (cf. seção 2.4). Na realidade, é muito fácil ver, através de um argumento variacional simples, que *qualquer* campo externo uniforme permite a ocorrência de processos de ionização (embora nem sempre rápidos), desde que mantido até distâncias suficientemente grandes do átomo. O argumento é também útil para que se possa ter uma imagem não perturbativa do que ocorre nas situações usualmente associadas ao efeito Stark.

Voltando à dinâmica interna do átomo de hidrogênio submetido ao campo elétrico externo uniforme \vec{E} , escrita em termos do Hamiltoniano relativo

$$H = \frac{p^2}{2\mu} - \frac{e^2}{r} + e\vec{E} \cdot \vec{r} \equiv K + V$$

e utilizando a função de onda de prova gaussiana (não normalizada), dependente do vetor \vec{a} como parâmetro

$$\langle \vec{r} | \varphi_{\vec{a}} \rangle = e^{-\frac{(\vec{r}-\vec{a})^2}{2b^2}}$$

é fácil verificar que o valor médio do termo de energia cinética de H

$$\langle K \rangle_{\varphi_{\vec{a}}} = \frac{\langle \varphi_{\vec{a}} | \frac{p^2}{2\mu} | \varphi_{\vec{a}} \rangle}{\langle \varphi_{\vec{a}} | \varphi_{\vec{a}} \rangle}$$

é positivo e independente de \vec{a} , que corresponde à localização do máximo da gaussiana. Uma forma simples de obter esse resultado é notar que $\langle K \rangle_{\varphi_{\vec{a}}}$ é a energia cinética média de um estado coerente de um oscilador harmônico de massa μ e parâmetro b . Esta é, por sua vez, igual à energia cinética média do estado fundamental desse oscilador que, pelo teorema do virial, é a metade da energia do estado fundamental. Chamando ω a frequência do oscilador, $b = \sqrt{\hbar/\mu\omega}$, de modo que

$$\langle K \rangle_{\varphi_{\vec{a}}} = \frac{\hbar\omega}{4} = \frac{\hbar^2}{4\mu b^2}.$$

O valor médio da energia potencial, por outro lado, é dado pela soma de dois termos, que dependem respectivamente do potencial coulombiano do sistema elétron-próton e do potencial associado ao campo externo:

$$\langle V \rangle_{\varphi_{\vec{a}}} = \frac{\langle \varphi_{\vec{a}} | -\frac{e^2}{r} | \varphi_{\vec{a}} \rangle + \langle \varphi_{\vec{a}} | e\vec{E} \cdot \vec{r} | \varphi_{\vec{a}} \rangle}{\langle \varphi_{\vec{a}} | \varphi_{\vec{a}} \rangle}.$$

Ao contrário do que ocorre no caso de energia cinética, esses dois termos dependem de \vec{a} . O primeiro deles depende apenas de $a = |\vec{a}|$ e tende a zero com $1/|a|$ à medida que $|a|$ cresce, mas o segundo depende também da direção do vetor \vec{a} . Uma forma simples de ver isso é tomar $a \gg b$, pois nesse caso, a menos da dispersão dos valores de \vec{r} para os quais a gaussiana é apreciavelmente diferente de zero (que é da ordem de b em torno de $\vec{r} = \vec{a}$) os dois termos podem ser aproximados respectivamente por

$$\frac{\langle \varphi_{\vec{a}} | -\frac{e^2}{r} | \varphi_{\vec{a}} \rangle}{\langle \varphi_{\vec{a}} | \varphi_{\vec{a}} \rangle} \sim -\frac{e^2}{a} \quad \text{e} \quad \frac{\langle \varphi_{\vec{a}} | e\vec{E} \cdot \vec{r} | \varphi_{\vec{a}} \rangle}{\langle \varphi_{\vec{a}} | \varphi_{\vec{a}} \rangle} \sim e\vec{E} \cdot \vec{a}.$$

Portanto, escolhendo \vec{a} na direção oposta a \vec{E} , isto é, $e\vec{E} \cdot \vec{a} = -eEa$, é possível tornar arbitrariamente *negativo* o valor médio de H na função de prova, $\langle H \rangle_{\varphi_{\vec{a}}} = \langle K \rangle_{\varphi_{\vec{a}}} + \langle V \rangle_{\varphi_{\vec{a}}}$, sempre que a puder ser tomado como sendo arbitrariamente grande. Como esse valor médio é um limite superior para a energia do estado fundamental segue que, nessas condições, o sistema descrito por H *não possui um estado fundamental*. A razão disso é que a energia potencial associada ao campo externo se torna arbitrariamente negativa na direção oposta ao campo, o que permite o escape do elétron nessa direção, tornando o espectro de H na realidade puramente *contínuo* na região de energias negativas onde ocorrem os estados discretos do átomo não perturbado.

Esse resultado evidentemente coloca em questão o significado do cálculo perturbativo, que dá valores definidos para a correção de energia de níveis *discretos*, juntamente com as respectivas funções de onda corrigidas. As diferenças radicais existentes entre o espectro contínuo e o conjunto discreto de estados cujos parâmetros são obtidos através de teoria de perturbações podem porém ser conciliadas através de argumentos baseados nas relações de incerteza. De fato, a natureza contínua do espectro do átomo perturbado em energias negativas está associada, em termos qualitativos, à instabilidade do sistema atômico como sistema ligado, que se traduz numa probabilidade de ionização por unidade de tempo diferente de zero que é introduzida pela dinâmica do sistema para um estado inicialmente preparado com características de um estado atômico ligado (que deve, em particular, ser representado por um vetor de estado normalizado $|I\rangle$). Para valores pequenos do campo externo aplicado, no entanto, o valor dessa taxa de ionização é extremamente pequeno ou, equivalentemente, o tempo médio de vida τ do estado inicial não ionizado $|I\rangle$ é extremamente grande no sentido que \hbar/τ é uma energia até muitas ordens de grandeza menor que os efeitos de estrutura fina no espectro atômico.

Por outro lado, o estado normalizado $|I\rangle$ pode também em princípio ser expandido nos autoestados $|\psi_E\rangle$ exatos que constituem o espectro contínuo do átomo perturbado, o que permite escrever a sua evolução temporal sob a forma

$$e^{-\frac{i}{\hbar}Ht}|I\rangle = \int dE e^{-\frac{i}{\hbar}Et} |\psi_E\rangle \langle \psi_E|I\rangle.$$

A ordem de grandeza do tempo característico τ associado a essa evolução temporal está relacionado com a dispersão em energia ΔE da distribuição $|\langle \psi_E|I\rangle|^2$, definida da forma usual como

$$\Delta E^2 = \int dE E^2 |\langle \psi_E|I\rangle|^2 - \left(\int dE E |\langle \psi_E|I\rangle|^2 \right)^2,$$

através da relação de incerteza $\tau \Delta E \sim \hbar$, ou seja

$$\Delta E \sim \frac{\hbar}{\tau}.$$

Isso significa, para τ grande no sentido especificado acima, que apenas os estados do espectro contínuo pertencentes a uma estreita faixa de largura ΔE , até muitas ordens de grandeza menor que os efeitos de estrutura fina no espectro atômico, poderão contribuir significativamente para a expansão de $|I\rangle$.

O que se vê desse modo é que processos de ionização de vida média τ muito longa devem poder ser associados a estados iniciais preparados $|I\rangle$ cuja expansão em termos dos autoestados exatos do espectro contínuo depende essencialmente apenas dos estados localizados em faixas muito bem definidas em energia, e que tem portanto, eles próprios, energia muito bem definida. Permanece então aberta a possibilidade de que o cálculo perturbativo de autovetores e autovalores forneça aproximações para os estados do tipo $|I\rangle$ e para a sua respectiva localização em energia, embora o argumento dado não seja, evidentemente, uma prova de tal resultado.

Uma outra forma de exprimir o resultado do cálculo variacional acima é dizendo que o processo de ionização através da aplicação de um campo elétrico externo uniforme é um processo “sem limiar”, isto é, não existe um valor mínimo, diferente de zero, do módulo do campo aplicado abaixo do qual a ionização não poderia ocorrer. Estritamente falando, a ionização é possível com *qualquer* campo uniforme não nulo, e efeito de um aumento do campo aplicado é apenas o de diminuir o tempo de ionização, que é em geral diferente para diferentes estados iniciais. É aliás este último fato que desempenha um papel central no funcionamento dos detectores de ionização mencionados na seção 2.4, nos quais diferentes estados atômicos são distinguidos através de seus diferentes tempos de ionização. Como a ocorrência ou não de um determinado processo de ionização em condições externas de campo e num intervalo de tempo *dados* é, segundo a mecânica quântica, um processo intrinsecamente estatístico, o poder discriminatório

de tais detectores é também intrinsecamente limitado estatisticamente, e de uma forma tão mais grave quanto menor seja a diferença entre os tempos de ionização envolvidos em seu funcionamento.

6.4.2 Efeito Zeeman no hidrogênio.

O chamado efeito Zeeman corresponde, em geral, aos efeitos de um campo magnético externo uniforme sobre o espectro atômico. Um campo magnético externo, uniforme \vec{B} pode ser descrito por meio do potencial vetor

$$\vec{A}(\vec{r}) = \frac{1}{2} \vec{B} \times \vec{r}$$

pois, como pode ser facilmente verificado, $\nabla \times \vec{A}(\vec{r}) = \vec{B}$. Ignorando, por simplicidade, os efeitos do recuo do próton, o hamiltoniano que descreve a interação do modelo simples (isto é, ignorando efeitos de spin) do átomo de hidrogênio com esse campo é

$$H_B = \frac{1}{2m_e} \left(\vec{p} + \frac{e}{c} \vec{A} \right)^2 - \frac{e^2}{r}$$

onde m_e é a massa do elétron, de carga $-e$. Como o potencial vetor $\vec{A}(\vec{r})$ satisfaz a relação $\nabla \cdot \vec{A}(\vec{r}) = 0$, ele comuta com \vec{p} , de modo que o primeiro termo de H_0 pode ser transformado em

$$\frac{1}{2m_e} \left(\vec{p} + \frac{e}{c} \vec{A} \right)^2 = \frac{p^2}{2m_e} + \frac{e}{2m_e c} \vec{B} \times \vec{r} \cdot \vec{p} + \frac{e^2}{8m_e c^2} (\vec{B} \times \vec{r})^2$$

onde o termo linear no campo \vec{B} pode ser expresso em termos do momento angular orbital \vec{L} usando o fato de que

$$\vec{B} \times \vec{r} \cdot \vec{p} = \vec{B} \cdot \vec{r} \times \vec{p} = \vec{B} \cdot \vec{L}.$$

Desse modo a forma resultante para H_0 é

$$H_B = \frac{p^2}{2m_e} - \frac{e^2}{r} + \frac{e}{2m_e c} \vec{B} \cdot \vec{L} + \frac{e^2}{8m_e c^2} (\vec{B} \times \vec{r})^2 \equiv H_0 + H'_B$$

onde

$$H_0 = \frac{p^2}{2m_e} - \frac{e^2}{r}$$

é simplesmente o hamiltoniano correspondente ao modelo simples do átomo de hidrogênio.

Como foi dito na seção 6.4, para incluir efeitos associados aos spins é preciso acrescentar a esse hamiltoniano, além dos termos correspondentes à interação spin-órbita H_{LS} (incluindo eventualmente a correção relativística à energia cinética H_K e o termo de Darwin H_D) e à interação hiperfina H_{SS} , termos correspondentes à interação dos spins do elétron $\vec{S}^{(e)}$ e do próton $\vec{S}^{(p)}$ com o campo externo que, usando a notação (6.18) para os momentos magnéticos, são dados por

$$H_S^{\text{ext}} = -\mu_e \frac{\vec{B} \cdot \vec{S}^{(e)}}{\hbar} - \mu_p \frac{\vec{B} \cdot \vec{S}^{(p)}}{\hbar} = g_e \left(\frac{e}{2m_e c} \right) \vec{B} \cdot \vec{S}^{(e)} - g_p \left(\frac{e}{2m_p c} \right) \vec{B} \cdot \vec{S}^{(p)}.$$

Isso leva finalmente ao hamiltoniano

$$H = H_0 + H_{LS} + H_K + H_D + H_{SS} + H'_B + H_S^{\text{ext}}$$

que contém um número apreciável de termos. Como discutido a seguir, no entanto, esses termos correspondem a efeitos de ordem de grandeza bastante diferente, o que pode ser usado para simplificar o tratamento de situações específicas.

Os termos independentes do campo externo já foram discutidos anteriormente, quando foi observado que os efeitos das interações hiperfinas H_{SS} são desprezíveis na escala dos efeitos de estrutura fina introduzidos pela interação spin-órbita H_{LS} . A razão básica disso é o valor pequeno do momento magnético do próton, proporcional ao magneton nuclear. Pelo mesmo motivo, a razão dos efeitos devidos aos termos de H_S^{ext} correspondentes respectivamente ao spin do próton e ao spin do elétron contém um fator $m_e/m_p \sim 1/2000$, o que torna os efeitos associados a $\vec{S}^{(p)}$ muito menores que os associados a $\vec{S}^{(e)}$, para um valor dado de \vec{B} . Para valores do campo magnético tais que os efeitos do termo eletrônico de H_S^{ext} tem a ordem de grandeza dos efeitos de estrutura fina, tanto o termo nuclear de H_S^{ext} como as interações hiperfinas são desprezíveis.

Resta, portanto, neste caso, examinar os efeitos dos dois termos de H'_B . O termo que envolve o momento angular orbital, proporcional a $\vec{B} \cdot \vec{L}$, tem claramente a mesma ordem de grandeza que o termo eletrônico de H_S^{ext} , e deve portanto ser considerado juntamente com este último. A ordem de grandeza dos efeitos do termo quadrático em \vec{B} pode ser estimada escrevendo

$$\frac{e^2}{8m_e c^2} (\vec{B} \times \vec{r})^2 \rightarrow \frac{e\hbar B}{m_e c} \times \frac{e\hbar B}{m_e c} \times \frac{1}{8} \frac{\langle r \rangle^2}{a_0^2} \frac{m_e a_0^2}{\hbar^2}$$

onde $\langle r \rangle$ é o raio médio do estado atômico considerado, e $a_0 = \hbar^2/m_e e^2$ é o primeiro raio de Bohr do hidrogênio. O último fator pode ser reescrito sob a forma mais conveniente

$$\frac{1}{8} \frac{\langle r \rangle^2}{a_0^2} \frac{m_e a_0^2}{\hbar^2} = \frac{1}{8} \frac{\langle r \rangle^2}{a_0^2} \frac{\alpha^{-2}}{m_e c^2}$$

onde $\alpha = e^2/\hbar c \simeq 1/137$ é a constante de estrutura fina. Os dois primeiros fatores tem a ordem de grandeza do termo linear em \vec{B} para valores do momento angular orbital da ordem de \hbar . Quando os efeitos deste termo têm a ordem de grandeza dos efeitos associados à estrutura fina, isto é,

$$\frac{e\hbar B}{m_e c} \sim \alpha^2 R,$$

onde R é a constante de Rydberg, resulta

$$\frac{e^2}{8m_e c^2} (\vec{B} \times \vec{r})^2 \rightarrow (\alpha^2 R)^2 \times \frac{1}{8} \frac{\langle r \rangle^2}{a_0^2} \frac{\alpha^{-2}}{m_e c^2} = \frac{1}{8} \frac{\langle r \rangle^2}{a_0^2} \frac{R}{m_e c^2} \alpha^2 R \simeq 10^{-5} \frac{\langle r \rangle^2}{a_0^2} \alpha^2 R.$$

Portanto, para estados pouco excitados, para os quais o raio médio não é muito maior que o raio de Bohr, e cujo momento angular orbital não é muito maior que \hbar , o termo quadrático é também desprezível na escala de energia dos efeitos de estrutura fina. Essa estimativa mostra ainda que, para tais estados pouco excitados do átomo de hidrogênio, existem valores de \vec{B} para os quais os efeitos do termo linear são *maiores* que os efeitos da interação spin-órbita, enquanto os efeitos do termo quadrático ainda são relativamente pequenos. Esse limite de campos intensos corresponde, de fato, ao chamado “efeito Paschen-Back”.

As situações específicas típicas a considerar podem portanto ser classificadas inicialmente em dois grupos distintos. No primeiro desses grupos, o valor do campo magnético externo é tal que o termo linear de acoplamento ao momento angular orbital é suficientemente grande para que as interações hiperfinas e os efeitos ligados ao spin do próton sejam desprezíveis, mas ainda suficientemente pequeno para que o termo quadrático em \vec{B} possa ser ignorado. Os termos de H que é preciso considerar nestes casos são

$$H \simeq H_0 + H_{LS} + H_K + H_D + \frac{e}{2m_e c} \vec{B} \cdot (\vec{L} + g_e \vec{S}^{(e)}). \quad (6.35)$$

A característica mais relevante deste hamiltoniano é a existência de termos perturbativos que envolvem as componentes do momento angular orbital \vec{L} e do spin do elétron $\vec{S}^{(e)}$ na direção do campo externo aplicado. Isso corresponde a um acoplamento dessas variáveis dinâmicas *com uma direção externa dada*, que deve competir com o seu acoplamento mútuo através do termo de interação spin-órbita, e leva à consideração de duas situações distintas. A primeira delas corresponde ao limite de “campos fracos”, no sentido que os efeitos da interação spin-órbita

dominam os efeitos dos acoplamentos externos, ao passo que a segunda é a situação de “campos fortes”, dominada pelos acoplamentos externos, que corresponde ao efeito Paschen-Back.

Um segundo grupo de situações, cuja relevância cresceu consideravelmente nos últimos anos devido aos desenvolvimentos experimentais ligados ao aprisionamento de um gás de átomos alcalinos em armadilhas magnéticas, envolve o efeito Zeeman da estrutura hiperfina do estado fundamental de átomos alcalinos (incluindo o átomo de hidrogênio!). Os campos magnéticos externos envolvidos nestas situações são muito mais fracos, de modo que a sua interação com o sistema atômico compete com os efeitos das interações hiperfinas. O hamiltoniano a ser considerado no caso particular do estado fundamental do átomo de hidrogênio pode prescindir tanto do termo de interação spin-órbita como do termo linear de acoplamento ao campo externo que envolve o momento angular orbital \vec{L} , dado que nesse caso o momento angular orbital é zero. É essencial, no entanto, incluir as interações hiperfinas, de modo que os termos de H a considerar são

$$H \simeq H_0 + H_{SS} + H_K + H_D - \mu_e \frac{\vec{B} \cdot \vec{S}^{(e)}}{\hbar} - \mu_p \frac{\vec{B} \cdot \vec{S}^{(p)}}{\hbar}. \quad (6.36)$$

O tratamento perturbativo das características resultantes em cada um desses três casos é desenvolvido nos itens seguintes.

1. Campos fracos. No limite de campos fracos das situações descritas pelo hamiltoniano (6.35), os efeitos do termo proporcional ao campo externo são pequenos na escala de energia da estrutura fina. Neste caso é possível partir do problema “não perturbado” definido por $H_1 \equiv H_0 + H_{LS} + H_K + H_D$ e incluir perturbativamente os efeitos devidos ao termo envolvendo o campo externo \vec{B} . Escolhendo o eixo de quantização do momento angular na direção do campo magnético externo este termo assume a forma

$$h' = \frac{eB}{2m_e c} (L_3 + 2S_3^{(e)}) \quad (6.37)$$

onde foi usado explicitamente o valor $g_e = 2$ do fator giromagnético do elétron.

Os autovetores de H_1 podem ser escolhidos como autovetores simultâneos de L^2 , S^2 , de J^2 (com $\vec{J} = \vec{L} + \vec{S}$) e de J_3 , como no tratamento da estrutura fina na seção 6.3.1. Os valores da energia dependem portanto dos números quânticos l , $s = 1/2$ e j , além do número quântico principal n , sendo que os dois valores possíveis de j para cada valor de n e l correspondem a energias que diferem pelos efeitos do acoplamento spin-órbita. O caso dos níveis com $n = 2$ está ilustrado na figura 6.3. As autofunções correspondentes a tais níveis tem portanto a forma

$$\langle \vec{r}, s | \phi_{nl\frac{1}{2}jm} \rangle = \frac{u_{nl\frac{1}{2}j}(r)}{r} \sum_{m_s} C_{m-m_s}^l \frac{1}{m_s} \frac{j}{m} Y_{l,m-m_s}(\hat{r}) \chi_{\frac{1}{2}m_s}(s). \quad (6.38)$$

Para obter as correções perturbativas de ordem mais baixa devidas a h' é preciso ter em conta a degenerescência desses estados. Como valores diferentes de n correspondem a energias muito diferentes na escala da estrutura fina, basta considerar um único valor do número quântico principal. Os elementos de matriz relevantes são, desse modo,

$$\langle \phi_{nl'\frac{1}{2}j'm'} | h' | \phi_{nl\frac{1}{2}jm} \rangle = \frac{eB}{2m_e c} \langle \phi_{nl'\frac{1}{2}j'm'} | L_3 + 2S_3^{(e)} | \phi_{nl\frac{1}{2}jm} \rangle.$$

Como h' comuta com L^2 , S^2 e J_3 , esses elementos de matriz se anulam a menos que $l = l'$ e $m = m'$, mas como h' não comuta com J^2 , há elementos de matriz diferentes de zero para valores diferentes de j . No entanto, estados que diferem apenas pelo valor de j tem energias diferentes na escala da estrutura fina, e podem ser tratados como não degenerados no caso presente de campos fracos. As correções de primeira ordem se reduzem portanto aos valores médios

$$\Delta E_{nl\frac{1}{2}jm} = \frac{eB}{2m_e c} \langle \phi_{nl\frac{1}{2}jm} | L_3 + 2S_3^{(e)} | \phi_{nl\frac{1}{2}jm} \rangle = \frac{eB}{2m_e c} (\hbar m + \langle \phi_{nl\frac{1}{2}jm} | S_3^{(e)} | \phi_{nl\frac{1}{2}jm} \rangle)$$

pois $L_3 + S_3^{(e)} = J_3$, cujo autovalor é $\hbar m$.

O elemento de matriz de $S_3^{(e)}$ pode ser calculado explicitamente usando a expressão (6.38). De fato

$$\begin{aligned} \langle \phi_{nl\frac{1}{2}jm} | S_3^{(e)} | \phi_{nl\frac{1}{2}jm} \rangle &= \sum_{m_s, m'_s} C_{m-m_s}^l \frac{1}{2} \frac{j}{m_s} C_{m-m'_s}^l \frac{1}{2} \frac{j}{m} \int d\hat{r} Y_{lm-m_s}^*(\hat{r}) Y_{lm-m'_s}(\hat{r}) \hbar m_s \delta_{m_s, m'_s} \\ &= \frac{\hbar}{2} \left[\left(C_{m-\frac{1}{2}}^l \frac{1}{2} \frac{j}{m} \right)^2 - \left(C_{m+\frac{1}{2}}^l \frac{1}{2} \frac{j}{m} \right)^2 \right]. \end{aligned}$$

Os valores dos coeficientes de Clebsch-Gordan que aparecem nessa expressão são dados pelas expressões analíticas conhecidas

$$\begin{aligned} C_{m-m_s}^l \frac{1}{2} \frac{j}{m_s} : \quad & m_s = \frac{1}{2} & m_s = -\frac{1}{2} \\ j = l + \frac{1}{2} & \left[\frac{l+m+\frac{1}{2}}{2l+1} \right]^{\frac{1}{2}} & \left[\frac{l-m+\frac{1}{2}}{2l+1} \right]^{\frac{1}{2}} \\ j = l - \frac{1}{2} & - \left[\frac{l-m+\frac{1}{2}}{2l+1} \right]^{\frac{1}{2}} & \left[\frac{l+m+\frac{1}{2}}{2l+1} \right]^{\frac{1}{2}} \end{aligned}$$

que podem, em particular, ser obtidas sem dificuldades excessivas a partir do método indicado na seção 6.2.1. Com esses valores resulta então

$$\langle \phi_{nl\frac{1}{2}jm} | S_3^{(e)} | \phi_{nl\frac{1}{2}jm} \rangle = \frac{\hbar m}{2} \frac{j(j+1) - l(l+1) + \frac{3}{4}}{j(j+1)}$$

de modo que as energias do multipletto Zeeman são dadas, no caso de campos fracos e em primeira ordem no campo externo, por

$$E_{nl\frac{1}{2}jm} \simeq E_{nlj} + \frac{e\hbar B}{2m_e c} \left[1 + \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \right] m.$$

O efeito do campo externo é portanto separar uniformemente os diferentes valores de m correspondentes a cada valor de j , de uma forma que depende de l e de j através do coeficiente de m no último termo, conhecido como *coeficiente de Landé*. Os estados correspondentes a cada valor de m são freqüentemente chamados, por esse motivo, *sub-estados magnéticos* do nível j .

2. Campos fortes: efeito Paschen-Back. Quando o campo externo é suficientemente intenso para que os efeitos do termo (6.37) sejam maiores que os devidos à interação spin órbita, um tratamento possível é partir de um hamiltoniano “não perturbado” dado por $H_2 \equiv H_0 + H_K + H_D + h'$ e tratar perturbativamente os efeitos de H_{LS} . Essa situação é a que corresponde ao chamado efeito Paschen-Back. As autofunções de H_2 têm a forma

$$\langle \vec{r}, s | \phi_{nlm_l\frac{1}{2}m_s} \rangle = \frac{u_{nl}(r)}{r} Y_{lm_l}(\hat{r}) \chi_{\frac{1}{2}m_s}(s)$$

e os seus respectivos autovalores⁹ são

$$E_{nlm_l m_s} = E_{nl} + \frac{e\hbar B}{2m_e c} (m_l + 2m_s).$$

Em geral haverá valores diferentes de m_l e m_s que conduzem ao mesmo valor de $m_l + 2m_s$, o que implica que esse espectro ainda contém estados degenerados (por exemplo, no caso de um

⁹Na realidade existe, por assim dizer, uma contaminação perturbativa na forma de escrever estes resultados, que consiste em escrever a função radial $u_{nl}(r)$ e a parte E_{nl} dos autovalores que não depende explicitamente de B como dependendo apenas dos números quânticos n e l . O fato de que o campo magnético introduz uma dependência explícita de m_l e m_s nos autovalores implica em que as funções de onda radiais correspondentes aos estados ligados devam depender também desses números quânticos, e devem, a rigor, ser escritas como $u_{nlm_l m_s}(r)$. Concomitantemente a esse fato, os valores da energia para os quais a função radial é normalizável adquirem uma dependência implícita (isto é, além da dependência explícita que aparece no termo dependente de B) de m_l e m_s . Desse modo, E_{nl} a rigor depende também de m_l e m_s . Estas dependências implícitas de m_l e de m_s são, nos regimes perturbativos considerados aqui, relativamente muito menores que a dependência explícita desses números quânticos, o que pode servir de justificativa para que tenham sido ignoradas.

nível correspondente a $l = 1$, $m_l = 1$ e $m_s = -1/2$ levam à mesma energia que $m_l = -1$ e $m_s = 1/2$).

O termo de interação spin-órbita pode ser escrito em termos de L^2 , S^2 e de J^2 , que não comuta com L_3 e com S_3 e portanto tem elementos de matriz não diagonais diferentes de zero entre os estados $|\phi_{nlm_l\frac{1}{2}m_s}\rangle$. Ele comuta, porém, com $J_3 = L_3 + S_3$, de modo que os elementos de matriz diferentes de zero são aqueles que envolvem dois estados diferentes com o mesmo autovalor de J_3 . Como os estados não perturbados degenerados são aqueles que tem o mesmo valor de $L_3 + 2S_3$ e portanto valores *diferentes* de $L_3 + S_3$, todos esses elementos de matriz não diagonais envolvem estados não perturbados de energias diferentes, e portanto contribuem apenas em segunda ordem de teoria de perturbações. Os efeitos de primeira ordem se reduzem portanto aos valores esperados

$$\Delta E_{nlm_lm_s}^{(1)} = \langle \phi_{nlm_l\frac{1}{2}m_s} | H_{LS} | \phi_{nlm_l\frac{1}{2}m_s} \rangle$$

que podem ser calculados de forma simples escrevendo o termo de interação spin-órbita como

$$H_{LS} = \frac{1}{2m_e^2c^2} \frac{1}{r} \frac{dV}{dr} \left(\frac{S_+^{(e)}L_+ + S_-^{(e)}L_-}{2} + S_3^{(e)}L_3 \right). \quad (6.39)$$

Os termos envolvendo $S_{\pm}^{(e)}$ e L_{\pm} mudam os valores de m_l e de m_s , e portanto não contribuem para tais valores esperados. A única contribuição provém portanto do último termo e é dada por

$$\Delta E_{nlm_lm_s}^{(1)} = \frac{\hbar^2}{2m_e^2c^2} m_l m_s \int_0^\infty dr |u_{nl}(r)|^2 \frac{1}{r} \frac{dV}{dr}.$$

Essas correções de primeira ordem removem apenas parcialmente a degenerescência do espectro não perturbado. De fato, as correções $\Delta E_{n11-\frac{1}{2}}^{(1)}$ e $\Delta E_{n1-1\frac{1}{2}}^{(1)}$ que correspondem, como visto acima, a dois níveis degenerados, são por sua vez iguais.

Os resultados do item anterior e os deste item se referem a duas situações limite, em que os efeitos devidos ao campo externo são respectivamente *muito menores* ou *muito maiores* que a separação entre níveis da estrutura fina com valores dados de n e l mas com valores diferentes de j (embora, no segundo caso, os efeitos de \vec{B} devam ser ainda muito menores que a separação de níveis com valores diferentes de n). Eles não descrevem, portanto, o comportamento do espectro na região de valores intermediários do campo magnético, na qual os seus efeitos são da mesma ordem que os efeitos de estrutura fina. Um tratamento alternativo que cobre também essa região intermediária, além de permanecer válido nos dois limites já estudados, consiste em diagonalizar os dois termos de interação H_{LS} e h' (v. eq. (6.37)) no multiplete degenerado constituído pelos auto-estados $|\phi_{nlm_l\frac{1}{2}m_s}\rangle$ de H_0 , com n e l fixos. A interação que envolve o

campo externo h' contribui apenas para a diagonal da matriz secular, enquanto a interação spin-órbita tem também elementos não diagonais entre estados m_l , $m_s = 1/2$ e $m_l + 1$, $m_s = -1/2$ que correspondem ao mesmo autovalor de J_3 , e que podem ser facilmente calculados usando a forma (6.39) de H_{LS} . A matriz secular resultante pode ser montada de forma a ter apenas blocos 2×2 ao longo da diagonal principal, podendo portanto ser facilmente diagonalizada. Esse procedimento é explorado no Problema 6.9 e também no item seguinte, para o tratamento do efeito Zeeman da estrutura hiperfina dos estados $n = 1$ do átomo de hidrogênio.

3. Efeito Zeeman da estrutura hiperfina dos estados $n = 1$. Os efeitos da aplicação de um campo magnético externo sobre os quatro estados de spin (tendo agora em conta não apenas o spin do elétron, mas também o spin do próton) correspondentes ao orbital $n = 1$, $l = 0$ do átomo de hidrogênio são descritos pelo hamiltoniano (6.36). De uma forma análoga ao que acontece com o momento angular orbital \vec{L} e com o spin do elétron $\vec{S}^{(e)}$ no caso do efeito Zeeman da estrutura fina, existe neste caso uma competição entre o acoplamento mútuo dos dois spins $\vec{S}^{(p)}$ e $\vec{S}^{(e)}$ através de H_{SS} e o seu acoplamento com o campo externo \vec{B} , descrito pelos dois últimos termos de H . Usando os resultados da seção 6.3.2 é possível neste caso escrever H_{SS} como

$$H_{SS} = \frac{\Delta E_{SS}}{\hbar^2} \vec{S}^{(e)} \cdot \vec{S}^{(p)}$$

onde ΔE_{SS} é a quantidade calculada em (6.23), que caracteriza a diferença de energia entre os estados de spin total $S = 1$ e $S = 0$ no caso em que $\vec{B} = 0$. Os termos H_0 , H_K e H_D em (6.36) têm apenas o papel de determinar uma energia \tilde{E}_{10} comum aos quatro estados de spin, de modo que o hamiltoniano de interesse para esta caso pode ser reduzido a

$$H \simeq \tilde{E}_{10} + \frac{\Delta E_{SS}}{\hbar^2} \vec{S}^{(e)} \cdot \vec{S}^{(p)} - \mu_e B \frac{S_3^{(e)}}{\hbar} - \mu_p B \frac{S_3^{(p)}}{\hbar}$$

onde a direção de quantização do momento angular foi tomada como a direção do campo magnético externo.

Ainda em analogia ao efeito Zeeman da estrutura fina, existe neste caso um limite de campos fracos, no qual dominam os efeitos associados a ΔE_{SS} com correções perturbativas devidas aos termos restantes, e um limite “Paschen-Back”, no qual dominam os efeitos do campo externo, com correções perturbativas devidas a ΔE_{SS} . Um tratamento que dá conta não só desses dois limites mas também das situações intermediárias consiste em diagonalizar H no subespaço dos quatro estados de spin escritos na representação em que $S_3^{(e)}$ e $S_3^{(p)}$ são diagonais, que podem ser denotados simplesmente como $|m_s^{(e)}, m_s^{(p)}\rangle$, com $m_s^{(e)}$ e $m_s^{(p)}$ podendo ser $\pm 1/2$. O único termo de H que tem elementos de matriz não diagonais entre esses estados é o termo

de interação spin-spin, que contém o operador $\vec{S}^{(e)} \cdot \vec{S}^{(p)}$. De fato, esse operador pode ser escrito como

$$\vec{S}^{(e)} \cdot \vec{S}^{(p)} = \frac{S_+^{(e)} S_-^{(p)} + S_-^{(e)} S_+^{(p)}}{2} + S_3^{(e)} S_3^{(p)}$$

de modo que

$$\left\langle \frac{1}{2}, -\frac{1}{2} \right| \vec{S}^{(e)} \cdot \vec{S}^{(p)} \left| -\frac{1}{2}, \frac{1}{2} \right\rangle = \frac{\hbar^2}{2} \quad \text{e} \quad \langle m_s^{(e)}, m_s^{(p)} | \vec{S}^{(e)} \cdot \vec{S}^{(p)} | m_s^{(e)}, m_s^{(p)} \rangle = \hbar^2 m_s^{(e)} m_s^{(p)}.$$

Para escrever a matriz a ser diagonalizada é conveniente agrupar os estados com $m_s^{(e)} = m_s^{(p)}$, que não aparecem em nenhum elemento de matriz não diagonal diferente de zero. Adotando a ordenação de estados $\{(--), (++) , (-+), (+-)\}$ para as linhas e colunas da matriz, a equação secular resultante é

$$\det \begin{pmatrix} A + \frac{1}{2}B\mu_+ - \lambda & 0 & 0 & 0 \\ 0 & A - \frac{1}{2}B\mu_+ - \lambda & 0 & 0 \\ 0 & 0 & -A + \frac{1}{2}B\mu_- - \lambda & 2A \\ 0 & 0 & 2A & -A - \frac{1}{2}B\mu_- - \lambda \end{pmatrix} = 0 \quad (6.40)$$

onde foram usadas as abreviaturas

$$A \equiv \frac{\Delta E_{SS}}{4} \quad \text{e} \quad \mu_{\pm} \equiv \mu_e \pm \mu_p = -\frac{e\hbar}{2c} \left(\frac{g_e}{m_e} \mp \frac{g_p}{m_p} \right).$$

Devido ao desacoplamento dos estados $(--)$ e $(++)$ é possível identificar imediatamente os seus autovalores como sendo

$$\lambda_{--} = A + \frac{1}{2}B\mu_+ \quad \text{e} \quad \lambda_{++} = A - \frac{1}{2}B\mu_+.$$

Esses autovalores se separam linearmente com B e são degenerados para $B = 0$. Os outros dois autovalores resultam da diagonalização da sub-matriz 2×2 correspondente ao acoplamento dos estados $(-+)$ e $(+-)$ e são dados por

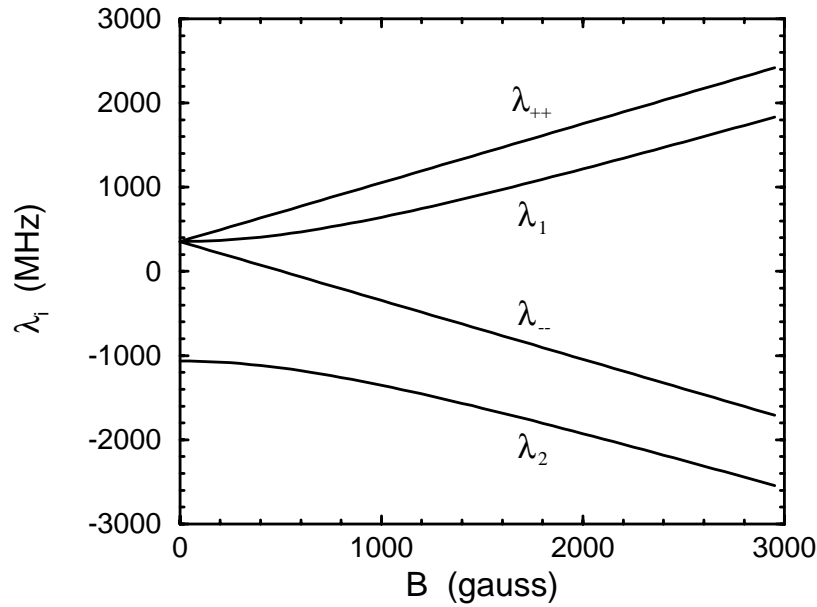


Figura 6.6: Efeito Zeeman dos estados $n = 1$, $l = 0$ do átomo de hidrogênio: autovalores λ_i (em megahertz) dos quatro estados $i = ++, 1, --, 2$ como função da intensidade do campo magnético externo B (em gauss). As regiões à extrema esquerda e à extrema direita do gráfico correspondem respectivamente aos limites de campos fracos e “Paschen-Back”.

$$\lambda_1 = -A + \sqrt{4A^2 + \frac{B^2\mu_-^2}{4}} \quad \text{e} \quad \lambda_2 = -A - \sqrt{4A^2 + \frac{B^2\mu_-^2}{4}} \quad (6.41)$$

sendo que λ_1 coincide com λ_{--} e com λ_{++} no limite de campo nulo. O exame dos autovetores associados a esses tres estados no limite de campo nulo revela que eles correspondem aos tres autovetores de S_3 com spin total $S = 1$, enquanto o autovetor associado a λ_2 corresponde, no limite de campo nulo, ao autovetor com spin total $S = 0$.

O comportamento dos autovalores está representado na figura 6.6 como função do campo magnético externo B . O limite de campo nulo reproduz o resultado da figura 6.4, e a região em que $B \lesssim 200$ gauss corresponde claramente ao limite de campos fracos, em que os efeitos do campo externo sobre os níveis hiperfinos pode ser tratada perturbativamente. O extremo oposto da figura ($B \sim 3000$ gauss) corresponde a um limite do tipo Paschen-Back, em que os efeitos do campo externo dominam os da interação spin-spin. Os *autovetores* correspondentes

aos autovalores que dependem linearmente de B , λ_{--} e λ_{++} , são autovetores tanto de S^2 e S_3 como de $S_3^{(e)}$ e $S_3^{(p)}$, e portanto não se alteram à medida que o campo é aumentado. A estrutura dos autovetores correspondentes aos autovalores λ_1 e λ_2 , por outro lado, varia continuamente a partir dos autovetores de S^2 e S_3 com autovalores $S = 1, 0$ e $m_S = 0$ no limite de campo zero, aproximando-se progressivamente dos estados $|m_s^{(e)}, m_s^{(p)} = -m_s^{(e)}\rangle$ para valores grandes de B .

É interessante analisar o papel desempenhado pelo momento magnético do próton neste contexto. Por um lado, devido à diferença de valor do magneton nuclear e do magneton de Bohr, a sua energia de acoplamento ao campo externo é muito menor que a energia de acoplamento do momento magnético do elétron, à qual aparece sempre associada seja na combinação $B\mu_+$ (no caso de λ_{--} e λ_{++}) seja na combinação $B\mu_-$ (no caso de $\lambda_{1,2}$). Desse modo, os efeitos resultantes da ocorrência *explícita* do momento magnético do próton na equação secular (6.40) através dessas quantidades representam correções relativas da ordem de 10^{-3} do que se obterá ignorando μ_p completamente. Por outro lado, o momento magnético do próton dá lugar ao acoplamento spin-spin e intervêm também no problema secular, embora de forma menos explícita, através de $A = \Delta E_{SS}/4$, que é uma energia associada à interação *mútua* dos dois momentos magnéticos. O valor de ΔE_{SS} depende, no entanto, de uma quantidade adicional, que é o valor da função de onda do estado $n = 1, l = 0$ na origem, o qual reflete, através da condição de normalização, a escala de tamanho do sistema atômico (v. eq. (6.22)). Essa escala é tal que o valor resultante para A tem a mesma ordem de grandeza da interação de um momento magnético da ordem de um *magneton de Bohr* com um campo externo da ordem de um milhar de gauss, o que define a escala de B na figura 6.6. Desse modo, as características do efeito Zeeman da estrutura hiperfina podem ser vistos como resultando essencialmente da competição entre a interação do momento magnético do elétron com o campo externo e com o campo magnético produzido pelo momento magnético do próton, enquanto este interage apenas muito mais fracamente com o campo externo.

Problemas.

6.1 - a) Se A e B são duas matrizes quaisquer de dimensão N , chamando *traço* de A a soma dos elementos diagonais

$$\text{Tr} A \equiv \sum_{n=1}^N A_{nn}$$

mostre que $\text{Tr}(AB) = \text{Tr}(BA)$.

b) Mostre que o resultado anterior se generaliza para um produto de várias matrizes A, B, C, \dots de dimensão N , no sentido de que o traço do produto das matrizes não se altera por qualquer *permutação circular* das matrizes, isto é

$$\text{Tr}(ABC \dots) = \text{Tr}(BC \dots A) = \text{Tr}(C \dots AB) = \dots$$

c) Se U é uma matriz unitária também de dimensão N , mostre que $\text{Tr}(U^\dagger AU) = \text{Tr}A$, isto é o traço é invariante por transformações unitárias da matriz.

6.2 - Usando as propriedades gerais das matrizes de Pauli $\vec{\sigma} \equiv \{\sigma_1, \sigma_2, \sigma_3\}$ mostre que vale a identidade

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) \equiv \vec{A} \cdot \vec{B} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B})$$

onde \vec{A} e \vec{B} são dois operadores vetoriais quaisquer, cujas componentes comutam com as componentes de $\vec{\sigma}$. Verifique como uma consequência dessa identidade que $(\vec{\sigma} \cdot \vec{A})^2 = A^2$ para qualquer operador vetorial \vec{A} cujas componentes comutem com as matrizes de Pauli, e que, em particular, $(\vec{\sigma} \cdot \hat{u})^2 = 1$ sendo \hat{u} um vetor unitário qualquer.

6.3 - a) Um elétron está ligado ao campo coulombiano de um núcleo de carga positiva Ze . Calcule $\sqrt{\langle v^2/c^2 \rangle}$ para o estado fundamental, sendo v a velocidade do elétron. A função de onda normalizada para o estado fundamental desse sistema é

$$\phi_{1s}(\vec{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0}$$

onde $a_0 = \hbar^2/\mu e^2$, sendo $\mu = m_e M/m_e + M$ a massa reduzida do sistema elétron-núcleo.

b) Supondo que o núcleo tenha um raio finito R e que a distribuição de carga nuclear seja uniforme, calcule o efeito do raio finito da distribuição de carga sobre a energia do estado fundamental usando teoria de perturbação de primeira ordem.

6.4 - a) Num dado sistema de coordenadas, o estado do spin de um elétron é dado por $\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] / \sqrt{2}$. Esse spin tem uma direção definida? Em caso positivo, determine a direção.

b) Construa um estado de spin ortogonal ao estado anterior (isto é, seu produto escalar com o estado anterior é nulo). Ele tem uma direção definida? Em caso positivo, determine também a direção deste estado.

6.5 - a) Calcule o elemento de matriz do momento de dipolo $d_3 = ez$ entre os estados $2s$ e $2p$ do átomo de hidrogênio. Esses elementos de matriz são relevantes para o efeito Stark.

b) Calcule o elemento de matriz do operador r^{-3} em um estado $2p$ do átomo de hidrogênio. Esta quantidade é relevante para a discussão da estrutura fina do espectro.

As funções de onda para os estados $2s$ e $2p$ do átomo de hidrogênio são

$$\phi_{2s}(\vec{r}) = \frac{1}{\sqrt{8\pi}} e^{-r/a_0} \left(1 - \frac{r}{2a_0} \right),$$

$$\phi_{2p,m}(\vec{r}) = \frac{1}{2\sqrt{6}} e^{-r/a_0} \frac{r}{a_0} Y_{1,m}(\theta, \varphi).$$

Nessas expressões a_0 é o raio de Bohr.

6.6 - Sejam $|m_s = \pm 1/2\rangle$ os autovetores da componente S_3 do spin de uma partícula de spin $1/2$. Determine explicitamente todos os autovetores do *spin total* I^2 para o spin de *duas* partículas de spin $1/2$, isto é, $\vec{I} = \vec{S}^{(1)} + \vec{S}^{(2)}$, em termos dos estados $|m_s^{(i)} = \pm 1/2\rangle$ para cada um dos dois spins $1/2$.

6.7 - Estime o valor da magnitude do campo elétrico externo E para o qual o deslocamento Stark linear dado por (6.31) tem a ordem de grandeza dos deslocamentos de nível associados à estrutura fina. Os resultados para o efeito Stark linear obtidos usando o modelo simples para o átomo de hidrogênio certamente *não* são realísticos para valores de E menores que o valor estimado dessa forma. **Sugestão:** use como ordem de grandeza dos efeitos de estrutura fina $\sim R\alpha^2/40$, cf. figura 6.2.

6.8 - Considere um sistema de dois níveis $|1\rangle$ e $|2\rangle$ cuja dinâmica é descrita pelo Hamiltoniano

$$H = H_0 + H_1$$

com

$$\begin{aligned} H_0 &= |1\rangle(\epsilon + \Delta)\langle 1| + |2\rangle(\epsilon - \Delta)\langle 2|, & \Delta > 0, & \quad \epsilon, \Delta \text{ reais,} \\ H_1 &= |1\rangle\gamma\langle 2| + |2\rangle\gamma^*\langle 1|. \end{aligned}$$

a) Obtenha os autovalores de H em termos de ϵ , Δ e γ .

b) Mostre que no caso em que $\Delta = 0$ os autovalores dependem *linearmente* de $|\gamma|$.

c) Mostre que no caso em que $\Delta \neq 0$ e $|\gamma|/\Delta \ll 1$ os autovalores dependem de $|\gamma|$ de uma forma que é aproximadamente quadrática.

d) Ainda com $\Delta \neq 0$, mas supondo $|\gamma|/\Delta \gg 1$, mostre que os autovalores dependem de $|\gamma|$ de uma forma que é aproximadamente linear.

e) Tomando $2\Delta \simeq h\nu$, usando para ν o deslocamento de Lamb ($\sim 10^9 \text{ s}^{-1}$), e tomando ainda

$$\gamma = -\frac{eE}{3} \int_0^\infty dr u_{2s\frac{1}{2}}(r) r u_{2p\frac{1}{2}}(r)$$

onde e é a carga do elétron, E é a intensidade do campo elétrico externo e u_{nlj} são orbitais do átomo de hidrogênio, *estime* o valor de E a partir do qual a dependência do deslocamento Stark com E aparece como sendo aproximadamente linear.

6.9 - Estude o efeito de um campo magnético externo \vec{B} constante e uniforme, orientado na direção $\hat{3}$, sobre os estados $n = 1$, $l = 1$, $-1 \leq m_l \leq 1$, $-1/2 \leq m_s^{(e)} \leq 1/2$ do átomo de hidrogênio seguindo um procedimento análogo ao usado no item **3** acima, para poder cobrir também a região intermediária entre a região de campos fracos e a do efeito Paschen-Back. Como naquele caso, a base conveniente é aquela em que L_3 e $S_3^{(e)}$ são diagonais. Os seis estados $\{|m_l, m_s^{(e)}\rangle\}$ envolvidos levam a uma matriz secular 6×6 cuja estrutura é suficientemente simples para que possa ser diagonalizada analiticamente sem dificuldade.

a) Escreva o hamiltoniano de interação sob a forma

$$H' = \frac{\mu_B B}{\hbar} (L_s + g_e S_3^{(e)}) + \frac{\gamma_{LS}}{\hbar^2} \left(\frac{L_+ S_-^{(e)} + L_- S_+^{(e)}}{2} + L_3 S_3^{(e)} \right)$$

obtendo uma expressão para a constante γ_{LS} a ser usada nesse caso, e relacione seu valor com a separação observada dos níveis da estrutura fina correspondentes a $n = 1$, $l = 1$ no átomo de hidrogênio. A constante μ_B que aparece no primeiro termo é o magneton de Bohr.

b) Verifique que os únicos elementos de matriz não diagonais diferentes de zero de H' são aqueles que têm o mesmo valor de $m_l + m_s^{(e)}$, embora valores de m_l e de $m_s^{(e)}$ diferentes, e que isso permite ordenar as linhas e colunas da matriz secular de modo que ela tenha uma estrutura semelhante à da matriz (6.40). **Sugestão:** Um dos ordenamentos possíveis dos estados $(m_l, m_s^{(e)})$ é $(1, +)$, $(-1, -)$, $(1, -)$, $(0, +)$, $(0, -)$, $(-1, +)$.

c) Obtenha os autovalores da matriz secular e os respectivos autovetores. Mostre que nos limites de campos fracos e de campos fortes são recuperados os resultados dos itens **1** e **2** respectivamente.

d) Mostre o comportamento do efeito Zeeman na região intermediária através de um gráfico semelhante ao da figura 6.6.

e) Inclua no gráfico os níveis $n = 1$, $l = m_l = 0$, $m_s^{(e)} = \pm 1/2$, supondo que no caso em que o campo magnético externo é zero o esquema de níveis é o indicado à direita da figura

6.3. Porque o efeito Zeeman dos estados $n = 1, l = 1$ e $n = 1, l = 0$ podem ser tratados independentemente, apesar da degenerescência que ocorre para $B = 0$?

Referências Gerais.

O trabalho no qual foi introduzido o spin para dar conta das propriedades observadas dos espectros atômicos apareceu em *Nature*, em 1926, pouco antes do trabalho de Thomas (op. cit.):

★ **G. E. Uhlenbeck e S Goudsmith**, *Spinning Electrons and the Structure os Spectra*, *Nature* **117**, 264 (1926);

Discussões das propriedades dos estados de spin $1/2$ bem como da álgebra de momentos angulares podem ser encontradas em

★ **A. Messiah**, op. cit., Vol II, Capítulo XIII

bem como em livros especialmente dedicados ao tratamento quântico do momento angular:

★ **M. E. Rose**, *Elementary Theory of Angular Momentum*, John Wyley, 1957;

★ **A. R. Edmonds**, *Angular Momentum in Quantum Mechanics*, Princeton U. Press, 1957;

★ **U. Fano e G. Racah**, *Irreducible Tensorial Sets*, Academic Press, 1959;

★ **D. Brink e G. Satchler**, *Angular Momentum*, Oxford University Press, 19XX.

Um tratamento informal e simples das correções relativísticas à estrutura fina dos esperctros, embora incompleto no que se refere ao termo de Darwin, pode ser encontrado nas notas manuscritas de Fermi

★ **E. Fermi**, *Notes on Quantum Mechanics*, The University of Chicago Press, 1961.

Um tratamento mais completo, feito no entanto a partir da equação de Dirac, pode ser encontrado em

★ **J. J. Sakurai**, *Advanced Quantum Mechanics*, Addison-Wesley Publishing Company, 1967, p. 88.

Tratamentos dos efeitos Stark e Zeeman podem ser encontrados na maioria dos textos, e em particular em

★ **L. D. Landau e E. M. Lifshitz**, Op. cit.

★ **C. Cohen-Tanoudji, B. Diu e F. Laloë**, Op. cit.

Parte II

Capítulo 7

Sistemas de partículas idênticas.

A indistinguibilidade das partículas constituintes de um sistema físico, isto é, a inexistência de propriedades *intrínsecas* que permitam identifica-las individualmente, é um fato que tem consequências muito mais radicais no contexto da mecânica quântica que no da mecânica clássica, e as razões para isso se ligam mais uma vez à diferença existente entre os substratos cinemáticos dessas duas teorias. De fato, a cinemática clássica permite, em princípio, estabelecer sempre uma “distinguibilidade relativa” de partículas idênticas, através do fato de que, embora as partículas não possam ser distinguidas individualmente com base em propriedades intrínsecas, é sempre possível acompanhá-las individualmente através da continuidade de suas trajetórias no espaço de fases clássico, o que permite estender a qualquer tempo posterior alguma rotulação convencional que tenha sido estabelecida inicialmente. Isso não é em geral possível, porém, no contexto da mecânica quântica, onde o objeto cinemático básico é um vetor de estado interpretado como uma amplitude de probabilidade, pois neste caso há limitações para a localizabilidade no espaço de fases clássico expressas por relações de incerteza que inviabilizam a idéia de trajetórias univocamente associadas às partículas.

Em termos mais formais, considerando inicialmente, por simplicidade, apenas um par de partículas idênticas, um estado clássico no tempo t_0 , $a(t_0)$, é dado em termos das posições e dos momentos das partículas nesse instante, o que pode ser representado como

$$a(t_0) = \begin{cases} q_1, p_1 & \rightarrow \text{part. 1} \\ q_2, p_2 & \rightarrow \text{part. 2} \end{cases}$$

onde os q_i, p_i , $i = 1, 2$ são coordenadas do espaço de fases clássico. A indistinguibilidade das partículas corresponde ao fato de que o estado obtido de $a(t)$ simplesmente pela troca das duas partículas, isto é

$$a'(t_0) = \begin{cases} q_1, p_1 & \rightarrow \text{part. 2} \\ q_2, p_2 & \rightarrow \text{part. 1} \end{cases}$$

é na realidade indistinguível do primeiro. Os rótulos 1 e 2 atribuídos às partículas têm portanto um valor puramente convencional. As equações de movimento clássicas dão, para condições iniciais dadas por $a(t_0)$, uma solução única

$$a(t) = \begin{cases} q_1(t), p_1(t) & \rightarrow \text{part. 1} \\ q_2(t), p_2(t) & \rightarrow \text{part. 2} \end{cases}$$

enquanto para condições iniciais $a'(t_0)$ a solução, também única, será

$$a'(t) = \begin{cases} q_1(t), p_1(t) & \rightarrow \text{part. 2} \\ q_2(t), p_2(t) & \rightarrow \text{part. 1} \end{cases}$$

que difere da anterior apenas pela re-rotulação das partículas. Isso se deve a que a Hamiltoniana que caracteriza a dinâmica do par de partículas idênticas deve ser necessariamente *simétrica* nas variáveis canônicas de cada uma delas. Este requisito se deve simplesmente ao fato de que a não simetria da Hamiltoniana, nesse sentido, forneceria um critério dinâmico para distinguir as duas partículas, que não poderiam mais, por isso, ser consideradas idênticas. Qualquer que tenha sido a rotulação convencional adotada em t_0 , no entanto, ela pode ser propagada para outros valores de t através da solução das equações de movimento. Assim, no caso $a'(t)$, a partícula “1” será, em t a que se encontra na posição $q_2(t)$ com momento $p_2(t)$.

Na mecânica quântica, por outro lado, os estados de um sistema de duas partículas são representados por vetores de estado no espaço produto dos espaços lineares que funcionam como espaços de fase para cada uma delas. Designando respectivamente por $\{|n^{(1)}\rangle\}$, $n^{(1)} = 1, 2, \dots$ e $\{|n^{(2)}\rangle\}$, $n^{(2)} = 1, 2, \dots$ bases ortonormais (discretas) em cada um desses espaços, um estado genérico (que pode em particular ser entendido como descrevendo o sistema num dado instante t_0) é um vetor da forma

$$|a\rangle = \sum_{n_1, n_2} a_{n_1 n_2} |n_1^{(1)}\rangle \otimes |n_2^{(2)}\rangle \equiv \sum_{n_1, n_2} a_{n_1 n_2} |n_1^{(1)} n_2^{(2)}\rangle. \quad (7.1)$$

A operação de troca das duas partículas aplicada a esse estado leva a um vetor que é em geral diferente de $|a\rangle$ e dado por

$$|a'\rangle = \sum_{n_1, n_2} a_{n_1 n_2} |n_1^{(2)} n_2^{(1)}\rangle.$$

Essa operação de permutação do papel desempenhado por cada uma das partículas pode ser descrita introduzindo um operador P_{12} definido em termos da base produto como

$$P_{12} |n_1^{(1)} n_2^{(2)}\rangle = |n_1^{(2)} n_2^{(1)}\rangle.$$

A relação entre $|a\rangle$ e $|a'\rangle$ fica, em termos desse operador,

$$|a'\rangle = P_{12}|a\rangle.$$

Ele tem ainda as propriedades facilmente verificáveis

$$P_{12}^2 = \hat{1} \quad \text{e} \quad P_{12}^\dagger = P_{12}$$

que o caracterizam como um operador hermiteano com autovalores ± 1 , e portanto também unitário, pois $P_{12}^\dagger = P_{12}^{-1}$.

A indistinguibilidade das duas partículas quânticas não foi utilizada até aqui. Ela pode ser introduzida na discussão através da exigência de que pares de estados diferindo apenas pela troca das duas partículas sejam indistinguíveis, no sentido que, qualquer que seja o estado $|a\rangle$, se tenha, para todo observável g ,

$$\langle a|g|a\rangle = \langle a'|g|a'\rangle = \langle a|P_{12}^\dagger g P_{12}|a\rangle. \quad (7.2)$$

Uma condição *necessária e suficiente* para que essa relação seja válida para um observável g dado é a de que ele seja *simétrico* com relação à troca das duas partículas, no sentido que

$$g = P_{12}^\dagger g P_{12} = P_{12}^{-1} g P_{12} \quad \text{ou seja} \quad [g, P_{12}] = 0. \quad (7.3)$$

De fato, é imediato que essa condição é suficiente para a validade da relação (7.2) para g . Para ver que ela é também necessária (isto é, que a validade de (7.2) para qualquer $|a\rangle$ implica em (7.3)) basta considerar vetores de estado da forma $N(|a\rangle + \alpha|b\rangle)$, onde N é um fator de normalização, em (7.2), que dá então

$$\langle a|g|a\rangle + |\alpha|^2 \langle b|g|b\rangle + \alpha \langle a|g|b\rangle + \alpha^* \langle b|g|a\rangle = \langle a'|g|a'\rangle + |\alpha|^2 \langle b'|g|b'\rangle + \alpha \langle a'|g|b'\rangle + \alpha^* \langle b'|g|a'\rangle.$$

Usando o fato de que os dois primeiros termos de cada um dos lados dessa relação são também iguais, segue que essa condição equivale a

$$\alpha (\langle a|g|b\rangle - \langle a'|g|b'\rangle) - \alpha^* (\langle b|g|a\rangle - \langle b'|g|a'\rangle) = 0$$

donde segue que, fazendo sucessivamente as escolhas $\alpha = 1$ e $\alpha = i$,

$$\langle a|g|b\rangle = \langle a'|g|b'\rangle = \langle a|P_{12}^\dagger g P_{12}|b\rangle$$

para estados $|a\rangle$ e $|b\rangle$ quaisquer, o que equivale a (7.3). A indistinguibilidade das duas partículas no sentido proposto acima exige, portanto, que *todos* os observáveis sejam simétricos no sentido

da relação (7.3), isto é, o operador P_{12} deve ser compatível com *todos* os observáveis do sistema de duas partículas idênticas.

Esse fato tem consequências importantes para a dinâmica desse sistema, que podem ser explicitadas notando que o operador identidade do espaço de estados pode ser decomposto, usando o operador P_{12} , como

$$\hat{1} = \frac{\hat{1} - P_{12}}{2} + \frac{\hat{1} + P_{12}}{2} \equiv \mathcal{A} + \mathcal{S},$$

sendo fácil verificar que \mathcal{A} e \mathcal{S} são operadores de projeção complementares, isto é

$$\mathcal{A} = \mathcal{A}^\dagger = \mathcal{A}^2, \quad \mathcal{S} = \mathcal{S}^\dagger = \mathcal{S}^2, \quad \text{e} \quad \mathcal{AS} = \mathcal{SA} = 0.$$

Para entender o significado desses projetores basta notar que um estado qualquer $|a\rangle$ admite a decomposição

$$|a\rangle = \mathcal{A}|a\rangle + \mathcal{S}|a\rangle$$

e que as componentes satisfazem

$$P_{12}\mathcal{A}|a\rangle = -\mathcal{A}|a\rangle \quad \text{e} \quad P_{12}\mathcal{S}|a\rangle = \mathcal{S}|a\rangle$$

isto é, são autovetores de P_{12} com autovalores -1 e $+1$, isto é, são estados respectivamente *anti-simétricos* e *simétricos* com relação à troca das duas partículas idênticas. Um estado qualquer de duas partículas pode portanto ser sempre decomposto em uma componente anti-simétrica e uma componente simétrica com relação à troca das partículas, e \mathcal{A} e \mathcal{S} projetam respectivamente sobre a componente antissimétrica e sobre a componente simétrica do estado. No caso de partículas idênticas, em vista da compatibilidade de P_{12} com *todos* os observáveis, segue que

$$\mathcal{A}g\mathcal{S} = \mathcal{S}g\mathcal{A} = 0$$

para qualquer g , isto é, *nenhum* observável conecta vetores de simetrias diferentes. Essa decomposição do espaço de vetores de estado em dois setores ortogonais que não são conectados por nenhum observável é um exemplo do que se chama uma *regra de super-seleção*. No caso, a regra de super-seleção permite considerar separadamente estados simétricos e anti-simétricos, já que nenhum efeito observável provém da superposição coerente desses dois tipos de estado. Como o Hamiltoniano, responsável pela dinâmica do sistema, deve ser também simétrico (sendo um observável), a evolução temporal das componentes simétrica e anti-simétrica é independente, e a simetria ou anti-simetria é conservada ao longo da evolução temporal. Isso significa que

existem na realidade duas dinâmicas independentes para sistemas de duas partículas idênticas, associadas respectivamente aos sub-espços correspondentes aos projetores \mathcal{A} e \mathcal{S} .

As propriedades físicas de estados simétricos e anti-simétricos são porém muito diferentes. Tomando, como exemplo, as projeções do estado geral (7.1)

$$|a_{\pm}\rangle \equiv \frac{\mathcal{S}}{\mathcal{A}} |a\rangle = \frac{1}{2} \sum_{n_1, n_2} a_{n_1 n_2} \left(|n_1^{(1)} n_2^{(2)}\rangle \pm |n_1^{(2)} n_2^{(1)}\rangle \right)$$

é fácil ver que amplitudes do tipo $\langle n^{(1)} n^{(2)} | a_{\pm} \rangle$ (cujo módulo quadrado se relaciona com a probabilidade de encontrar as duas partículas no mesmo estado $|n\rangle$ dos respectivos espaços fatores) são necessariamente *nulas* no caso da componente anti-simétrica $|a_{-}\rangle$ mas não no caso da componente simétrica $|a_{+}\rangle$.

7.1 Férmions e bosons.

A discussão precedente mostra que existem essencialmente duas situações independentes a considerar com relação à questão do tratamento quântico de um sistema de duas partículas idênticas, as quais envolvem respectivamente o setor anti-simétrico e o setor simétrico do espaço produto utilizado na expressão (7.1). A questão da realização efetiva de cada uma dessas duas possibilidades em casos específicos não pode ser respondida dentro dos limites da teoria quântica não relativística. Existe, no entanto, uma correlação *experimental* completa entre a simetria dos estados quânticos e o spin das partículas: partículas de spin semi-inteiro estão *sempre* associadas a estados anti-simétricos, enquanto partículas de spin inteiro estão sempre associadas a estados simétricos. Em outras palavras, o vetor de estado que descreve um sistema de duas partículas idênticas de spin semi-inteiro pertence sempre ao setor anti-simétrico do espaço produto, isto é, chamando $|f\rangle$ esse vetor de estado

$$\mathcal{A}|f\rangle = |f\rangle; \quad \mathcal{S}|f\rangle = 0 \quad \text{e} \quad P_{12}|f\rangle = -|f\rangle.$$

Analogamente, chamando $|b\rangle$ um vetor de estado qualquer que descreve um sistema de duas partículas idênticas de spin inteiro, o que se tem é

$$\mathcal{S}|b\rangle = |b\rangle; \quad \mathcal{A}|b\rangle = 0 \quad \text{e} \quad P_{12}|b\rangle = |b\rangle.$$

No primeiro caso, diz-se que as partículas obedecem à estatística de Fermi-Dirac, enquanto no segundo caso obedecem à estatística de Bose-Einstein. Partículas de spin semi-inteiro são chamadas *férmions*, enquanto partículas de spin inteiro são chamadas *bosons*.

Utilizando os operadores de projeção \mathcal{A} e \mathcal{S} é fácil construir explicitamente bases ortonormais nos sub-espços correspondentes a partir da base produto utilizada na expansão (7.1). Os

vetores dessas bases são, de fato, as projeções normalizadas dos vetores da base produto geral, isto é

$$|n_1 n_2\rangle_- \equiv \sqrt{2}\mathcal{A}|n_1^{(1)} n_2^{(2)}\rangle = \frac{1}{\sqrt{2}} \left(|n_1^{(1)} n_2^{(2)}\rangle - |n_1^{(2)} n_2^{(1)}\rangle \right)$$

para o caso de férmions e

$$|n_1 n_2\rangle_+ \equiv \begin{cases} \sqrt{2}\mathcal{S}|n_1^{(1)} n_2^{(2)}\rangle & \text{se } n_1 \neq n_2 \\ \mathcal{S}|n_1^{(1)} n_1^{(2)}\rangle = |n_1^{(1)} n_1^{(2)}\rangle & \text{se } n_1 = n_2. \end{cases}$$

para o caso de bosons. Deve ser notada a necessidade de introduzir um fator adicional $1/\sqrt{2}$ para a normalização dos estados do setor simétrico quando $n_1 = n_2$. Esse fator resulta do fato de que

$$P_{12}|n_1^{(1)} n_1^{(2)}\rangle = |n_1^{(1)} n_1^{(2)}\rangle$$

enquanto, se $n_1 \neq n_2$, o estado $P_{12}|n_1^{(1)} n_2^{(2)}\rangle = |n_1^{(2)} n_2^{(1)}\rangle$ é ortogonal a $|n_1^{(1)} n_2^{(2)}\rangle$. De certa forma como uma contrapartida desse peso diferenciado, os estados “de ocupação dupla” $|n_1^{(1)} n_1^{(2)}\rangle$ estão completamente ausentes do setor anti-simétrico. Uma forma, embora um tanto peculiar, de escrever o estado simétrico $|n_1 n_2\rangle_+$ que é válida tanto para $n_1 = n_2$ como para $n_1 \neq n_2$ é

$$|n_1 n_2\rangle_+ \equiv \frac{\sqrt{2}}{1 + (\sqrt{2} - 1)\delta_{n_1 n_2}} \mathcal{S}|n_1^{(1)} n_2^{(2)}\rangle = \frac{1}{1 + (\sqrt{2} - 1)\delta_{n_1 n_2}} \frac{1}{\sqrt{2}} \left(|n_1^{(1)} n_2^{(2)}\rangle + |n_1^{(2)} n_2^{(1)}\rangle \right),$$

onde o pré-fator envolvendo $\delta_{n_1 n_2}$ dá conta das normalizações apropriadas a cada um dos dois casos.

Os vetores $|n_1 n_2\rangle_{\pm}$ assim definidos são normalizados e mutuamente ortogonais. Isso decorre imediatamente da ortonormalidade das bases $\{|n^{(i)}\rangle\}$ tomadas nos espaços fatores. As duas famílias de vetores $\{|n_1 n_2\rangle_{\pm}\}$ são também completas nos sub-espacos associados aos projetores \mathcal{A} e \mathcal{S} respectivamente. De fato, a projeção $\mathcal{A}|a\rangle$ de um vetor geral do espaço produto é um vetor geral do setor anti-simétrico e, utilizando a expansão (7.1) para $|a\rangle$, essa projeção pode ser escrita como

$$\mathcal{A}|a\rangle = \sum_{n_1 n_2} a_{n_1 n_2} \mathcal{A}|n_1^{(1)} n_2^{(2)}\rangle = \sum_{n_1 n_2} a_{n_1 n_2} \frac{1}{\sqrt{2}} |n_1 n_2\rangle_-$$

e pode portanto ser sempre expandida em termos dos $\{|n_1 n_2\rangle_-$. A última expressão pode ser escrita ainda da forma alternativa

$$\mathcal{A}|a\rangle = \sum_{n_1 < n_2} \frac{1}{\sqrt{2}} (a_{n_1 n_2} - a_{n_2 n_1}) |n_1 n_2\rangle_- \quad (7.4)$$

que faz uso da antissimetria dos vetores de base, $|n_1 n_2\rangle_- = -|n_2 n_1\rangle_-$. Como a anti-simetria exclui os vetores de ocupação dupla ($n_1 = n_2$) da expansão, é claro que para o vetor de estado geral de dois férmions idênticos as amplitudes $\langle nn|\mathcal{A}|a\rangle$ são sempre nulas, o que significa que dois férmions idênticos não podem ocupar simultaneamente o mesmo estado de uma partícula, um resultado conhecido como “princípio de exclusão de Pauli”.

De forma completamente análoga se pode mostrar que a componente simétrica de um vetor geral do espaço produto escrito sob a forma (7.1) admite a expansão

$$\begin{aligned} \mathcal{S}|a\rangle &= \sum_{n_1 n_2} a_{n_1 n_2} \mathcal{S}|n_1^{(1)} n_2^{(2)}\rangle = \sum_{n_1 n_2} a_{n_1 n_2} \frac{1 + (\sqrt{2} - 1)\delta_{n_1 n_2}}{\sqrt{2}} |n_1 n_2\rangle_+ \\ &= \sum_{n_1 \leq n_2} \frac{1 + (\sqrt{2} - 1)\delta_{n_1 n_2}}{\sqrt{2}} (a_{n_1 n_2} + a_{n_2 n_1}) |n_1 n_2\rangle_+ \end{aligned}$$

na qual, ao contrário do que acontece no caso do setor anti-simétrico, existem em geral termos com $n_1 = n_2$.

Existe uma forma simples e muito usada de escrever os vetores das bases anti-simétrica e simétrica, $|n_1 n_2\rangle_{\pm}$, em termos dos vetores $|n^{(i)}\rangle$, $i = 1, 2$ das bases de uma partícula. Os vetores da base anti-simétrica podem ser escritos sob a forma de um determinante

$$|n_1 n_2\rangle_- = \frac{1}{\sqrt{2}} (|n_1^{(1)} n_2^{(2)}\rangle - |n_1^{(2)} n_2^{(1)}\rangle) = \frac{1}{\sqrt{2}} \det \begin{pmatrix} |n_1^{(1)}\rangle & |n_1^{(2)}\rangle \\ |n_2^{(1)}\rangle & |n_2^{(2)}\rangle \end{pmatrix}$$

em que índices de linhas e colunas se referem a estados e partículas respectivamente. A anti-simetria é realizada através da propriedade de determinantes de mudança de sinal por uma troca de duas linhas ou de duas colunas. Analogamente os vetores da base simétrica podem ser escritos sob a forma de um “permanente”

$$|n_1 n_2\rangle_+ = \frac{1}{\sqrt{2}[1 + (\sqrt{2} - 1)\delta_{n_1 n_2}]} \text{perm} \begin{pmatrix} |n_1^{(1)}\rangle & |n_1^{(2)}\rangle \\ |n_2^{(1)}\rangle & |n_2^{(2)}\rangle \end{pmatrix}$$

que é definido de forma semelhante a um determinante, mas com todos os termos positivos.

7.1.1 Sistemas de muitas partículas idênticas.

A extensão do tratamento anterior para sistemas de três ou mais partículas idênticas não envolve maiores dificuldades, a menos da complicação técnica da existência de outros tipos de simetria além dos dois tipos que ocorrem no caso de apenas duas partículas. A simplicidade do caso de apenas duas partículas se deve basicamente ao fato de que o operador identidade $\hat{1}$ e o operador de troca P_{12} esgotam as possibilidades de permutação das partículas, o que não ocorre no caso de três ou mais partículas. Em geral existem $N!$ permutações distintas de N partículas, e cada uma dessas permutações pode ser associada a um operador P_i , $i = 1, \dots, N!$ (incluindo o operador identidade $\hat{1}$ para a “permutação idêntica”). Esse conjunto de $N!$ operadores forma um grupo chamado *grupo simétrico* das permutações das N partículas. No caso $N = 2$ esse grupo é constituído apenas dos dois elementos $P_1 \equiv \hat{1}$ e $P_2 \equiv P_{12} = P_{12}^{-1} = P_{12}^\dagger$. No caso $N \geq 3$ os $N!$ operadores P_i são também operadores unitários (embora em geral não hermiteanos), pois a ação de cada um deles sobre os vetores da base produto equivale a uma re-rotulação dos vetores da base. O critério de indistinguibilidade (7.2) deve agora ser estendido exigindo que

$$\langle a|g|a\rangle = \langle a|P_i^\dagger g P_i|a\rangle = \langle a|P_i^{-1} g P_i|a\rangle$$

para qualquer permutação P_i , o que leva à condição de simetria

$$[g, P_i] = 0$$

para os observáveis. Utilizando técnicas de teoria de grupos é possível identificar, a partir dessa condição, sub-espacos cujas propriedades de simetria são invariantes sob qualquer permutação P_i , e que não são conectados por *nenhum* observável. Além de estados completamente anti-simétricos e completamente simétricos, isto é, tais que a troca de duas partículas *qualquer* equivale à multiplicação do estado por -1 ou por $+1$, respectivamente, existem setores de “simetria intermediária”, isto é, que são simétricos com relação a algumas permutações, anti-simétricos com relação a outras e que não têm simetria definida com relação a ainda outras permutações. No entanto, a correlação entre spin e estatística mostra que estados de muitos férmions (bosons) idênticos devem ser sempre *completamente anti-simétricos (simétricos)*. No caso de duas partículas idênticas esses dois tipos de simetria correspondem aos projetores \mathcal{A} e \mathcal{S} , que satisfazem a relação de complementaridade com relação ao espaço produto completo

$$\mathcal{A} + \mathcal{S} = \hat{1}.$$

No caso $N \geq 3$ é fácil construir os operadores correspondentes usando os $N!$ elementos do grupo simétrico. Isso se faz pondo

$$\mathcal{A} = \frac{1}{N!} \sum_{i=1}^{N!} (-1)^{P_i} P_i \quad \text{e} \quad \mathcal{S} = \frac{1}{N!} \sum_{i=1}^{N!} P_i$$

onde os P_i são as diferentes permutações das N partículas e $(-1)^{P_i}$ é uma fase igual a $+1$ ou -1 para permutações pares e ímpares respectivamente. Além de se reduzir, no caso $N = 2$, à definição dada anteriormente, é possível mostrar diretamente (embora isso seja algo trabalhoso) que esses operadores são também operadores de projeção ortogonais no espaço produto de N partículas, isto é

$$\mathcal{A} = \mathcal{A}^\dagger = \mathcal{A}^2, \quad \mathcal{S} = \mathcal{S}^\dagger = \mathcal{S}^2 \quad \text{e} \quad \mathcal{AS} = \mathcal{SA} = 0.$$

Quando $N \geq 3$, no entanto, a existência de estados de simetria intermediária implica que

$$\mathcal{A} + \mathcal{S} \neq \hat{1}$$

devido à existência de vetores no espaço produto geral que não podem ser decompostos em componentes completamente simétricas e completamente anti-simétricas apenas. Devido à correlação observada entre spin e estatística, no entanto, tais estados de simetria intermediária não são realizados fisicamente, o que dispensa sua explicitação no presente contexto.

De uma forma inteiramente análoga à empregada no caso $N = 2$, é possível usar os operadores gerais \mathcal{A} e \mathcal{S} para construir bases nos setores anti-simétrico e simétrico do espaço produto a partir de uma base de vetores do tipo

$$|n_1^{(1)} \dots n_N^{(N)}\rangle \equiv |n_1^{(1)}\rangle \otimes \dots \otimes |n_N^{(N)}\rangle$$

onde os vetores $|n^{(j)}\rangle$ são elementos de uma base ortonormal no espaço de vetores de estado para a j -ésima partícula. De fato, escrevendo um vetor geral do sub-espaço completamente anti-simétrico como projeção de um vetor geral do espaço produto

$$\mathcal{A}|a\rangle = \sum_{n_1 \dots n_N} a_{n_1 \dots n_N} \mathcal{A}|n_1^{(1)} \dots n_N^{(N)}\rangle.$$

Os vetores anti-simétricos $\mathcal{A}|n_1^{(1)} \dots n_N^{(N)}\rangle$ podem ser escritos como determinantes de dimensão N

$$\mathcal{A}|n_1^{(1)} \dots n_N^{(N)}\rangle = \frac{1}{N!} \det \begin{pmatrix} |n_1^{(1)}\rangle & \dots & |n_1^{(N)}\rangle \\ \dots & \dots & \dots \\ |n_N^{(1)}\rangle & \dots & |n_N^{(N)}\rangle \end{pmatrix}.$$

Dois determinantes desse tipo (que são freqüentemente chamados *determinantes de Slater*) serão ortogonais quando eles diferirem pelo menos por um dos estados de uma partícula envolvidos na sua construção, e da ortonormalidade das bases de uma partícula $\{|n^{(j)}\rangle\}$ resulta que o quadrado se sua norma é

$$\langle n_1^{(1)} \dots n_N^{(N)} | \mathcal{A} | n_1^{(1)} \dots n_N^{(N)} \rangle = \frac{1}{N!}$$

sempre que os n_i forem todos diferentes entre si. Se dois ou mais estados de uma partícula forem iguais a norma é zero, o que corresponde ao princípio de exclusão de Pauli para o estado de N férmions. Portanto um vetor geral do sub-espço completamente anti-simétrico pode ser expandido em termos dos vetores de base

$$|n_1 \dots n_N\rangle_- = \sqrt{N!} \mathcal{A} |n_1^{(1)} \dots n_N^{(N)}\rangle$$

nos quais todos os n_i são diferentes como

$$\mathcal{A}|a\rangle = \sum_{n_1 \dots n_N} a_{n_1 \dots n_N} \frac{1}{\sqrt{N!}} |n_1 \dots n_N\rangle_- = \sum_{n_1 < n_2 < \dots < n_N} \left(\frac{1}{\sqrt{N!}} \sum_{i=1}^{N!} (-1)^{P_i} a_{n_{p_1} \dots n_{p_N}} \right) |n_1 \dots n_N\rangle_-.$$

Esta última expressão, em particular, generaliza a relação (7.4), e mostra que estados de N férmions idênticos podem ser expandidos em termos de uma base de determinantes formados a partir de N estados de uma partícula *diferentes*, identificados pelos índices n_i e com $n_1 < n_2 < \dots < n_N$.

Um procedimento completamente análogo se aplica ao caso de N bosons idênticos. A expansão de um vetor geral do sub-espço completamente simétrico, escrito como a projeção apropriada de um vetor geral do espaço produto, é

$$\mathcal{S}|a\rangle = \sum_{n_1 \dots n_N} a_{n_1 \dots n_N} \mathcal{S} |n_1^{(1)} \dots n_N^{(N)}\rangle$$

e os vetores completamente simétricos $\mathcal{S} |n_1^{(1)} \dots n_N^{(N)}\rangle$ podem ser escritos como permanentes de ordem N

$$\mathcal{S} |n_1^{(1)} \dots n_N^{(N)}\rangle = \frac{1}{N!} \text{perm} \begin{pmatrix} |n_1^{(1)}\rangle & \dots & |n_1^{(N)}\rangle \\ \dots & \dots & \dots \\ |n_N^{(1)}\rangle & \dots & |n_N^{(N)}\rangle \end{pmatrix}.$$

Para que dois vetores desse tipo sejam ortogonais é necessário e suficiente que eles difiram pelo menos por um dos estados de uma partícula envolvidos na construção do permanente, e as

projeções completamente simétricas de vetores da base produto que diferem apenas por uma permutação dos estados de uma partícula são iguais. A norma das projeções completamente simétricas dos vetores da base produto depende do número de repetições existentes entre os n_i , generalizando o que foi encontrado no caso $N = 2$, e os vetores normalizados da base completamente simétrica são

$$|n_1 \dots n_N\rangle_+ = \sqrt{\frac{N!}{\nu_1! \nu_2! \dots \nu_q!}} \mathcal{S} |n_1^{(1)} \dots n_N^{(N)}\rangle$$

onde foi suposto que os N vetores de uma partícula $|n_i\rangle$, $i = 1, \dots, N$ são de q tipos distintos ($q \leq N$), com multiplicidades ν_1, \dots, ν_q , de modo que $\nu_1 + \dots + \nu_q = N$. Isso está ligado ao fato de que no caso do setor completamente simétrico a repetição de índices *não* é excluída. Finalmente, utilizando a simetria destes últimos vetores de base o estado geral completamente simétrico admite uma expansão do tipo

$$\mathcal{S}|a\rangle = \sum_{n_1 \leq n_2 \leq \dots \leq n_N} \left(\frac{1}{\sqrt{N! \nu_1! \dots \nu_q!}} \sum_{i=1}^{N!} a_{n_{p_1} \dots n_{p_N}} \right) |n_1 \dots n_N\rangle_+$$

de modo que o conjunto dos permanentes normalizados, formados a partir de N estados de uma partícula identificados pelos índices $n_1 \leq n_2 \leq \dots \leq n_N$.

Um ponto que exige atenção na construção da base completamente *anti-simétrica* de determinantes $|n_1 \dots n_N\rangle_-$, $n_1 < n_2 < \dots < n_N$, é a particular ordem em que os vetores da base de uma partícula são rotulados. Uma mudança de convenção de rotulação pode levar a uma mudança no ordenamento das linhas do determinante que implica numa mudança de sinal de vetores da base anti-simétrica. A rotulação dos vetores da base de uma partícula é arbitrária, mas deve ser cuidadosamente observada, uma vez definida, bem como a convenção de ordenação adotada para a colocação de linhas e colunas na construção dos determinantes, sob pena de erros de sinal.

No caso dos vetores da base completamente simétrica os problemas de ordenação não existem. Um vetor dessa base é completamente caracterizado pelos estados da base de uma partícula utilizados na sua construção apenas, independentemente de critérios de rotulação.

7.1.2 Exemplos simples. Indistinguibilidade quântica e clássica.

1. As propriedades de simetria dos estados de sistemas quânticos de muitas partículas idênticas dão origem a alguns efeitos muito importantes, conhecidos genericamente como “efeitos de

troca". Um exemplo simples desse tipo de efeito aparece quando se considera os elementos de matriz do potencial coulombiano

$$V_c = \frac{e^2}{|\vec{r}_1 - \vec{r}_2|}$$

numa base de determinantes de Slater para o setor anti-simétrico do espaço de estados para um sistema de dois electrons, que são férmions de spin $1/2$. Em primeiro lugar, cabe observar que uma forma mais completa de caracterizar a ação desse operador é escreve-lo como $V_c \otimes \hat{1}_S$, onde o fator $\hat{1}_S$ enfatiza o fato de que ele age como o operador identidade sobre os spins dos electrons. Em seguida, deve ser observado que $[P_{12}, V_c \otimes \hat{1}_S] = 0$, isto é, $V_c \otimes \hat{1}_S$ é devidamente simétrico sob a operação de troca dos dois electrons.

Os vetores ortonormais da base anti-simétrica podem ser representados por funções de onda do tipo

$$\langle \vec{r}_1 s_1, \vec{r}_2 s_2 | n_1 n_2 \rangle_- = \frac{1}{\sqrt{2}} \det \begin{pmatrix} \langle \vec{r}_1 s_1 | n_1 \rangle & \langle \vec{r}_2 s_2 | n_1 \rangle \\ \langle \vec{r}_1 s_1 | n_2 \rangle & \langle \vec{r}_2 s_2 | n_2 \rangle \end{pmatrix}, \quad n_1 < n_2$$

onde $\{\langle \vec{r} s | n \rangle\}$ são uma base ortonormal de funções de onda de um electron, dependentes portanto da posição \vec{r} e da variável de spin $s = \pm 1/2$. Para fixar idéias é conveniente considerar funções de onda escritas como produtos de uma função da posição por uma função associada ao spin e explicitar os rótulos n como $n \equiv \{k, m\}$, onde k identifica a função de posição e $m = \pm 1/2$ caracteriza o autovalor da terceira componente do spin. Desse modo

$$\langle \vec{r} s | n \rangle \rightarrow \langle \vec{r} s | k, m \rangle = u_k(\vec{r}) \chi_m(s)$$

onde

$$\chi_{\frac{1}{2}}(s) = \begin{cases} 1 & \text{para } s = 1/2 \\ 0 & \text{para } s = -1/2 \end{cases} \quad \text{e} \quad \chi_{-\frac{1}{2}}(s) = \begin{cases} 0 & \text{para } s = 1/2 \\ 1 & \text{para } s = -1/2 \end{cases}.$$

Os elementos de matriz $-\langle n_1 n_2 | V_c \otimes \hat{1}_S | n_3 n_4 \rangle_-$ podem então ser escritos explicitamente em termos das funções de onda de um electron como

$$\begin{aligned} & -\langle n_1 n_2 | V_c \otimes \hat{1}_S | n_3 n_4 \rangle_- = \\ & \frac{1}{2} \sum_{s_1 s_2} \int d^3 r_1 \int d^3 r_2 \left(u_{k_1}^*(\vec{r}_1) \chi_{m_1}^*(s_1) u_{k_2}^*(\vec{r}_2) \chi_{m_2}^*(s_2) - u_{k_1}^*(\vec{r}_2) \chi_{m_1}^*(s_2) u_{k_2}^*(\vec{r}_1) \chi_{m_2}^*(s_1) \right) \times \\ & \times \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \left(u_{k_3}(\vec{r}_1) \chi_{m_3}(s_1) u_{k_4}(\vec{r}_2) \chi_{m_4}(s_2) - u_{k_3}(\vec{r}_2) \chi_{m_3}(s_2) u_{k_4}(\vec{r}_1) \chi_{m_4}(s_1) \right). \end{aligned}$$

Os quatro termos dessa expressão se reduzem na realidade a dois, pois os dois termos positivos são iguais entre si, o mesmo acontecendo com os dois termos negativos. Colecionando os termos iguais o fator $1/2$ é cancelado, levando ao resultado

$$\begin{aligned}
& -\langle n_1 n_2 | V_c \otimes \hat{1}_S | n_3 n_4 \rangle_- = \\
& \sum_{s_1 s_2} \int d^3 r_1 \int d^3 r_2 u_{k_1}^*(\vec{r}_1) \chi_{m_1}^*(s_1) u_{k_2}^*(\vec{r}_2) \chi_{m_2}^*(s_2) \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} u_{k_3}(\vec{r}_1) \chi_{m_3}(s_1) u_{k_4}(\vec{r}_2) \chi_{m_4}(s_2) - \\
& - \sum_{s_1 s_2} \int d^3 r_1 \int d^3 r_2 u_{k_1}^*(\vec{r}_2) \chi_{m_1}^*(s_2) u_{k_2}^*(\vec{r}_1) \chi_{m_2}^*(s_1) \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} u_{k_3}(\vec{r}_1) \chi_{m_3}(s_1) u_{k_4}(\vec{r}_2) \chi_{m_4}(s_2).
\end{aligned} \tag{7.5}$$

O primeiro dos dois termos desta última expressão é usualmente chamado “termo direto”, enquanto o segundo é chamado “termo de troca”. De um ponto de vista de tipo perturbativo, eles podem ser entendidos como associados a “transições” experimentadas pelos dois electrons entre os vetores da base produto, devidas à ação de sua interação coulombiana. O elemento de matriz completo corresponde, nesse contexto, à amplitude que caracteriza uma transição na qual os electrons que ocupam inicialmente os estados de uma partícula $\{k_3, m_3\}$ e $\{k_4, m_4\}$ passam a ocupar os estados $\{k_1, m_1\}$ e $\{k_2, m_2\}$. Devido ao fato de que os dois electrons são férmions idênticos, essa amplitude é dada como a diferença de uma amplitude direta, em que o estado $\{k_3, m_3\}$ dá lugar ao estado $\{k_1, m_1\}$, e o estado $\{k_4, m_4\}$ ao estado $\{k_2, m_2\}$, e uma amplitude de troca, em que o estado $\{k_3, m_3\}$ dá lugar ao estado $\{k_2, m_2\}$, e o estado $\{k_4, m_4\}$ ao estado $\{k_1, m_1\}$. No caso de duas partículas distinguíveis, essas duas amplitudes correspondem a dois processos fisicamente distintos, no sentido de que existem observáveis do sistema de duas partículas que permitem distingui-los um do outro. A identidade dos dois electrons equivale porém à impossibilidade de distinguir entre os dois processos, de modo que as respectivas amplitudes devem ser combinadas numa amplitude única. O sinal relativo obtido entre os dois termos se deve ao caráter anti-simétrico dos estados de dois férmions idênticos. De fato, é fácil ver que um cálculo correspondente para dois bosons idênticos também leva a uma amplitude global escrita em termos de uma amplitude direta e a uma amplitude de troca, mas com sinal relativo oposto ao que resulta no caso fermiônico.

Uma consequência importante da ocorrência de efeitos de troca na interação coulombiana de dois electrons é o fato de que, embora a interação coulombiana seja independente de spin, os elementos de matriz dependem dos estados de spin dos electrons envolvidos. Isso pode ser visto de forma particularmente clara no caso de elementos de matriz diagonais, isto é, com $\{k_1, m_1\} = \{k_3, m_3\}$ e $\{k_2, m_2\} = \{k_4, m_4\}$. De fato, usando o fato de que a interação coulombiana age sobre os spins como o operador identidade, o termo direto fica nesse caso dado por

$$-\langle n_1 n_2 | V_c \otimes \hat{1}_S | n_3 n_4 \rangle_-^{(\text{dir})} = \int d^3 r_1 \int d^3 r_2 |u_{k_1}(\vec{r}_1) u_{k_2}(\vec{r}_2)|^2 \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \quad (7.6)$$

enquanto o termo de troca é dado por

$$\begin{aligned} & -\langle n_1 n_2 | V_c \otimes \hat{1}_S | n_3 n_4 \rangle_-^{(\text{troca})} = \\ & - \int d^3 r_1 \int d^3 r_2 u_{k_1}^*(\vec{r}_1) u_{k_2}^*(\vec{r}_2) u_{k_2}(\vec{r}_1) u_{k_1}(\vec{r}_2) \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \sum_{s_1 s_2} \chi_{m_1}^*(s_1) \chi_{m_2}^*(s_2) \chi_{m_2}(s_1) \chi_{m_1}(s_2) \end{aligned} \quad (7.7)$$

que depende dos estados de spin através do último fator, que pode ser escrito simplesmente em termos do quadrado de um produto escalar como

$$|(\chi_{m_1}, \chi_{m_2})|^2 = \delta_{m_1 m_2}$$

isto é, ele é igual a 1 para spins paralelos e igual a zero para spins anti-paralelos. Essa dependência de spin proveniente de efeitos de troca é crucial para as estudar as propriedades magnéticas dos materiais, por exemplo.

Um fato de importância prática considerável para cálculos, como o precedente, de elementos de matriz de operadores simétricos entre estados completamente simétricos ou anti-simétricos é que, a menos de fatores de normalização, basta que um dos dois estados envolvidos seja devidamente simetrizado ou anti-simetrizado. Isso provém de que um operador simétrico comuta com o operador de projeção sobre o sub-espço simétrico ou anti-simétrico. De fato, como $[g, \mathcal{A}] = 0$ segue que

$$\langle a | \mathcal{A} g \mathcal{A} | b \rangle = \langle a | g \mathcal{A}^2 | b \rangle = \langle a | g \mathcal{A} | b \rangle$$

sendo que relações análogas valem para o caso de \mathcal{S} . No caso acima dos elementos de matriz da interação coulombiana entre dois electrons, o cálculo pode ser feito de forma mais simples através de (deve ser notada a não anti-simetrização do bra)

$$\begin{aligned} & \langle n_1 n_2 | V_c | n_3 n_4 \rangle_- = \\ & \sum_{s_1 s_2} \int d^3 r_1 \int d^3 r_2 \langle n_1 | \vec{r}_1 s_1 \rangle \langle n_2 | \vec{r}_2 s_2 \rangle \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \frac{1}{\sqrt{2}} \det \begin{pmatrix} \langle \vec{r}_1 s_1 | n_3 \rangle & \langle \vec{r}_2 s_2 | n_3 \rangle \\ \langle \vec{r}_1 s_1 | n_4 \rangle & \langle \vec{r}_2 s_2 | n_4 \rangle \end{pmatrix} \end{aligned}$$

que reproduz o resultado já obtido a menos de um fator de normalização. Este pode porém ser facilmente recuperado notando que

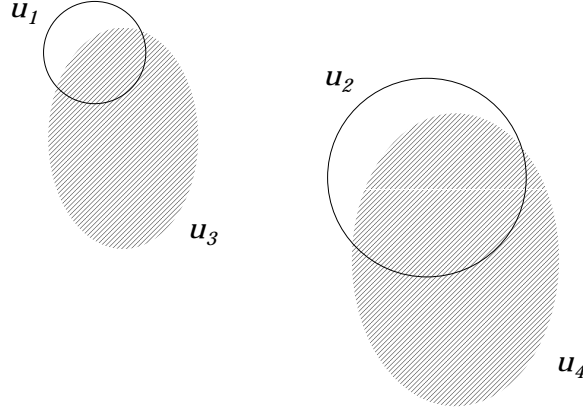


Figura 7.1: Domínios em que as funções localizadas de uma partícula $u_i(\vec{r})$ são diferentes de zero.

$$|n_1 n_2\rangle_- = \sqrt{2}\mathcal{A}|n_1 n_2\rangle$$

de modo que

$$-\langle n_1 n_2 | V_c | n_3 n_4 \rangle_- = \sqrt{2} \langle n_1 n_2 | V_c | n_3 n_4 \rangle_-.$$

A importância quantitativa das contribuições de troca em elementos de matriz envolvendo partículas idênticas depende também do grau de localização ou deslocalização das partículas nos estados envolvidos, e isso permite, em particular, fazer contacto com a distinguibilidade relativa de partículas clássicas referida inicialmente. A forma mais simples de ver isso é através de um exemplo concreto em que, na expressão (7.5), os estados “iniciais” de uma partícula $u_3(\vec{r})$ e $u_4(\vec{r})$ por um lado, e os estados “finais” $u_1(\vec{r})$ e $u_2(\vec{r})$ por outro lado, são localizados em regiões distintas do espaço da forma indicada na figura 7.1. É fácil ver então que esse arranjo de domínios é tal que o termo de troca se anula quaisquer que sejam as funções de spin $\chi_{m_i}(s)$

$$\int d^3 r_1 \int d^3 r_2 u_1^*(\vec{r}_2) u_2^*(\vec{r}_1) \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} u_3(\vec{r}_1) u_4(\vec{r}_2) \sum_{s_1 s_2} \chi_{m_1}^*(s_2) \chi_{m_2}^*(s_1) \chi_{m_3}(s_1) \chi_{m_4}(s_2) = 0$$

pois, qualquer que seja \vec{r} , tem-se que

$$u_1^*(\vec{r}) u_4(\vec{r}) = 0, \quad u_2^*(\vec{r}) u_3(\vec{r}) = 0.$$

O termo direto, porém, é diferente de zero, dado que

$$u_1^*(\vec{r})u_3(\vec{r}) \neq 0, \quad \text{e} \quad u_2^*(\vec{r})u_4(\vec{r}) \neq 0$$

o que permite identificar, por exemplo, a partícula que ocupa o estado final $u_3(\vec{r})$ com a que ocupava o estado inicial $u_1(\vec{r})$. Esse exemplo permite entender de que forma a inexistência, na teoria clássica, de qualquer limite para a localizabilidade das partículas leva à não existência de efeitos de troca, bem como à distinguibilidade relativa de partículas idênticas através da continuidade de suas trajetórias no espaço de fases.

2. A existência de amplitudes de troca tem conseqüências importantes também para as seções de choque de espalhamento de um par de partículas idênticas. No caso de partículas *distinguíveis*, o espalhamento devido a um potencial que dependente da distância relativa entre as partículas e independente dos spins (como é o caso de V_c , por exemplo), cujos momentos iniciais (no sistema de centro de massa) e estados de spin são $\vec{k}_i^{(1)}$, $\chi^{(1)}(s)$ e $\vec{k}_i^{(2)} = -\vec{k}_i^{(1)}$, $\chi^{(2)}(s)$, e cujos momentos e estados de spin finais são $\vec{k}_f^{(1)}$, $\chi^{(1)}(s)$ e $\vec{k}_f^{(2)} = -\vec{k}_f^{(1)}$, $\chi^{(2)}(s)$, é descrito por uma amplitude de espalhamento $f(\theta)$ em termos da qual a seção de choque diferencial é dada por

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2$$

onde θ é o ângulo de espalhamento, dado por

$$\cos \theta = \frac{\vec{k}_i^{(1)} \cdot \vec{k}_f^{(1)}}{k_i^{(1)} k_f^{(1)}}.$$

Como foi suposto que a interação é independente de spin, os estados de spin finais são iguais aos estados iniciais. É importante notar que esses estados $\chi^{(1)}(s)$ e $\chi^{(2)}(s)$ não são necessariamente autoestados de s_3 . Em geral eles são estados de spin que dependem do particular processo de preparação que a que tenham sido submetidos o projétil e o alvo.

Para passar ao caso do espalhamento de duas partículas idênticas, é útil escrever a amplitude de espalhamento em termos de um operador de transição T definido por

$$\langle \vec{k}_f^{(1)} \chi^{(1)}, \vec{k}_f^{(2)} \chi^{(2)} | T | \vec{k}_i^{(1)} \chi^{(1)}, \vec{k}_i^{(2)} \chi^{(2)} \rangle \equiv f(\theta)$$

que, como a interação responsável pelo espalhamento, é também independente dos spins. Como no caso de partículas idênticas os estados inicial e final devem ser devidamente simetrizados (para bosons) ou anti-simetrizados (para férmions), a amplitude de espalhamento adquire contribuições de troca, sendo dada por

$$\begin{aligned}
f_{\pm}(\theta) &= \langle \vec{k}_f^{(1)} \chi^{(1)}, \vec{k}_f^{(2)} \chi^{(2)} | T | \vec{k}_i^{(1)} \chi^{(1)}, \vec{k}_i^{(2)} \chi^{(2)} \rangle \pm \langle \vec{k}_f^{(1)} \chi^{(1)}, \vec{k}_f^{(2)} \chi^{(2)} | T | \vec{k}_i^{(2)} \chi^{(2)}, \vec{k}_i^{(1)} \chi^{(1)} \rangle \\
&= f(\theta) \pm f(\pi - \theta) \left| \langle \chi^{(1)} | \chi^{(2)} \rangle \right|^2
\end{aligned}$$

onde os sinais $+$ e $-$ se referem ao caso de bosons e f rmions respectivamente, e foi usado o fato de que a transi  o de $\vec{k}_i^{(2)}$ para $\vec{k}_f^{(1)}$ corresponde a um  ngulo de espalhamento $\pi - \theta$. A se  o de choque de espalhamento, tendo em conta os efeitos de troca associados   identidade das part culas,   portanto dada por

$$\frac{d\sigma}{d\Omega}(\theta) = \left| f(\theta) \pm f(\pi - \theta) \langle \chi^{(1)} | \chi^{(2)} \rangle \right|^2$$

que mostra efeitos de interfer ncia entre as duas amplitudes que, em particular, exp e   observa  o a sua fase relativa. Por exemplo, no caso do espalhamento coulombiano de dois electrons com spins paralelos (isto  , $\chi^{(1)} = \chi^{(2)}$), a se  o de choque diferencial   dada por

$$\frac{d\sigma}{d\Omega}(\theta) = |f_c(\theta) - f_c(\pi - \theta)|^2 \quad (7.8)$$

onde $f_c(\theta)$   a amplitude de espalhamento coulombiana dada por

$$f_c(\theta) = -\frac{\eta}{2k \sin^2 \frac{\theta}{2}} \exp \left[-i\eta \log \left(\sin^2 \frac{\theta}{2} \right) + 2i\sigma_0 \right] \quad (7.9)$$

sendo o par metro de Sommerfeld, η , e a defasagem coulombiana da onda s , σ_0 , dados respectivamente por

$$\eta = \frac{\mu e^2}{\hbar^2 k} \quad \text{e} \quad \sigma_0 = \arg \Gamma(1 + i\eta).$$

Nessas express es k   o vetor de onda associado ao movimento relativo dos dois electrons e μ   a sua massa reduzida. A se  o de choque (7.8)   marcadamente distinta da se  o de choque de Rutherford. De fato, devido aos efeitos de troca ela se torna sim trica em torno de $\theta = \pi/2$ e adquire um termo de interfer ncia

$$\left(\frac{d\sigma}{d\Omega} \right)^{(\text{int})} = -\frac{2\eta^2}{k^2 \sin^2 \theta} \cos \left[\eta \log \left(\tan^2 \frac{\theta}{2} \right) \right]$$

que tem um comportamento fortemente oscilat rio nas vizinhan as de $\theta = 0$ e $\theta = \pi$. Por outro lado, no caso de spins anti-paralelos (isto  , $\langle \chi^{(1)} | \chi^{(2)} \rangle = 0$), n o h  contribui  es de troca para

a amplitude de espalhamento e a seção de choque se reduz à seção de choque de Rutherford. Isso pode ser entendido notando que, sendo a interação coulombiana independente dos spins, o estado de spin dos electrons é preservado durante o processo de colisão, o que permite, nesse caso, a identificação das partículas através de seus estados de spin.

3. A natureza simétrica (anti-simétrica) dos estados de um sistema de muitos bosons (férmions) idênticos implica na existência de correlações entre as partículas, mesmo na ausência de qualquer interação entre elas. Essas correlações se manifestam de forma particularmente conspícua no caso de um estado de muitos férmions idênticos, pois a amplitude de probabilidade correspondente ao vetor de estado do setor anti-simétrico $\langle \vec{r}_1 s_1, \dots, \vec{r}_N s_N | \mathcal{A} | a \rangle$ se anula sempre que dois (ou mais) férmions tiverem as mesmas coordenadas, isto é $\vec{r}_i = \vec{r}_j$, $s_i = s_j$, o que implica evidentemente que o estado contém informação sobre as coordenadas *relativas* dos férmions.

Um sistema simples que exhibe efeitos dessas correlações é um gás de férmions sem interação, cujo Hamiltoniano é

$$H = \sum_{i=1}^N \frac{p_i^2}{2M}.$$

O estado fundamental desse sistema, chamado “gás de Fermi degenerado”, pode ser visto como um modelo simplificado para os electrons de condução de um metal, no qual as interações Coluombianas entre electrons bem como entre electrons e ions positivos da rede atômica são substituídas por interações de cada electron com potenciais médios uniformes devidos aos outros electrons e aos ions positivos da rede, balanceados de forma a garantir a neutralidade elétrica do sistema. Os estados de um electron podem então ser tomados convenientemente como autoestados de momento com condições de contorno periódicas num volume L^3 e de s_3 . As funções de onda de um electron são portanto

$$\langle \vec{r} s | \vec{k} m \rangle = \frac{1}{\sqrt{L^3}} e^{i\vec{k} \cdot \vec{r}} \chi_m(s)$$

onde $\hbar \vec{k}$ e $m = \pm 1/2$ são autovalores do momento e de s_3 respectivamente. Os autovetores do setor anti-simétrico do Hamiltoniano H que correspondem a um sistema de N electrons livres no volume L^3 podem portanto ser escritos como determinantes de Slater

$$\langle \vec{r}_1 s_1, \dots, \vec{r}_N s_N | \vec{k}_1 m_1, \dots, \vec{k}_N m_N \rangle = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \langle \vec{r}_1 s_1 | \vec{k}_1 m_1 \rangle & \dots & \langle \vec{r}_N s_1 | \vec{k}_1 m_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \vec{r}_1 s_N | \vec{k}_N m_N \rangle & \dots & \langle \vec{r}_N s_N | \vec{k}_N m_N \rangle \end{pmatrix}$$

sendo que os autovalores correspondentes são dados por

$$E = \sum_{i=1}^N \frac{\hbar^2 k_i^2}{2M}.$$

Como todos os estados de um electron envolvidos na construção do determinante devem ser diferentes sob pena de seu anulamento, o estado fundamental corresponde ao determinante construído com os N estados de um electron de menor momento, levando em conta a possibilidade de ter dois electrons com o mesmo \vec{k} desde que com funções de onda de spin diferentes. Isso leva então à utilização de todos os estados de momento até um determinado valor máximo k_F (chamado *momento de Fermi* do gás de degenerado) que depende do número de electrons N .

Para determinar k_F em termos de N basta notar que deve-se ter, supondo que os estados de um electron tenham sido rotulados em ordem crescente de k ,

$$2 \sum_{k \leq k_F} 1 = N$$

onde o fator 2 tem em conta a existência de dois estados de spin para cada valor do momento. Transformando essa soma numa integral com a prescrição usual para condições de contorno periódicas resulta

$$N = 2 \sum_{k \leq k_F} 1 = 2 \frac{L^3}{(2\pi)^3} \int_{k \leq k_F} 1 d^3k = 2 \frac{L^3}{(2\pi)^3} \frac{4\pi}{3} k_F^3.$$

Introduzindo a densidade de electrons $\rho \equiv N/L^3$ resulta portanto a relação com o momento de Fermi

$$\rho = \frac{N}{L^3} = \frac{k_F^3}{3\pi^2}.$$

Essa relação pode ser entendida qualitativamente com um argumento que envolve as relações de incerteza posição momento. De fato, a restrição do momento dos electrons a uma faixa de largura $\sim \hbar k_F$, a sua posição só poderá ser definida dentro de intervalos cuja dimensão é da ordem de $1/k_F$. Por outro lado, como as correlações impostas pela anti-simetria impedem dois electrons de se encontrarem na mesma posição, é necessário alocar um volume espacial da ordem de $1/k_F^3$ por electron do gás, o que corresponde a um volume total, para N electrons,

$$V \sim \frac{N}{k_F^3} \quad \text{ou seja} \quad \rho \sim k_F^3.$$

A energia cinética de um electron com o momento de Fermi é chamada a *energia de Fermi* E_F . Ela está relacionada com a densidade de electrons por

$$E_F = \frac{\hbar^2 k_F^2}{2M} = \frac{\hbar^2}{2M} (3\pi^2 \rho)^{\frac{2}{3}}. \quad (7.10)$$

A energia E_0 do estado fundamental do gás de electrons livres pode ser também facilmente expressa em termos do momento de Fermi (ou, alternativamente, da densidade ou da energia de Fermi) calculando a soma das energias cinéticas por meio da integral

$$E_0 = 2 \frac{L^3}{(2\pi)^3} \int_{k \leq k_F} \frac{\hbar^2 k^2}{2M} d^3 k = \frac{L^3}{5\pi^2} \frac{\hbar^2 k_F^5}{2M} = \frac{3}{5} N E_F$$

que mostra que a energia média por partícula é $3/5$ da energia de Fermi e, em vista da relação (7.10), depende da potência $2/3$ da densidade. Essa dependência da energia do estado fundamental com a densidade pode ser interpretada em termos de uma pressão exercida pelo gás, que pode ser calculada através da relação

$$P_F = -\frac{dE_0}{d(L^3)} = \frac{3}{5} \frac{\hbar^2}{2M} (3\pi^2)^{\frac{2}{3}} \rho^{\frac{5}{3}} \quad (7.11)$$

obtida igualando a variação da energia do estado fundamental do gás ao trabalho efetuado contra a pressão num processo adiabático de compressão. Essa *pressão de Fermi* pode ser vista como um resultado do “princípio de exclusão”, que força a ocupação de estados de momento diferente de zero no estado fundamental (temperatura zero) do gás de Fermi.

7.2 A tabela periódica. Aproximações autoconsistentes para a estrutura atômica.

Um átomo, visto como um sistema composto de um núcleo compacto, carregado positivamente, ao qual estão ligados electrons carregados negativamente, é um sistema quântico cuja estrutura é determinada de forma essencial pela identidade e pelo caráter fermiônico dos electrons, que desempenha também um papel crucial nos processos de interação entre vários sistemas atômicos que dão origem, entre outras coisas, às estruturas moleculares. Esta seção se limita no entanto a aspectos qualitativos gerais da estrutura atômica revelados nas regularidades que permitem a construção da tabela periódica dos elementos, bem como a aproximações que permitem tratar quantitativamente essas regularidades.

Omitindo complicações ligadas às interações dependentes de spin e outros efeitos de natureza magnética, o Hamiltoniano mais simples que se pode considerar para um sistema atômico com

um núcleo de carga Ze e com Z electrons envolve apenas interações coulombianas. Ignorando efeitos de recuo com base na ordem de grandeza relativa das massas nuclear e eletrônica, ele pode ser escrito como

$$\begin{aligned} H &= \sum_{i=1}^Z \left(\frac{p_i^2}{2m_e} - \frac{Ze^2}{r_i} \right) + \frac{1}{2} \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} = \\ &= \sum_{i=1}^Z \left(\frac{p_i^2}{2m_e} + V(r_i) \right) + \left[\frac{1}{2} \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} - \sum_{i=1}^Z \frac{Ze^2}{r_i} - \sum_{i=1}^Z V(r_i) \right]. \end{aligned} \quad (7.12)$$

Na segunda forma do Hamiltoniano foi introduzido um potencial auxiliar de uma partícula $V(r_i)$ com a idéia que, quando convenientemente escolhido, possa tornar “pequenos” os efeitos do último parênteses. É importante notar desde logo que esse objetivo deve ser fortemente qualificado por razões de princípio, pois esse parênteses inclui interações coulombianas de dois corpos (a repulsão coulombiana entre dois electrons) que dá lugar a correlações dinâmicas entre os electrons que não podem ser descritas por um potencial de um corpo como $V(r)$. A expectativa deve se restringir portanto à de que propriedades atômicas que não dependam dessas correlações de forma essencial possam ser bem descritas considerando um potencial médio de um corpo $V(r)$ apenas. As correlações *cinemáticas* devidas ao caráter fermiônico dos electrons devem, em qualquer caso, ser tratadas exatamente. A utilidade de um tratamento aproximado desse tipo provém de que de fato existem várias propriedades atômicas de interesse que sobrevivem ao truncamento das correlações dinâmicas.

Existem vários métodos desenvolvidos para a determinação do potencial médio $V(r)$. Em termos qualitativos, o que se deve esperar para o potencial médio experimentado por um dos electrons de um átomo neutro é que o seu comportamento se aproxime do potencial coulombiano nuclear $-Ze^2/r$ para valores de r pequenos com relação ao raio atômico e do potencial coulombiano de uma única unidade de carga $-e^2/r$ para valores grandes de r , devido à blindagem de parte da carga positiva nuclear pelos $Z - 1$ electrons restantes (figura 7.2). Supondo que o sistema atômico seja esfericamente simétrico, esse potencial será central e terá estados estacionários de partícula independente rotulados por números quânticos n , l e m_l , que são respectivamente um “número quântico principal” análogo ao introduzido no espectro do átomo de hidrogênio, o momento angular orbital e sua projeção no eixo de quantização em unidades de \hbar . Para caracterizar completamente os estados eletrônicos é preciso introduzir ainda o estado de spin, caracterizado pelo autovalor $m_s = \pm 1/2$ de sua componente ao longo do eixo de quantização, s_3 (também em unidades de \hbar). O modelo mais simples para a estrutura de um átomo de Z electrons consiste em tratá-lo como um sistema de Z electrons *independentes* (isto é, sem interações entre eles), que se movem em um potencial médio central com as características

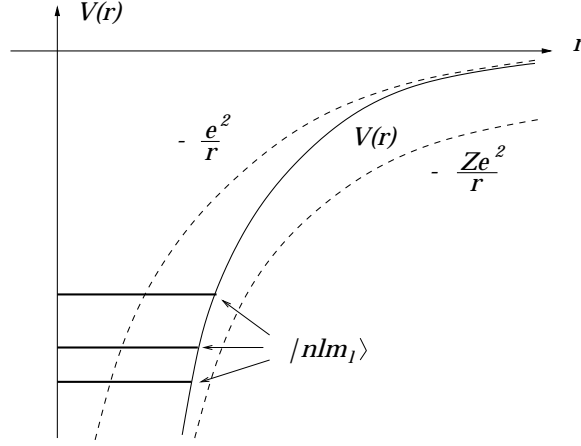


Figura 7.2: Propriedades qualitativas do potencial médio para um electron de um átomo neutro.

indicadas. Os estados estacionários do átomo são, nessa aproximação, os estados estacionários do Hamiltoniano

$$H = \sum_{i=1}^Z \left(\frac{p_i^2}{2m_e} + V(r_i) \right)$$

que, devido ao caráter fermiônico dos electrons, são determinantes de Slater formados a partir de Z estados estacionários distintos de um electron no potencial médio $V(r)$

$$\langle \vec{r}s | nlm_l \ 1/2 m_s \rangle = \phi_{nlm_l}(\vec{r}) \chi_{m_s}(s) \quad (7.13)$$

onde os chamados *orbitais* $\phi_{nlm_l}(\vec{r})$ satisfazem

$$\left(\frac{p^2}{2m_e} + V(r) \right) \phi_{nlm_l}(\vec{r}) = \epsilon_{nl} \phi_{nlm_l}(\vec{r}).$$

Devido ao spin do electron, cada um dos orbitais $\phi_{nlm_l}(\vec{r})$ pode ser ocupado por até dois electrons. Os autovalores de H correspondentes a esses estados determinantis é simplesmente a soma dos autovalores ϵ_{nl} associados aos orbitais incluídos no determinante, de modo que o estado de menor autovalor corresponde ao determinante formado a partir dos Z estados de um electron (7.13) de menor autovalor ϵ_{nl} .

Em termos também qualitativos, o espectro de orbitais no potencial médio tem as características indicadas na figura 7.3. Como o potencial médio se aproxima do potencial coulombiano

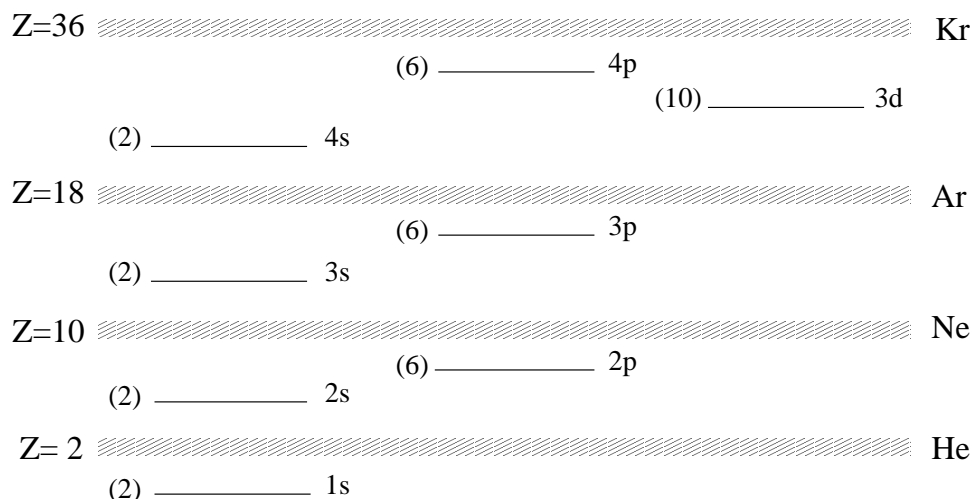


Figura 7.3: Propriedades qualitativas dos orbitais de um potencial médio com as características indicadas na figura 7.2.

do átomo de hidrogênio para valores grandes de r e é *mais atrativo* que esse potencial para valores menores de r , os efeitos repulsivos do potencial centrífugo faz com que orbitais com valores menores de l (para os quais a região de maior atração é mais acessível) sejam relativamente mais ligados que no caso do átomo de hidrogênio. As características importantes desse esquema de níveis são o grau de degenerescência de cada nível e a irregularidade dos espaçamentos, que faz com que apareçam intervalos maiores (assinalados na figura 7.3 pelas faixas hachuradas) entre determinados grupos de níveis. Os números entre parênteses à esquerda de cada nível indicam a degenerescência total (incluindo a que é associada ao spin dos electrons) que corresponde ao maior número de electrons que ele pode conter. Esses números são totalizados até cada uma das faixas hachuradas, e correspondem aos elementos indicados.

Essas propriedades dão conta, qualitativamente, do fato experimental de que as energias de ligação e as energias de ionização dos estados fundamentais dos átomos não variam suavemente como função de Z , mas apresentam descontinuidades para valores particulares do número atômico conhecidos como “números mágicos”. Experimentalmente os números mágicos são 2, 10, 18, 36, ... que correspondem à ocupação total dos níveis abaixo de cada uma das faixas hachuradas. A existência dos números mágicos está portanto vinculada à existência de “camadas” disjuntas no espectro de partícula independente do potencial médio, bem como ao caráter fermiônico dos electrons que implica nas restrições impostas pelo princípio de exclusão para a ocupação dos estados de uma partícula.

Correlações dinâmicas induzidas pelos termos de interação de dois corpos no Hamiltonia-

no, não consideradas na aproximação de partículas independentes, têm em geral o efeito de atenuar os “efeitos de camada” que dão origem, em particular, aos “números mágicos”, de modo que a observação de tais efeitos de camada pode ser tida como uma evidência importante em favor da validade de uma aproximação de partículas independentes submetidas a um potencial médio. Isso se aplica também a outros sistemas, além dos átomos, em que efeitos de camada são também observados, notadamente aos núcleos atômicos, onde o campo médio é devido unicamente às interações atrativas entre os nucleons. Devido ao caráter de curto alcance dessas interações, bem como a outros elementos interpretativos que apontam na direção da importância de efeitos das correlações dinâmicas nesse caso (notadamente o sucesso da hipótese de “núcleo composto” introduzida por N. Bohr para descrever o comportamento do espalhamento de neutrons lentos por núcleos), a resistência para a adoção de um modelo de partículas independentes no caso nuclear foi considerável, tendo se imposto apenas na década de 50, devido, em particular, à formulação por Mayer, Jensen e Wigner de um potencial médio capaz de dar conta de regularidades em propriedades nucleares interpretadas em termos de efeitos de camada.

7.2.1 Potenciais médios auto-consistentes. Aproximação de Thomas-Fermi.

A distribuição de electrons no átomo é um ingrediente essencial para qualquer determinação que se pretenda fazer do potencial médio $V(\vec{r})$. Por outro lado, na medida em que uma descrição em termos de um potencial médio seja adequada, ele é um ingrediente essencial para a determinação da distribuição de electrons. Essa interdependência mútua do potencial médio e da densidade eletrônica conduz naturalmente ao *critério de auto-consistência* para a determinação do potencial médio, que consiste na exigência de que a densidade eletrônica resultante do potencial seja igual à densidade eletrônica que o determina.

Para implementar esse critério é necessário portanto dispor, por um lado, de um procedimento que permita determinar um potencial médio a partir de uma distribuição eletrônica dada e também, por outro lado, de um segundo procedimento que determine a densidade eletrônica a partir de um potencial médio dado. Uma forma simples de obter um potencial médio a partir de uma distribuição de cargas consiste em utilizar a equação de Poisson da eletrostática, enquanto a determinação da distribuição de cargas no átomo pode ser feita a partir das funções de onda de um electron determinadas pelo potencial médio considerado. A idéia de obter um potencial médio auto-consistente dessa forma constitui a base do chamado método de Hartree (e de Hartree-Fock) que será discutido na seção seguinte. Embora conceitualmente simples, eles são computacionalmente trabalhosos, pois dependem da obtenção explícita e reiterada das funções de onda de um electron no potencial médio considerado, com o objetivo de satisfazer

o critério de auto-consistência. Existe, no entanto, um método devido a Thomas¹ e a Fermi² que, mediante certas aproximações, permite simplificar esse processo de forma substancial.

O método de Thomas-Fermi pode ser formulado através da suposição de que a relação (7.10) entre a energia de Fermi e a densidade de um gás de Fermi degenerado seja válida *localmente* para a nuvem eletrônica do átomo. Dessa forma ela é re-escrita como

$$E_F(\vec{r}) \simeq \frac{\hbar^2}{2m_e} \left(3\pi^2 \rho(\vec{r}) \right)^{\frac{2}{3}} \quad (7.14)$$

o que corresponde a admitir que a densidade da nuvem eletrônica num determinado ponto pode ser associada a uma “energia de Fermi local” através dessa relação. Uma forma de justificar esse procedimento é associa-lo à suposição de que os electrons contidos em um pequeno volume dV em torno do ponto \vec{r} , dentro do qual a densidade possa ser tratada como praticamente uniforme, se comportam essencialmente como os electrons de um gás de Fermi degenerado com a mesma densidade. Um ponto que limita a validade dessa suposição é o de que devem ser levadas em conta restrições impostas pelas relações de incerteza, segundo as quais as dimensões dl associadas a dV implicam numa indefinição dos vetores de onda da ordem de $\Delta k = 1/dl$. A importância dessa limitação pode ser medida pela relação entre essa indefinição dos momentos e o suposto momento de Fermi local $k_F(\vec{r}) = [3\pi^2 \rho(\vec{r})]^{1/3}$, e será menos grave quando $\Delta k \ll k_F(\vec{r})$ ou seja $3\pi^2 \rho(\vec{r}) dV \gg 1$. Isso torna o método de Thomas-Fermi mais preciso nas regiões de maior densidade eletrônica, e pouco confiável na região superficial do átomo, em que a densidade eletrônica se aproxima de zero.

Admitida a relação (7.14), no entanto, é possível estabelecer uma relação entre o potencial médio $V(\vec{r})$ e a densidade pondo, para um átomo neutro

$$V(\vec{r}) + E_F(\vec{r}) = V(\vec{r}) + \frac{\hbar^2}{2m_e} \left(3\pi^2 \rho(\vec{r}) \right)^{\frac{2}{3}} \simeq 0$$

em qualquer ponto \vec{r} da nuvem eletrônica. Isso significa supor que a energia total máxima em cada ponto, dada pela soma da máxima energia cinética local (representada por $E_F(\vec{r})$) com a energia potencial local (atrativa, representada por $V(\vec{r})$) seja essencialmente nula em todos os pontos. O potencial $V(\vec{r})$ está sendo tomado portanto como o potencial mais fracamente atrativo que, dentro das aproximações feitas, é ainda capaz de ligar todos os electrons. O fato de que o átomo neutro é ligado implica também que $V(\vec{r}) \rightarrow 0$ e também $E_F(\vec{r}) \rightarrow 0$ para $r \rightarrow \infty$, pois nessas condições $\rho(\vec{r}) \rightarrow 0$ para $r \rightarrow \infty$. Esta última relação pode então ser utilizada como uma forma aproximada de obter a densidade eletrônica a partir do potencial independentemente do cálculo de funções de onda eletrônicas. A densidade assim obtida é

¹L. H. Thomas, *Proc. Cambridge Phil. Soc.* **23**, 543 (1927).

²E. Fermi, *Z. Physik* **48**, 73 (1928).

$$\rho(\vec{r}) = \frac{1}{3\pi^2} \left(-\frac{2m_e}{\hbar^2} V(\vec{r}) \right)^{\frac{3}{2}}.$$

Por outro lado, o potencial $V(\vec{r})$ pode ser obtido de uma densidade eletrônica dada utilizando a equação de Poisson. De fato, a densidade eletrônica $\rho(\vec{r})$ dá lugar a uma densidade de carga

$$\rho_c(\vec{r}) = -e\rho(\vec{r})$$

com a qual a equação de Poisson para o potencial eletrostático $\Phi(\vec{r})$ devido à distribuição de cargas no átomo é

$$\begin{aligned} \nabla^2 \Phi(\vec{r}) &= -4\pi \left[Ze\delta(\vec{r}) - \frac{e}{3\pi^2} \left(-\frac{2m_e}{\hbar^2} V(\vec{r}) \right)^{\frac{3}{2}} \right] \\ &= -4\pi \left[Ze\delta(\vec{r}) - \frac{e}{3\pi^2} \left(\frac{2m_e e}{\hbar^2} \Phi(\vec{r}) \right)^{\frac{3}{2}} \right] \end{aligned}$$

onde foi utilizada a relação $V(\vec{r}) = -e\Phi(\vec{r})$ entre o potencial eletrostático e o potencial médio experimentado por um electron.

Supondo a simetria esférica do potencial $\Phi(\vec{r})$ (e conseqüentemente também da distribuição eletrônica), essa equação se reduz a

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Phi(r)}{dr} \right) = \frac{4e}{3\pi} \left(\frac{2m_e e}{\hbar^2} \Phi(r) \right)^{\frac{3}{2}} \quad (7.15)$$

com as condições de contorno

$$\Phi(r) \rightarrow \begin{cases} \frac{Ze}{r} & \text{para } r \rightarrow 0 \\ 0 & \text{para } r \rightarrow \infty \end{cases} \quad (7.16)$$

que decorrem da presença da carga nuclear Ze na origem e da neutralidade do átomo, respectivamente. Como o átomo é um sistema não apenas neutro, mas também ligado, o anulamento do potencial eletrostático no infinito deve na realidade ser tal que $r\Phi(\vec{r}) \rightarrow 0$ para $r \rightarrow \infty$. Essa condição pode ser entendida como indicando que não existe carga nuclear que não tenha sido blindada por carga eletrônica a distâncias suficientemente grandes do núcleo.

Uma propriedade interessante da equação (7.15) é a de que ela possui propriedades de escalonamento que permitem escrever a solução, para *qualquer* Z , em termos de uma mesma “função universal”. Essa propriedade se torna manifesta introduzindo a variável adimensional

$$x = Z^{\frac{1}{3}} \frac{r}{b} \quad \text{com} \quad b = \frac{1}{2} \left(\frac{3\pi}{4} \right)^{\frac{2}{3}} \frac{\hbar^2}{m_e e^2} \simeq .5 \times 10^{-8} \text{ cm}$$

e escrevendo o potencial eletrostático como

$$\Phi(r) = \frac{Ze}{r} \chi(x)$$

com o que a equação (7.15) adquire a forma

$$\sqrt{x} \frac{d^2 \chi(x)}{dx^2} = \chi^{\frac{3}{2}}(x)$$

e deve ser resolvida com as condições de contorno, em vista de (7.16),

$$\chi(x) \rightarrow \begin{cases} 1 & \text{para } x \rightarrow 0 \\ 0 & \text{para } x \rightarrow \infty \end{cases} .$$

A tabulação da solução³ permite escrever o potencial autoconsistente (no sentido de Thomas-Fermi) para qualquer número atômico Z através da escolha da escala de distâncias b .

Essa enorme regularidade na variação do potencial autoconsistente de Thomas-Fermi na realidade reflete uma de suas grandes limitações, que é a de eliminar completamente os efeitos de camada responsáveis pela estrutura de tabela periódica dos elementos. A origem dessa limitação é a substituição da contribuição para a densidade eletrônica dos níveis eletrônicos individuais por quantidades obtidas a partir das propriedades de um gás de Fermi degenerado, que ignoram completamente as peculiaridades do espectro do potencial médio. Para evita-la, é preciso ter essas peculiaridades na devida conta, por exemplo, obtendo autofunções eletrônicas a partir de uma equação de autovalores com o potencial de Thomas-Fermi. Esse procedimento não é no entanto completamente autoconsistente, no sentido que a densidade eletrônica obtida a partir dessas autofunções não reproduz o potencial de Thomas-Fermi utilizado para determiná-las. Para obter efeitos de camada de forma autoconsistente é preciso, portanto, recorrer aos métodos menos esquemáticos discutidos a seguir.

³v. e.g. V. Bush e S. Caldwell, *Phys. Rev.* **38** 1898 (1931).

7.2.2 Aproximações de Hartree e de Hartree-Fock.

Na aproximação auto-consistente de Thomas-Fermi o potencial médio é relacionado à densidade eletrônica utilizando expressões gerais e simples, mas a rigor válidas apenas para um gás de Fermi degenerado. Uma forma clara de obter potenciais médios mais precisos e associados a átomos específicos consiste em substituir esse procedimento pelo cálculo efetivo da distribuição de carga devida aos Z electrons da espécie atômica considerada em termos dos estados estacionários de uma partícula submetida ao potencial médio considerado. Uma aproximação baseada nesse procedimento foi proposta por Hartree. Para implementá-la basta observar que, uma vez feita alguma escolha para o potencial médio $V(\vec{r})$ em (7.12) é possível, ignorando o último parênteses nessa expressão de H , resolver a equação de autovalores que determina os estados estacionários de uma partícula nesse potencial, e associá-los a funções de onda $\phi_\nu(\vec{r})\chi_m(s)$ (incluindo as funções de onda de spin). Com essas funções de onda se pode então construir uma função de onda anti-simetrizada de Z electrons sob a forma de um determinante de Slater

$$\langle \vec{r}_1 s_1, \dots, \vec{r}_Z s_Z | \Psi_Z \rangle = \frac{1}{\sqrt{Z!}} \det \begin{pmatrix} \phi_{k_1}(\vec{r}_1)\chi_{m_1}(s_1) & \dots & \phi_{k_1}(\vec{r}_Z)\chi_{m_1}(s_Z) \\ \dots & \dots & \dots \\ \phi_{k_Z}(\vec{r}_1)\chi_{m_Z}(s_1) & \dots & \phi_{k_Z}(\vec{r}_Z)\chi_{m_Z}(s_Z) \end{pmatrix} \quad (7.17)$$

que determina a distribuição espacial dos Z electrons através da densidade de probabilidade $|\langle \vec{r}_1 s_1, \dots, \vec{r}_Z s_Z | \Psi_Z \rangle|^2$. A anti-simetria da função de onda determinantal implica em que essa densidade de probabilidade é *simétrica* nas variáveis eletrônicas. A densidade eletrônica $\rho(\vec{r})$ em um dado ponto \vec{r} do espaço pode então ser obtida como a probabilidade de encontrar um electron em \vec{r} independentemente da posição dos outros electrons que, em vista da simetria da densidade de Z electrons, pode ser calculada como

$$\rho(\vec{r}) = \sum_{s_1 \dots s_Z} \int d^3 r_2 \dots \int d^3 r_Z |\langle \vec{r} s_1, \vec{r}_2 s_2, \dots, \vec{r}_Z s_Z | \Psi_Z \rangle|^2 = \sum_{i=1}^Z |\phi_{k_i}(\vec{r})|^2. \quad (7.18)$$

A redução das somas e integrações múltiplas à forma simples da última expressão envolve apenas a ortonormalidade das diferentes funções de onda de um electron utilizadas na construção do determinante. Ela pode ser feita também utilizando o recurso de substituir um dos determinantes pelo produto não anti-simetrizado das funções de onda de um electron e ajustando as constantes de normalização, como observado na seção 7.1.2. Deve ser observado, porém, que *a mesma* distribuição eletrônica é dada também por uma função de onda escrita simplesmente como o produto *não anti-simetrizado* das mesmas Z funções de onda de um electron utilizadas na construção do determinante. Portanto, o único ingrediente referente ao caráter fermiônico dos electrons que é preservado em $\rho(\vec{r})$ é o fato de que essas Z funções de onda devem ser diferentes, devido às restrições ligadas ao princípio de exclusão. Efeitos de troca, em particular, estão completamente ausentes em $\rho(\vec{r})$.

A densidade eletrônica $\rho(\vec{r})$ implica numa distribuição espacial de carga $\rho_c(\vec{r}) = -e\rho(\vec{r})$ que pode ser usada para determinar um potencial eletrostático $\Phi(\vec{r})$ através da equação de Poisson

$$\nabla^2 \Phi(\vec{r}) = -4\pi [Ze\delta(\vec{r}) + \rho_c(\vec{r})] \quad (7.19)$$

com as condições de contorno

$$\Phi(\vec{r}) \rightarrow \begin{cases} \frac{Ze}{r} & \text{para } r \rightarrow 0 \\ 0 & \text{para } r \rightarrow \infty \end{cases}.$$

A condição de auto-consistência para o potencial $V(\vec{r})$ pode então ser expressa pela relação

$$-e\Phi(\vec{r}) = V(\vec{r}) \quad (7.20)$$

isto é, o potencial eletrostático gerado pela distribuição de cargas resultante de $V(\vec{r})$ reproduz esse mesmo potencial.

A solução da equação de Poisson (7.19) com as condições de contorno requeridas pode ser escrita como

$$\Phi(\vec{r}) = \frac{Ze}{r} + \int d^3r' \frac{\rho_c(\vec{r}')}{|\vec{r} - \vec{r}'|}$$

que, usando a condição (7.20) e exprimindo a densidade de carga $\rho_c(\vec{r})$ em termos da densidade eletrônica (7.18) dá, para o potencial médio auto-consistente de Hartree $V(\vec{r})$

$$V(\vec{r}) = -\frac{Ze^2}{r} + \int d^3r' \frac{e^2}{|\vec{r} - \vec{r}'|} \sum_{i=1}^Z |\phi_{k_i}(\vec{r}')|^2 \quad (7.21)$$

que tem a forma intuitivamente transparente da soma do potencial devido à carga nuclear com um termo adicional dado a média do potencial coulombiano repulsivo entre electrons tomada sobre a densidade eletrônica. Esse potencial determina, por outro lado, os estados de um electron $\phi_k(\vec{r})$ através da “equação de autovalores”

$$\left[\frac{p^2}{2m_e} - \frac{Ze^2}{r} + \int d^3r' \frac{e^2}{|\vec{r} - \vec{r}'|} \sum_{i=1}^Z |\phi_{k_i}(\vec{r}')|^2 \right] \phi_k(\vec{r}) = \epsilon_k \phi_k(\vec{r}) \quad (7.22)$$

que define na realidade um sistema de equações não lineares acopladas devido à presença dos orbitais $\phi_{k_i}(\vec{r})$ na contribuição eletrônica para o potencial médio de Hartree. Essas *equações de Hartree* são em geral tratadas apenas numericamente e através de procedimentos iterativos. A partir de uma escolha inicial para $V(\vec{r})$ (para o que é possível utilizar, por exemplo, o

potencial de Thomas-Fermi) é determinada uma versão inicial para os orbitais, com a qual um novo potencial é gerado usando a relação (7.21). Caso esse potencial seja diferente do utilizado inicialmente, é possível obter com ele novos orbitais e repetir o ciclo até obter auto-consistência, no sentido que o potencial calculado a partir dos orbitais através de (7.21) reproduza o potencial utilizado para gera-los.

As funções de onda de um electron que resolvem auto-consistentemente as equações (7.22) definem ao mesmo tempo um conjunto de “energias de Hartree” de um elétron ϵ_k que podem ser escritas também como

$$\epsilon_k = \int d^3r \phi_k^*(\vec{r}) \left(\frac{p^2}{2m_e} - \frac{Ze^2}{r} \right) \phi_k(\vec{r}) + \int d^3r \int d^3r' \sum_{i=1}^Z |\phi_k(\vec{r})|^2 \frac{e^2}{|\vec{r} - \vec{r}'|} |\phi_{k_i}(\vec{r}')|^2.$$

A soma das energias de Hartree correspondentes aos estados ocupados pode portanto ser escrita

$$\sum_{j=1}^Z \epsilon_{k_j} = \sum_{j=1}^Z \langle \phi_{k_j} | \frac{p^2}{2m_e} - \frac{Ze^2}{r} | \phi_{k_j} \rangle + \sum_{i,j=1}^Z \langle \phi_{k_i} \phi_{k_j} | \frac{e^2}{|\vec{r} - \vec{r}'|} | \phi_{k_i} \phi_{k_j} \rangle. \quad (7.23)$$

Os termos de um corpo dessa relação (correspondentes à energia cinética e ao potencial coulombiano do núcleo), como pode ser verificado pelo cálculo explícito, podem ser escritos como o valor médio de

$$H_0 = \sum_{i=1}^Z \left(\frac{p_i^2}{2m_e} - \frac{Ze^2}{r_i} \right)$$

no estado determinantal (7.17) ou, equivalentemente, no estado produto, não anti-simetrizado

$$\langle \vec{r}_1 s_1, \dots, \vec{r}_Z s_Z | \Psi_Z^{(\text{prod})} \rangle = \prod_{i=1}^Z \phi_{k_i}(\vec{r}_i) \chi_{m_i}(s_i),$$

isto é,

$$\langle \Psi_Z | H_0 | \Psi_Z \rangle = \langle \Psi_Z^{(\text{prod})} | H_0 | \Psi_Z^{(\text{prod})} \rangle = \sum_{j=1}^Z \langle \phi_{k_j} | \frac{p^2}{2m_e} - \frac{Ze^2}{r} | \phi_{k_j} \rangle. \quad (7.24)$$

O termo de dois corpos de (7.23) (correspondente à interação coulombiana entre elétrons), por outro lado, pode ser escrito como

$$\langle \Psi_Z^{(\text{prod})} | \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} | \Psi_Z^{(\text{prod})} \rangle = \int d^3r \int d^3r' \sum_{i,j=1}^Z |\phi_{k_i}(\vec{r})|^2 \frac{e^2}{|\vec{r} - \vec{r}'|} |\phi_{k_j}(\vec{r}')|^2 \quad (7.25)$$

e difere por termos de troca (cf. equações (7.6) e (7.7)) da média correspondente no estado determinantal $|\Psi_Z\rangle$, que é dada por

$$\begin{aligned} \langle \Psi_Z | \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} | \Psi_Z \rangle &= \int d^3 r \int d^3 r' \sum_{i,j=1}^Z |\phi_{k_i}(\vec{r})|^2 \frac{e^2}{|\vec{r} - \vec{r}'|} |\phi_{k_j}(\vec{r}')|^2 - \\ &- \int d^3 r \int d^3 r' \sum_{i,j=1}^Z \phi_{k_i}^*(\vec{r}) \phi_{k_j}^*(\vec{r}') \frac{e^2}{|\vec{r} - \vec{r}'|} \phi_{k_j}(\vec{r}) \phi_{k_i}(\vec{r}') \delta_{m_i m_j}. \end{aligned} \quad (7.26)$$

É importante notar que, como observado anteriormente em conexão com a equação (7.7), o termo de troca inclui uma dependência de spins, que aparece aqui através do delta de Kronecker das funções de onda de spin.

Usando esses resultados, a soma das energias de Hartree de uma partícula (7.23) pode ser escrita em termos de um valor médio no estado produto $|\Psi_Z^{\text{prod}}\rangle$ como

$$\begin{aligned} \sum_{j=1}^Z \epsilon_{k_j} &= \langle \Psi_Z^{\text{prod}} | \left(H_0 + \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \right) | \Psi_Z^{\text{prod}} \rangle \\ &= \langle \Psi_Z^{\text{prod}} | H | \Psi_Z^{\text{prod}} \rangle + \frac{1}{2} \langle \Psi_Z^{\text{prod}} | \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} | \Psi_Z^{\text{prod}} \rangle \end{aligned}$$

onde H é o Hamiltoniano atômico escrito em (7.12). Isso mostra que a soma das energias de Hartree correspondentes aos estados de um elétron incluídos na função de onda produto difere do valor médio do Hamiltoniano nessa mesma função produto por uma dupla contagem dos termos provenientes da interação elétron-elétron.

Por outro lado, não é difícil verificar que os orbitais que satisfazem às equações de Hartree (7.22) dão um extremo do funcional de energia

$$\begin{aligned} \langle \Psi_Z^{\text{prod}} | H | \Psi_Z^{\text{prod}} \rangle &= \sum_{j=1}^Z \langle \phi_{k_j} | \frac{p^2}{2m_e} - \frac{Ze^2}{r} | \phi_{k_j} \rangle \\ &+ \frac{1}{2} \int d^3 r \int d^3 r' \sum_{i,j=1}^Z |\phi_{k_i}(\vec{r})|^2 \frac{e^2}{|\vec{r} - \vec{r}'|} |\phi_{k_j}(\vec{r}')|^2. \end{aligned}$$

De fato, variando $\phi_{k_i}^*(\vec{r})$ (ou, equivalentemente, os bras $\langle \phi_{k_i} |$), e introduzindo multiplicadores de Lagrange ϵ_{k_i} para ter em conta a condição de normalização $\langle \phi_{k_i} | \phi_{k_i} \rangle = 1$, da condição de extremo variacional

$$\delta \left(\langle \Psi_Z^{\text{prod}} | H | \Psi_Z^{\text{prod}} \rangle - \sum_i \epsilon_{k_i} \langle \phi_{k_i} | \phi_{k_i} \rangle \right) = 0$$

resultam as equações de Hartree (7.22). O cancelamento do fator $1/2$ do termo de dois electrons que leva à contagem dupla da energia de interação elétron-elétron na soma das energias de Hartree se deve à existência de duas funções $\phi_{k_i}^*(\vec{r})$ nos elementos de matriz de dois electrons. A energia do átomo na aproximação de Hartree E_H deve portanto ser identificada com o valor desse funcional quando calculado com as soluções das equações de Hartree, e sua relação com as energias associadas aos orbitais de um elétron é dada pelas expressões equivalentes

$$E_H = \sum_i \epsilon_{k_i} - \frac{1}{2} \langle \Psi_Z^{\text{prod}} | \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} | \Psi_Z^{\text{prod}} \rangle = \frac{1}{2} \left(\langle \Psi_Z^{(\text{prod})} | H_0 | \Psi_Z^{(\text{prod})} \rangle + \sum_i \epsilon_{k_i} \right).$$

Esse procedimento variacional para obter as equações de Hartree sugere imediatamente uma forma de incluir na aproximação autoconsistente os efeitos de troca associados à interação elétron-elétron. Ela consiste simplesmente em substituir o funcional de energia calculado com a função de onda produto pelo que se obtém utilizando a função de onda determinantal $|\Psi_Z\rangle$. A equação (7.24) mostra que os termos de um elétron não são modificados por essa substituição, enquanto que a expressão (7.25) deve ser substituída pela (7.26) no caso dos termos que descrevem a interação elétron-elétron. Esta última expressão inclui os efeitos de troca, e a condição variacional (com os devidos multiplicadores de Lagrange para assegurar a normalização dos orbitais de um elétron)

$$\delta \left(\langle \Psi_Z | H | \Psi_Z \rangle - \sum_i \epsilon_{k_i} \langle \phi_{k_i} | \phi_{k_i} \rangle \right) = 0 \quad (7.27)$$

leva às *equações de Hartree-Fock*

$$\begin{aligned} \left[\frac{p^2}{2m_e} - \frac{Ze^2}{r} \right] \phi_k(\vec{r}) + \left(\int d^3 r' \frac{e^2}{|\vec{r} - \vec{r}'|} \sum_{i=1}^Z |\phi_{k_i}(\vec{r}')|^2 \right) \phi_k(\vec{r}) \\ - \int d^3 r' \left(\sum_{i=1}^Z \phi_{k_i}^*(\vec{r}') \frac{e^2}{|\vec{r} - \vec{r}'|} \phi_{k_i}(\vec{r}) \delta_{m_i m} \right) \phi_k(\vec{r}') = \epsilon_k \phi_k(\vec{r}). \end{aligned} \quad (7.28)$$

para os orbitais de um elétron. Essas equações coincidem com as equações de Hartree (7.22) a menos do último termo do lado esquerdo, proveniente das contribuições de troca. Esse termo traz uma dependência do spin através do delta de Kronecker das funções de onda de

spin. Segundo uma interpretação usual dessas equações, elas identificam um potencial de um corpo efetivo para os orbitais de um elétron que inclui, além do potencial coulombiano nuclear $-Ze^2/r$, um potencial médio proveniente da repulsão coulombiana entre elétrons que contém um termo direto idêntico ao que aparece nas equações de Hartree

$$V_d(\vec{r}) = \int d^3r' \frac{e^2}{|\vec{r} - \vec{r}'|} \sum_{i=1}^Z |\phi_{k_i}(\vec{r}')|^2 \quad (7.29)$$

e um termo de troca *não local e dependente de spin*

$$V_t(\vec{r}, \vec{r}', s') = \sum_{i=1}^Z \phi_{k_i}(\vec{r}) \frac{e^2}{|\vec{r} - \vec{r}'|} \phi_{k_i}^*(\vec{r}') \chi_{m_i}^*(s') \quad (7.30)$$

cujas ação sobre um estado $\phi_k(\vec{r})\chi_m(s)$ é dada por

$$[V_t\phi_k](\vec{r}) = \sum_{s'} \int d^3r' V_t(\vec{r}, \vec{r}', s') \phi_k(\vec{r}') \chi_m(s')$$

A soma sobre a variável de spin nesta última expressão reproduz o delta de Kronecker das funções de onda de spin que aparece em (7.28).

A dependência de spin que aparece nas equações de Hartree-Fock (7.28) através dos efeitos de troca merece alguns comentários adicionais. Em primeiro lugar, vale a pena enfatizar que essas equações foram obtidas especificamente para o Hamiltoniano (7.12) que é, em particular, independente dos spins. No caso de átomos com um número par de elétrons, e em que o determinante (7.17) é construído de forma que cada orbital seja ocupado por dois electrons com funções de onda de spin diferentes (um tal estado atômico é às vezes chamado “estado com spin saturado”) então o delta de Kronecker do termo de troca apenas exclui a contribuição de um dos elétrons de cada orbital, de modo que a dependência com \vec{r} e \vec{r}' é a mesma para as duas projeções possíveis do spin m . No caso de spin “não saturado”, no entanto, essa dependência será diferente para cada um dos dois valores possíveis de m . Nesse caso os orbitais $\phi_k(r)$ que correspondem à solução autoconsistente (isto é, que resolvem as equações de Hartree-Fock) adquirem uma dependência de m , e estados com valores diferentes de m não serão em geral ortogonais, pois são nesse caso autofunções de Hamiltonianos com potenciais diferentes. Isso porém não prejudica a construção do determinante (7.17), dado que as funções de onda de spin asseguram a ortogonalidade dos estados de um elétron *completos*, isto é, incluindo o fator de spin. Um exemplo concreto que envolve uma situação deste tipo aparece no problema **7.3**.

Embora de uso corrente no contexto da física atômica⁴, estados de um elétron com dependência de spin fatorizada (v. equação (7.17)), que levam à equação (7.28), não

⁴v., e.g., H. P. Kelly, Phys. Rev. **131**, 684 (1963), particularmente a equação (4.3).

correspondem à escolha mais geral que pode ser feita e implica, em princípio, na introdução de um vínculo restritivo que pode ser importante no caso de sistemas de spin não saturado. O tipo mais geral de estado de um elétron corresponde, na realidade, a “spin-orbitais” da forma

$$\begin{pmatrix} \phi_{k+}(\vec{r}) \\ \phi_{k-}(\vec{r}) \end{pmatrix}$$

que não são em geral autovetores de S_3 , ao contrário do que ocorre com os estados empregados em (7.17). A utilização desses “spin-orbitais” mais gerais permite o aparecimento de dependências de spin também mais gerais no termo de troca, o que permite descrever efeitos mais gerais de correlação spin-posição.

A energia do sistema atômico na aproximação de Hartree-Fock é dada por

$$E_{HF} = \langle \Psi_Z | H | \Psi_Z \rangle$$

supondo o estado determinantal normalizado de acordo com $\langle \Psi_Z | \Psi_Z \rangle = 1$. Usando as equações de Hartree-Fock e as expressões (7.24) e (7.26) essa energia pode ser relacionada com as energias de Hartree-Fock ϵ_k através das duas expressões equivalentes

$$E_{HF} = \sum_i \epsilon_{k_i} - \frac{1}{2} \langle \Psi_Z | \sum_{i,j=1}^Z \frac{e^2}{|\vec{r}_i - \vec{r}_j|} | \Psi_Z \rangle = \frac{1}{2} \left(\langle \Psi_Z | H_0 | \Psi_Z \rangle + \sum_i \epsilon_{k_i} \right)$$

que são análogas às que valem para a aproximação de Hartree.

Finalmente, cabe observar que o uso das aproximações de Thomas-Fermi, Hartree e de Hartree-Fock não se restringe ao problema da estrutura atômica. Aproximações de campo médio desse mesmo tipo são de fato intensamente usada para outros sistemas que envolvem muitas partículas idênticas, notadamente estados eletrônicos em sólidos e núcleos atômicos. Métodos análogos são também aplicáveis a sistemas de muitos bósons idênticos, e têm um papel importante no tratamento de condensados de bósons aprisionados em potenciais externos confinantes, a temperaturas muito baixas.

7.3 Métodos de “segunda quantização” para sistemas de muitas partículas idênticas.

O espaço de Hilbert que funciona na mecânica quântica como espaço de fases para um sistema de N bósons ou férmions idênticos pode sempre ser realizado como o setor simétrico ou anti-simétrico do produto de N espaços de uma partícula. Dessa forma, o número de partículas

N entra na formulação como um ingrediente cinemático “a priori” associado ao sistema a ser estudado. Em muitas situações (por exemplo, situações em que podem ocorrer a criação e a aniquilação de partículas, mas também, como será visto a seguir, para simplificar determinados aspectos do tratamento da cinemática quântica de sistemas de muitas partículas cujo número é conservado) é necessário ou pelo menos conveniente utilizar um tipo alternativo de tratamento em termos de “campos quantizados” no qual o número de partículas aparece não dessa forma, mas como uma variável dinâmica, e as propriedades de simetria ou anti-simetria são implementadas algébricamente.

Respectivamente nos casos de bosons e de férmions idênticos de spin s (inteiro no primeiro caso, semi-inteiro no segundo), os campos quantizados relevantes são conjuntos infinitos (de fato, continuamente infinitos) de operadores $\psi_m(\vec{r})$ onde o rótulo m se refere ao estado do spin (e.g. m sendo o autovalor de S_3 em unidades de \hbar , $-s \leq m \leq s$) e \vec{r} se refere a uma posição no espaço (e que não deve ser confundido com a variável dinâmica de posição da partícula usada na formulação usual). Cada um desses conjuntos continuamente infinitos consiste portanto de uma coleção de $2s + 1$ operadores associada a *cada ponto do espaço*. Os operadores $\psi_m(\vec{r})$ são operadores lineares mas *não hermiteanos*, isto é $\psi_m(\vec{r}) \neq \psi_m^\dagger(\vec{r})$, que agem sobre vetores de um espaço de Hilbert, que será chamado *espaço de Fock*, cujos elementos vão ser designados genericamente também por kets $|\rangle$. Esses operadores são caracterizados ainda pelas relações de comutação

$$\begin{aligned} \text{bosons :} \quad [\psi_m(\vec{r}), \psi_{m'}^\dagger(\vec{r}')] &= \delta_{mm'} \delta(\vec{r} - \vec{r}') \\ [\psi_m(\vec{r}), \psi_{m'}(\vec{r}')] &= 0 \end{aligned} \quad (7.31)$$

e

$$\begin{aligned} \text{férmions :} \quad \{\psi_m(\vec{r}), \psi_{m'}^\dagger(\vec{r}')\} &= \delta_{mm'} \delta(\vec{r} - \vec{r}') \\ \{\psi_m(\vec{r}), \psi_{m'}(\vec{r}')\} &= 0 \end{aligned} \quad (7.32)$$

onde o símbolo $\{a, b\}$ indica o *anticomutador* de a e b , $ab + ba$. Afim de evitar questões técnicas ligadas à natureza singular das equações (7.31) e (7.32) implicada pelas funções delta de Dirac, é conveniente expandir a dependência de \vec{r} dos operadores de campo $\psi_m(\vec{r})$ em termos de um conjunto ortonormal e completo de funções $\phi_k(\vec{r})$

$$\int d^3r \phi_k^*(\vec{r}) \phi_{k'}(\vec{r}) = \delta_{kk'}, \quad \sum_k \phi_k^*(\vec{r}) \phi_k(\vec{r}') = \delta(\vec{r} - \vec{r}')$$

escrevendo-os sob a forma

$$\psi_m(\vec{r}) = \sum_k \phi_k(\vec{r}) a_{km}. \quad (7.33)$$

Os coeficientes a_{km} (que na realidade são operadores) podem por sua vez ser expressos em termos dos $\psi_m(\vec{r})$ como

$$a_{km} = \int d^3r \phi_k^*(\vec{r}) \psi_m(\vec{r}) \quad (7.34)$$

e as relações de comutação e anticomutação, (7.31) e (7.32) juntamente com as relações de ortonormalidade e completeza das funções $\phi_k(\vec{r})$ mostram que eles, por sua vez, satisfazem também as relações de comutação e anti-comutação

$$\begin{aligned} \text{bosons :} \quad [a_{km}, a_{k'm'}^\dagger] &= \delta_{kk'} \delta_{mm'} \\ [a_{km}, a_{k'm'}] &= 0. \end{aligned} \quad (7.35)$$

e

$$\begin{aligned} \text{férmions :} \quad \{a_{km}, a_{k'm'}^\dagger\} &= \delta_{kk'} \delta_{mm'} \\ \{a_{km}, a_{k'm'}\} &= 0. \end{aligned} \quad (7.36)$$

Tanto no caso de bosons como no de férmions, uma classe importante de operadores que é possível construir, seja a partir dos operadores de campo $\psi_m(\vec{r})$ seja a partir de suas componentes definidas em termos de uma base ortonormal como na equação (7.34), é a dos chamados *operadores de número*. No primeiro caso os operadores interessantes são

$$N_m = \int d\vec{r} \psi_m^\dagger(\vec{r}) \psi_m(\vec{r}) \quad \text{e} \quad N = \sum_m N_m \quad (7.37)$$

enquanto que, utilizando as componentes definidas em (7.34), eles são

$$N_{km} = a_{km}^\dagger a_{km}, \quad N_m = \sum_k a_{km}^\dagger a_{km} \quad \text{e} \quad N = \sum_m N_m.$$

É claro que as duas definições dos operadores N_m e N coincidem, em vista da relação (7.34). A motivação para o nome dado a esses operadores aparece claramente quando se examina as suas relações de comutação com os operadores de campo ou com os a_{km}^\dagger, a_{km} . Essas relações podem ser obtidas facilmente a partir da definição dos operadores e das relações básicas de comutação (7.31), (7.32) e (7.35), (7.36) do que resulta, *tanto no caso de bosons como no de férmions*

$$\begin{aligned}
[N_m, \psi_{m'}(\vec{r})] &= -\delta_{mm'}\psi_{m'}(\vec{r}); & [N, \psi_{m'}(\vec{r})] &= -\psi_{m'}(\vec{r}); \\
[N_m, \psi_{m'}^\dagger(\vec{r})] &= \delta_{mm'}\psi_{m'}^\dagger(\vec{r}); & [N, \psi_{m'}^\dagger(\vec{r})] &= \psi_{m'}^\dagger(\vec{r})
\end{aligned}
\tag{7.38}$$

e ainda

$$\begin{aligned}
[N_{km}, a_{k'm'}] &= -\delta_{kk'}\delta_{mm'}a_{k'm'}; \\
[N_{km}, a_{k'm'}^\dagger] &= \delta_{kk'}\delta_{mm'}a_{k'm'}^\dagger;
\end{aligned}
\tag{7.39}$$

$$\begin{aligned}
[N_m, a_{km}] &= -\delta_{mm}a_{km}; & [N, a_{km}] &= -a_{km}; \\
[N_m, a_{km}^\dagger] &= \delta_{mm}a_{km}^\dagger; & [N, a_{km}^\dagger] &= a_{km}^\dagger.
\end{aligned}$$

Essas relações têm uma semelhança clara com as que foram utilizadas na seção **3.1.1** para determinar o espectro do oscilador harmônico. Elas podem por isso ser usadas para obter, de uma forma inteiramente análoga, o espectro dos operadores $N_{\nu m}$, N_m e N . Os ingredientes necessários para isso são, de fato, as relações de comutação e propriedades geométricas gerais (como a positividade da norma) dos vetores do espaço de Fock. Uma primeira observação a ser feita é a de que

$$N_{km}^\dagger = N_{km} \quad \text{e} \quad [N_{km}, N_{k'm'}] = 0 \quad (\text{bosons ou férmions}),$$

isto é, tanto N_m como N são escritos como uma soma de operadores hermiteanos que *comutam* entre si (compatíveis) e que portanto podem ser todos diagonalizados simultaneamente. Dessa forma, os autovalores desses dois operadores poderão ser sempre escritos como somas de autovalores dos N_{km} . Para obter estes últimos basta observar, como no caso do oscilador harmônico, que das primeiras duas relações (7.39), e supondo que $|n_{km}\rangle$ seja um autovetor de N_{km} com autovalor n_{km} , segue que

$$N_{km}a_{km}^\dagger|n_{km}\rangle = a_{km}^\dagger N_{km}|n_{km}\rangle + [N_{km}, a_{km}^\dagger]|n_{km}\rangle = (n_{km} + 1)a_{km}^\dagger|n_{km}\rangle \quad \text{se} \quad a_{km}^\dagger|n_{km}\rangle \neq 0$$

e

$$N_{km}a_{km}|n_{km}\rangle = a_{km} N_{km}|n_{km}\rangle + [N_{km}, a_{km}]|n_{km}\rangle = (n_{km} - 1)a_{km}|n_{km}\rangle \quad \text{se} \quad a_{km}|n_{km}\rangle \neq 0.$$

Por outro lado,

$$n_{km} = \langle n_{km} | N_{km} | n_{km} \rangle = \langle n_{km} | a_{km}^\dagger a_{km} | n_{km} \rangle \geq 0$$

pois essa quantidade é a norma do vetor $a_{km} | n_{km} \rangle$. Dessas relações decorre portanto que os autovalores n_{km} devem ser números inteiros não negativos, e que

$$a_{km} | n_{km} = 0 \rangle = 0.$$

Esses resultados são válidos tanto para bosons quanto para férmions. Neste último caso, no entanto, segue das relações de anti-comutação (7.36) que $a_{km}^{\dagger 2} \equiv 0$, de modo que

$$a_{km}^{\dagger 2} | n_{km} = 0 \rangle = a_{km}^\dagger | n_{km} = 1 \rangle = 0$$

de modo que os autovalores de N_{km} são apenas 0 e 1 no caso de férmions. Esse mesmo resultado pode ser obtido notando que, no caso de férmions, os operadores N_{km} são na realidade *operadores de projeção*, que satisfazem portanto as relações (v. problema 7.5)

$$N_{km}^2 = N_{km} = N_{km}^\dagger \quad (\text{férmions}).$$

Como consequência desses fatos e da definição de N_m e de N como soma dos N_{km} , os espectros de N_m e de N consistem dos inteiros 0, 1, 2, Esses operadores são interpretados como associados respectivamente ao *número de partículas no estado de spin m* e ao *número total de partículas*. Isso na realidade torna essa quantidade uma variável dinâmica associada a um operador no espaço de Fock, como mencionado inicialmente. O autovetor de N com autovalor 0, $|0\rangle$, é o autovetor simultâneo de todos os N_{km} com autovalor zero, o que de acordo com a interpretação de N corresponde ao “vácuo”, isto é, ao estado sem nenhuma partícula. Os operadores a_{km}^\dagger e a_{km} , por outro lado, funcionam como *operadores de criação* e de *aniquilação* de partículas. A natureza bosônica ou fermiônica das partículas é determinada pelas relações de comutação ou anti-comutação satisfeitas pelos operadores de campo. Autovetores de N_m e de N com autovalores $\nu_m > 0$, $\nu > 0$ são obtidos atuando sobre $|0\rangle$ com ν operadores de criação. No caso de férmions, esses operadores devem ser todos *diferentes*, já que o produto de dois operadores iguais é nulo devido à propriedade de anticomutatividade dos operadores de campo.

7.3.1 Estados no espaço de Fock e funções de onda de muitas partículas idênticas.

Para completar a interpretação desse esquema é preciso ainda explicitar a natureza dos estados quânticos de muitas partículas que são obtidos aplicando sucessivamente vários operadores de

criação sobre o vácuo. Isso pode ser feito interpretando o operador de campo $\psi_m^\dagger(\vec{r})$ como operador de criação de uma partícula com projeção de spin m no ponto \vec{r} , isto é

$$\psi_s^\dagger(\vec{r})|0\rangle = |\vec{r}s\rangle.$$

Consistentemente com isso, a_{km}^\dagger cria um férmion com spin m no estado cuja função de onda (amplitude de probabilidade) é $\phi_k(\vec{r})$ (cf. equação (7.34)):

$$|km\rangle = a_{km}^\dagger|0\rangle.$$

A consistência dessas interpretações pode ser verificada calculando a função de onda

$$\langle\vec{r}s|km\rangle = \langle 0|\psi_s(\vec{r})a_{km}^\dagger|0\rangle = \phi_k(\vec{r})\delta_{sm} = \phi_k(\vec{r})\chi_m(s) \quad (7.40)$$

o que resulta do uso da equação (7.34) e das relações (7.31) ou (7.32). Explicitamente

$$\langle 0|\psi_s(\vec{r})a_{km}^\dagger|0\rangle = \int d\vec{r}' \langle 0|\psi_s(\vec{r})\psi_m^\dagger(\vec{r}')|0\rangle \phi_k(\vec{r}') = \int d\vec{r}' \delta(\vec{r} - \vec{r}')\delta_{sm}\phi_k(\vec{r}') = \phi_k(\vec{r})\chi_m(s)$$

onde foi usado o resultado $\psi_s(\vec{r})|0\rangle = 0$, que pode ser deduzido de que $|0\rangle$ é o autovetor comum de todos os N_{km} com autovalor zero. Desse modo $a_{ks}|0\rangle = 0$ para quaisquer ks , e também $\psi_s(\vec{r})|0\rangle = 0$ para quaisquer s e \vec{r} em virtude de (7.33). Isso mostra ainda que qualquer operador de aniquilação dá um resultado nulo quando agindo sobre o vácuo.

Essa correspondência entre vetores do espaço de Fock e funções de onda pode ser estendida facilmente para estados com mais de uma partícula. Os vetores do espaço de Fock a considerar nesse caso são vetores do tipo

$$a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger |0\rangle. \quad (7.41)$$

No caso da estatística de Fermi, é fácil verificar que esses vetores têm norma 0 ou 1. De fato, se dois ou mais operadores a_{km}^\dagger são iguais, as propriedades de anti-comutação desses operadores garantem que a norma é nula. Se, no entanto, todos os operadores de criação utilizados são distintos, é possível escrever

$$\begin{aligned} \langle 0|a_{k_1 m_1} \dots a_{k_n m_n} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger |0\rangle &= \\ &= \langle 0|a_{k_1 m_1} \dots a_{k_{n-1} m_{n-1}} \left(1 - a_{k_n m_n}^\dagger a_{k_n m_n}\right) a_{k_{n-1} m_{n-1}}^\dagger \dots a_{k_1 m_1}^\dagger |0\rangle. \end{aligned}$$

O operador de aniquilação $a_{k_n m_n}$ pode porém ser levado neste caso a agir diretamente sobre o vácuo introduzindo apenas uma fase adicional $(-1)^{n-1}$, o que mostra que o termo que contém esse operador é nulo. Disso segue então que

$$\langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = \langle 0 | a_{k_1 m_1} \dots a_{k_{n-1} m_{n-1}} a_{k_{n-1} m_{n-1}}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle$$

que corresponde à norma de um vetor do mesmo tipo que o vetor de partida mas agora com $n - 1$ operadores de criação. O mesmo procedimento pode portanto ser repetido outras $n - 1$ vezes levando finalmente a

$$\langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = \langle 0 | 0 \rangle = 1 \quad (\text{férmions}).$$

No caso de bosons, a questão da normalização dos estados do tipo (7.41) é ligeiramente mais complicada devido à possibilidade de criar mais de um boson no mesmo estado de uma partícula. Procedendo como no caso de férmions, mas tendo em conta as relações de comutação bosônicas, a contribuição do termo envolvendo $a_{k_n m_n}^\dagger a_{k_n m_n}$ não é em geral nula, mas igual a $n_{k_n m_n} - 1$ onde $n_{k_n m_n}$ é o número de operadores de criação iguais a $a_{k_n m_n}^\dagger$ no estado inicialmente considerado. Um procedimento recursivo semelhante ao utilizado no caso de férmions leva então a

$$\langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = \prod_q (\nu_q!) \langle 0 | 0 \rangle = \prod_q (\nu_q!) \quad (\text{bosons})$$

onde os ν_q são as multiplicidades da ocorrência de cada tipo distinto de operador de criação no estado considerado, com $\sum_q \nu_q = n$. Os vetores normalizados são portanto

$$\left[\prod_q (\nu_q!) \right]^{-\frac{1}{2}} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle.$$

Vale a pena observar que esse resultado *se aplica também ao caso de férmions*, dado que todos os ν_q devem ser então iguais a 1.

Funções de onda correspondentes a tais estados do espaço de Fock podem ser obtidas através de uma extensão natural da relação (7.40). O objeto a ser considerado pode ser escrito, tanto no caso de bosons como no de férmions, como

$$\tilde{\Psi}_{k_1 m_1 \dots k_n m_n}(\vec{r}_1 s_1, \dots, \vec{r}_n s_n) = \langle 0 | \psi_{s_1}(\vec{r}_1) \dots \psi_{s_n}(\vec{r}_n) a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle \left[\prod_q (\nu_q!) \right]^{-\frac{1}{2}}$$

que corresponde ao produto escalar de um estado *normalizado* do tipo (7.41) com

$$\psi_{s_n}(\vec{r}_n)^\dagger \dots \psi_{s_1}(\vec{r}_1)^\dagger | 0 \rangle.$$

Em primeiro lugar, é preciso verificar as propriedades de normalização dessas funções de onda de muitas partículas. Isso é feito calculando

$$\begin{aligned}
\langle \tilde{\Psi}_{\{k_i m_i\}} | \tilde{\Psi}_{\{k_i m_i\}} \rangle &\equiv \sum_{s_i} \int d^3 r_1 \dots \int d^3 r_n \left| \tilde{\Psi}_{k_1 m_1 \dots k_n m_n}(\vec{r}_1 s_1, \dots, \vec{r}_n s_n) \right|^2 = \\
&= \left[\prod_q (\nu_q!) \right]^{-1} \sum_{s_i} \int d^3 r_1 \dots \int d^3 r_n \\
&\langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} \psi_{s_n}^\dagger(\vec{r}_n) \dots \psi_{s_1}^\dagger(\vec{r}_1) | 0 \rangle \langle 0 | \psi_{s_1}(\vec{r}_1) \dots \psi_{s_n}(\vec{r}_n) a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle \\
&= \left[\prod_q (\nu_q!) \right]^{-1} \sum_{s_i} \int d^3 r_1 \dots \int d^3 r_n \\
&\langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} \psi_{s_n}^\dagger(\vec{r}_n) \dots \psi_{s_1}^\dagger(\vec{r}_1) \psi_{s_1}(\vec{r}_1) \dots \psi_{s_n}(\vec{r}_n) a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle.
\end{aligned}$$

A eliminação da projeção sobre o vácuo $|0\rangle\langle 0|$ no integrando é possível porque os n operadores de criação que agem sobre o vácuo do lado direito são seguidos de n operadores de aniquilação (sob a forma dos operadores de campo $\psi_{s_i}(\vec{r}_i)$). Desse modo, a projeção sobre o vácuo pode ser substituída por uma resolução da unidade do espaço de Fock escrita como

$$\hat{1}_{\text{Fock}} = \sum_{\{n_{k_i m_i}\}} |\{n_{k_i m_i}\}\rangle \langle \{n_{k_i m_i}\}|$$

pois a única contribuição provém, de qualquer modo, do termo de vácuo, no qual todos os $n_{k_i m_i}$ são nulos. Usando em seguida a definição (7.37) do operador número N e as relações de comutação (7.38), o valor esperado no vácuo que aparece no integrando pode ser expresso como

$$\begin{aligned}
\langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} N(N-1) \dots 1 a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle &= n! \langle 0 | a_{k_1 m_1} \dots a_{k_n m_n} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = \\
&= n! \prod_q (\nu_q!).
\end{aligned}$$

Isso mostra finalmente que a forma *normalizada* das funções de onda de n partículas deve ser escrita, tanto no caso de bosons como no de férmions, como

$$\Psi_{k_1 m_1 \dots k_n m_n}(\vec{r}_1 s_1, \dots, \vec{r}_n s_n) = \left[n! \prod_q (\nu_q!) \right]^{-\frac{1}{2}} \langle 0 | \psi_{s_1}(\vec{r}_1) \dots \psi_{s_n}(\vec{r}_n) a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle. \quad (7.42)$$

No caso de férmions as multiplicidades ν_q são todas iguais a 1, enquanto no caso de bosons a única restrição existente é a de que a soma dos ν_q deve ser igual ao número total de partículas n .

Resta, finalmente, examinar a natureza das funções de onda normalizadas. Para isso é conveniente examinar antes um caso particular explícito, que pode ser em seguida ser estendido ao caso geral sem maiores dificuldades. A função de onda normalizada de *duas* partículas (bosons ou férmions)

$$\begin{aligned} & \frac{1}{\sqrt{2\nu!}} \langle 0 | \psi_{s_1}(\vec{r}_1) \psi_{s_2}(\vec{r}_2) a_{k_2 m_2}^\dagger a_{k_1 m_1}^\dagger | 0 \rangle = \\ & = \frac{1}{\sqrt{2\nu!}} \int d^3 r_1' \int d^3 r_2' \langle 0 | \psi_{s_1}(\vec{r}_1) \psi_{s_2}(\vec{r}_2) \psi_{m_2}^\dagger(\vec{r}_2') \psi_{m_1}^\dagger(\vec{r}_1') | 0 \rangle \phi_{k_2}(\vec{r}_2') \phi_{k_1}(\vec{r}_1') \end{aligned} \quad (7.43)$$

(onde $\nu = 2$ para de dois bosons com $k_1 m_1 = k_2 m_2$ e $\nu = 1$ em todos os outros casos) pode ser calculada, como foi feito para a função de uma partícula (7.40), usando as regras de comutação (ou anti-comutação) para levar os operadores de aniquilação a agirem sobre o vácuo, com resultado nulo. As contribuições não nulas provém portanto das funções delta dos comutadores ou anti-comutadores, que podem ser integradas imediatamente. De forma mais explícita, e considerando o caso de férmions para fixar idéias, o produto de quatro operadores de campo pode ser re-escrito como

$$\psi_{s_1}(\vec{r}_1) \psi_{s_2}(\vec{r}_2) \psi_{m_2}^\dagger(\vec{r}_2') \psi_{m_1}^\dagger(\vec{r}_1') = \quad (7.44)$$

$$= \psi_{s_1}(\vec{r}_1) \{ \psi_{s_2}(\vec{r}_2) \psi_{m_2}^\dagger(\vec{r}_2') \} \psi_{m_1}^\dagger(\vec{r}_1') - \psi_{s_1}(\vec{r}_1) \psi_{m_2}^\dagger(\vec{r}_2') \psi_{s_2}(\vec{r}_2) \psi_{m_1}^\dagger(\vec{r}_1') \quad (7.45)$$

$$= \psi_{s_1}(\vec{r}_1) \delta(\vec{r}_2 - \vec{r}_2') \delta_{s_2 m_2} \psi_{m_1}^\dagger(\vec{r}_1') - \psi_{s_1}(\vec{r}_1) \psi_{m_2}^\dagger(\vec{r}_2') \psi_{s_2}(\vec{r}_2) \psi_{m_1}^\dagger(\vec{r}_1'); \quad (7.46)$$

o primeiro termo da última linha pode ser ainda transformado de acordo com

$$\begin{aligned} & \delta(\vec{r}_2 - \vec{r}_2') \delta_{s_2 m_2} \psi_{s_1}(\vec{r}_1) \psi_{m_1}^\dagger(\vec{r}_1') = \\ & = \delta(\vec{r}_2 - \vec{r}_2') \delta_{s_2 m_2} \delta(\vec{r}_1 - \vec{r}_1') \delta_{s_1 m_1} - \delta(\vec{r}_2 - \vec{r}_2') \delta_{s_2 m_2} \psi_{m_1}^\dagger(\vec{r}_1') \psi_{s_1}(\vec{r}_1) \end{aligned}$$

sendo que a contribuição deste último termo é nula, pois ele contém um operador de aniquilação à direita, que vai agir diretamente sobre o vácuo. O último termo de (7.44) pode ser tratado da mesma forma, dando uma contribuição

$$-\delta(\vec{r}_1 - \vec{r}_2')\delta_{s_1 m_2}\delta(\vec{r}_2 - \vec{r}_1')\delta_{s_2 m_1}$$

mais termos nulos, por envolver operadores de aniquilação agindo sobre o vácuo. Levando os termos não nulos (produtos de funções delta) à equação (7.43) resulta finalmente

$$\frac{1}{\sqrt{2}}\langle 0|\psi_{s_1}(\vec{r}_1)\psi_{s_2}(\vec{r}_2)a_{k_2 m_2}^\dagger a_{k_1 m_1}^\dagger|0\rangle = \frac{1}{\sqrt{2}} \det \begin{pmatrix} \phi_{k_1}(\vec{r}_1)\chi_{m_1}(s_1) & \phi_{k_1}(\vec{r}_2)\chi_{m_1}(s_2) \\ \phi_{k_2}(\vec{r}_1)\chi_{m_2}(s_1) & \phi_{k_2}(\vec{r}_2)\chi_{m_2}(s_2) \end{pmatrix}.$$

Esse resultado mostra explicitamente que a expressão (7.43) é (no caso considerado, de férmions) a função de onda *antissimetrizada* de duas partículas dada como o determinante de Slater formado pelas funções de onda de uma partícula ϕ_{k_1} e ϕ_{k_2} e com os estados de spin m_1 e m_2 . De uma forma completamente análoga é possível verificar, no caso de dois bosons, que a função de onda é uma função *simetrizada* das duas partículas, descrita em termos de um permanente.

No caso geral de estados de n partículas o que se obtém para as funções de onda normalizadas é

$$\begin{aligned} \Psi_{k_1 m_1 \dots k_n m_n}(\vec{r}_1 s_1, \dots, \vec{r}_n s_n) &= \frac{1}{\sqrt{n!}} \det \begin{pmatrix} \phi_{k_1}(\vec{r}_1)\chi_{m_1}(s_1) & \dots & \phi_{k_1}(\vec{r}_n)\chi_{m_1}(s_n) \\ \dots & \dots & \dots \\ \phi_{k_n}(\vec{r}_1)\chi_{m_n}(s_1) & \dots & \phi_{k_n}(\vec{r}_n)\chi_{m_n}(s_n) \end{pmatrix} \\ &\equiv \frac{1}{\sqrt{n!}} \sum_P (-1)^P P \phi_{k_1}(\vec{r}_1)\chi_{m_1}(s_1) \dots \phi_{k_n}(\vec{r}_n)\chi_{m_n}(s_n) \end{aligned}$$

para o caso de férmions e

$$\begin{aligned} \Psi_{k_1 m_1 \dots k_n m_n}(\vec{r}_1 s_1, \dots, \vec{r}_n s_n) &= \frac{1}{\sqrt{n! \prod_q \nu_q!}} \text{perm} \begin{pmatrix} \phi_{k_1}(\vec{r}_1)\chi_{m_1}(s_1) & \dots & \phi_{k_1}(\vec{r}_n)\chi_{m_1}(s_n) \\ \dots & \dots & \dots \\ \phi_{k_n}(\vec{r}_1)\chi_{m_n}(s_1) & \dots & \phi_{k_n}(\vec{r}_n)\chi_{m_n}(s_n) \end{pmatrix} \\ &\equiv \frac{1}{\sqrt{n! \prod_q \nu_q!}} \sum_P P \phi_{k_1}(\vec{r}_1)\chi_{m_1}(s_1) \dots \phi_{k_n}(\vec{r}_n)\chi_{m_n}(s_n) \end{aligned}$$

para o caso de bosons.

Na seção **7.1.1** foram construídas bases, respectivamente no setor simétrico e anti-simétrico do espaço de estados de sistemas de muitas partículas, sob a forma de permanentes e de determinantes de estados de uma partícula. Os resultados acima permitem então associar esses estados de base a vetores dos espaços de Fock de bosons e de férmions de acordo com a prescrição

$$\sqrt{\frac{n!}{\prod_g \nu_g!}} \mathcal{S}|k_1 m_1, \dots, k_n m_n\rangle \leftrightarrow \frac{1}{\sqrt{\prod_g \nu_g!}} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger |0\rangle \quad (\text{bosons})$$

$$\sqrt{n!} \mathcal{A}|k_1 m_1, \dots, k_n m_n\rangle \leftrightarrow a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger |0\rangle \quad (\text{férmions}). \quad (7.47)$$

Um estado qualquer de muitos bosons ou férmions idênticos pode portanto ser representado em termos de um vetor do espaço de Fock escrito como uma combinação linear de vetores do tipo (7.47).

7.3.2 Variáveis dinâmicas no espaço de Fock.

Para completar a transição do espaço produto para o espaço de Fock é preciso ainda obter a representação das variáveis dinâmicas também em termos de operadores definidos nesse espaço. A condição que define a representação no espaço de Fock \hat{g} de uma variável dinâmica g definida no espaço produto é a de que os elementos de matriz calculados no espaço de Fock devem ser iguais aos elementos de matriz correspondentes calculados no espaço produto. Como estados do tipo (7.47) são bases em termos das quais é possível escrever estados gerais, basta verificar a equivalência para elementos de matriz envolvendo estados desse tipo, isto é

$$\frac{1}{\prod_g \nu_g!} \langle 0 | a_{k'_1 m'_1} \dots a_{k'_n m'_n} \hat{g} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle \equiv \langle \Psi_{\{k'_i m'_i\}} | g | \Psi_{\{k_i m_i\}} \rangle.$$

Para determinar a forma dos operadores \hat{g} no espaço de Fock é conveniente classificar as variáveis dinâmicas de sistemas de muitas partículas idênticas em operadores de um corpo, de dois corpos, etc. Como já foi discutido, a identidade das partículas exige que esses operadores sejam sempre *simétricos* em todas as partículas. Um operador típico de um corpo é a energia cinética

$$K = \sum_{i=1}^n \frac{p_i^2}{2M} = \sum_{i=1}^n K_i \quad (7.48)$$

que satisfaz obviamente a essa condição de simetria. O que caracteriza um operador de um corpo é o fato de ele ser escrito como uma soma de termos cada um dos quais se refere a apenas uma partícula. Um operador de dois corpos é o que se escreve como uma soma de termos, cada um dos quais se refere a duas partículas, e essa classificação se estende de maneira análoga a um número qualquer de corpos. Um exemplo típico de um operador de dois corpos é o que descreve um potencial agindo entre pares de partículas

$$V = \frac{1}{2} \sum_{i,j=1}^n v(i, j) \quad (7.49)$$

onde os argumentos i, j de v indicam as variáveis apropriadas das duas partículas envolvidas (posição, spin, etc.), $v(i, j) = v(j, i)$ e o fator $1/2$ evita contagem dupla, dado que ambos os índices são somados de 1 a n . A forma típica de um operador de três corpos é

$$W = \frac{1}{3!} \sum_{i,j,k=1}^n w(i, j, k)$$

onde $w(i, j, k)$ deve ser simétrico sob permutações de seus argumentos e o fator $1/3!$ evita contagens múltiplas. O número n de partículas que constituem o sistema aparece em cada uma dessas expressões através do limite superior para os índices das somas.

Uma propriedade geral importante de elementos de matriz de operadores de p corpos entre estados da base correspondente a n ($n \geq p$) partículas (escritos como determinantes ou permanentes de n estados de uma partícula) é a de que eles se anulam sempre que os dois determinantes ou permanentes envolvidos diferirem por mais que p estados de uma partícula. Isso significa, em outras palavras, que a ação de um operador de p corpos sobre um estado dado como um determinante ou permanente pode alterar o estado de *no máximo* p partículas, cabendo às restantes o papel de simples “espectadoras”. No caso de operadores de um corpo, tomando como exemplo a energia cinética (7.48), um elemento de matriz típico tem a forma

$$\begin{aligned} \langle \Psi' | K | \Psi \rangle &= \sum_{P P'} (\pm 1)^P (\pm 1)^{P'} \prod_{j=1}^n \left(\int d^3 r_j \sum_{s_j} \right) \frac{1}{n! \sqrt{\prod \nu_q! \prod \nu'_q!}} \times \\ &\times P' \phi_{k'_1}^*(\vec{r}_1) \chi_{m'_1}^*(s_1) \dots \phi_{k'_n}^*(\vec{r}_n) \chi_{m'_n}^*(s_n) \sum_i K_i P \phi_{k_1}(\vec{r}_1) \chi_{m_1}(s_1) \dots \phi_{k_n}(\vec{r}_n) \chi_{m_n}(s_n) \end{aligned} \quad (7.50)$$

em que as fases $(\pm 1)^P$ correspondem respectivamente ao caso de permanentes e de determinantes. Cada termo desta dupla soma sobre permutações P e P' contém um único operador K_i , de modo que $n - 1$ das n integrações (e somas sobre variáveis de spin) correspondem na realidade a simples produtos escalares de estados de uma partícula, e que portanto se anulam sempre que os estados envolvidos forem diferentes, devido às suas propriedades de ortonormalidade. Eles correspondem nesse caso às partículas espectadoras, que são portanto em número de $n - 1$ e devem necessariamente ocupar os mesmos estados de partícula independente nos dois determinantes ou permanentes. A integração restante leva ao elemento de matriz *de uma partícula*

$$\int d^3 r_i \sum_{s_i} \phi_{k'_i}^*(\vec{r}_i) \chi_{m'_i}^*(s_i) K_i \phi_{k_i}(\vec{r}_i) \chi_{m_i}(s_i) \equiv \langle k'_i m'_i | \frac{p^2}{2M} | k_i m_i \rangle$$

que será em geral diferente de zero *mesmo que* $\phi_{k'_i} \chi_{m'_i} \neq \phi_{k_i} \chi_{m_i}$. Desse modo, no caso em que $|\Psi\rangle$ difere de $|\Psi'\rangle$ por um estado de uma partícula, $\{k_i m_i\} \rightarrow \{k'_i m'_i\}$, o elemento de matriz (7.50) é dado por

$$\langle \Psi' | K | \Psi \rangle = \sqrt{\nu_i \nu'_i} \langle k'_i m'_i | \frac{p^2}{2M} | k_i m_i \rangle$$

onde ν_i e ν'_i correspondem ao número de vezes que os estados $k_i m_i$ e $k'_i m'_i$ ocorrem respectivamente em Ψ e em Ψ' (no caso de férmions esses dois números são necessariamente iguais a 1). Isso se deve a que o operador de um corpo deve agir necessariamente entre esses dois estados. No caso em que $|\Psi\rangle = |\Psi'\rangle$, por outro lado,

$$\langle \Psi | K | \Psi \rangle = \sum_i \nu_i \langle k_i m_i | \frac{p^2}{2M} | k_i m_i \rangle$$

onde os pesos ν_i correspondem ao número de vezes que o estado $k_i m_i$ ocorre em Ψ . Uma característica importante desses resultados é a de que os elementos de matriz de um operador de um corpo entre determinantes ou permanentes é dado *sempre* em termos de elementos de matriz *de uma partícula*.

O caso de operadores de dois ou mais corpos pode ser tratado de forma completamente análoga. Para um operador de p corpos haverá em cada termo da soma sobre permutações $n-p$ partículas espectadoras que deverão necessariamente ocupar os mesmos estados de partícula independente nos determinantes ou permanentes inicial e final para que sua contribuição seja diferente de zero, e sempre que isso acontece o resultado final se exprime em termos de elementos de matriz de p partículas do tipo

$$\begin{aligned} & \langle k_1 m_1 \dots k_p m_p | w | k'_1 m'_1 \dots k'_p m'_p \rangle = \\ & = \prod_{i=1}^p \left(\int d^3 r_i \sum_{s_i} \right) \phi_{k_1}^*(\vec{r}_1) \chi_{m_1}^*(s_1) \dots \phi_{k_p}^*(\vec{r}_p) \chi_{m_p}^*(s_p) w \phi_{k'_1}(\vec{r}_1) \chi_{m'_1}(s_1) \dots \phi_{k'_p}(\vec{r}_p) \chi_{m'_p}(s_p). \end{aligned}$$

Uma propriedade característica de elementos de matriz de operadores dois ou mais corpos entre estados de muitas partículas idênticas que não existe no caso de elementos de matriz de operadores de um corpo é o aparecimento de efeitos de troca. Para entender o mecanismo que dá origem a esses efeitos é suficiente considerar um elemento de matriz de um operador de

dois corpos como (7.49) entre dois determinantes $|\Psi\rangle$ e $|\Psi'\rangle$ que diferem em *dois* estados de uma partícula, $\{k_i m_i, k_j m_j\} \rightarrow \{k'_i m'_i, k'_j m'_j\}$. Nesse caso há $n - 2$ partículas espectadoras que devem ocupar os mesmos estados de uma partícula para que o elemento de matriz não se anule. As duas funções de onda de uma partícula que sofrem a ação do operador de dois corpos dão origem, no entanto, contribuem através de diferentes permutações P e P' levando a dois tipos distintos de contribuição que diferem por uma permutação de variáveis. Explicitamente

$$\begin{aligned} \langle \Psi' | V | \Psi \rangle &= \\ &= \int d^3 r_1 \int d^3 r_2 \sum_{S_1 S_2} \phi_{k'_i}^*(\vec{r}_1) \chi_{m'_i}(s_1) \phi_{k'_j}^*(\vec{r}_2) \chi_{m'_j}(s_2) v(1, 2) \phi_{k_i}^*(\vec{r}_1) \chi_{m_i}(s_1) \phi_{k_j}^*(\vec{r}_2) \chi_{m_j}(s_2) \pm \\ &\pm \int d^3 r_1 \int d^3 r_2 \sum_{S_1 S_2} \phi_{k'_i}^*(\vec{r}_1) \chi_{m'_i}(s_1) \phi_{k'_j}^*(\vec{r}_2) \chi_{m'_j}(s_2) v(1, 2) \phi_{k_i}^*(\vec{r}_2) \chi_{m_i}(s_2) \phi_{k_j}^*(\vec{r}_1) \chi_{m_j}(s_1) \end{aligned}$$

onde o duplo sinal corresponde respectivamente aos casos de bosons e de férmions. Esses dois termos provêm na realidade da contribuição de quatro combinações distintas de permutações P e P' que dão resultados iguais duas a duas, levando ao cancelamento do fator $1/2$ que aparece em (7.49).

Os operadores do espaço de Fock que levam a esses mesmos resultados quando calculados com os estados (7.47) são dados por

$$K \rightarrow \hat{K} = \sum_{k_1 m_1 k_2 m_2} \langle k_2 m_2 | \frac{p^2}{2M} | k_1 m_1 \rangle a_{k_2 m_2}^\dagger a_{k_1 m_1} \quad (7.51)$$

para um operador de um corpo como (7.48) e

$$V \rightarrow \hat{V} = \frac{1}{2} \sum_{k_1 m_1 k_2 m_2 k_3 m_3 k_4 m_4} \langle k_3 m_3 k_4 m_4 | v | k_1 m_1 k_2 m_2 \rangle a_{k_3 m_3}^\dagger a_{k_4 m_4}^\dagger a_{k_2 m_2} a_{k_1 m_1} \quad (7.52)$$

para um operador de dois corpos como (7.49). Os elementos de matriz que aparecem nessas expressões são os elementos de matriz de uma e duas partículas das variáveis dinâmicas correspondentes que aparecem nos resultados calculados em termos de determinantes ou permanentes, isto é

$$\langle k_2 m_2 | \frac{p^2}{2M} | k_1 m_1 \rangle = \int d^3 r \sum_s \phi_{k_2}^*(\vec{r}) \chi_{m_2}^*(s) \frac{p^2}{2M} \phi_{k_1}(\vec{r}) \chi_{m_1}(s)$$

e

$$\begin{aligned}
& \langle k_3 m_3 k_4 m_4 | v | k_1 m_1 k_2 m_2 \rangle = \\
& = \int d^3 r_1 \int d^3 r_2 \sum_{s_1 s_2} \phi_{k_3}^*(\vec{r}_1) \chi_{m_3}^*(s_1) \phi_{k_4}^*(\vec{r}_2) \chi_{m_4}^*(s_2) v(1, 2) \phi_{k_1}(\vec{r}_1) \chi_{m_1}(s_1) \phi_{k_2}(\vec{r}_2) \chi_{m_2}(s_2).
\end{aligned}$$

Não é difícil verificar explicitamente a adequação dessas prescrições. Tomando como exemplo o elemento de matriz de operador de um corpo (7.50), o que deve ser calculado no espaço de Fock é

$$\langle \Psi' | K | \Psi \rangle \rightarrow \frac{1}{\sqrt{\prod \nu_q! \prod \nu'_q!}} \langle 0 | a_{k'_1 m'_1} \dots a_{k'_n m'_n} \hat{K} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle$$

com \hat{K} dado por (7.51), ou seja

$$\langle \Psi' | K | \Psi \rangle \rightarrow \frac{1}{\sqrt{\prod \nu_q! \prod \nu'_q!}} \sum_{k m, k' m'} \langle k' m' | \frac{p^2}{2M} | k m \rangle \langle 0 | a_{k'_1 m'_1} \dots a_{k'_n m'_n} a_{k' m}^\dagger a_{k m} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle. \quad (7.53)$$

Essa expressão tem a forma de uma combinação linear de elementos de matriz de uma partícula com pesos determinados pelo último fator, dado como valor esperado no vácuo de um produto de operadores de criação e de aniquilação. O cálculo desse objeto é feito, em última análise, utilizando as relações de comutação desses operadores, juntamente com o fato de que a ação de um operador de aniquilação sobre o vácuo dá um resultado nulo.

O processo, frequentemente longo, envolvido num procedimento desse tipo, pode ser abreviado em casos particulares através do uso de outras propriedades, como as propriedades de ortogonalidade e normalização de estados do tipo (7.47). Como ilustração do uso dessas outras propriedades, o último fator que aparece em (7.53) pode ser inicialmente re-escrito, usando as relações de comutação, como

$$\mp \langle 0 | a_{k'_1 m'_1} \dots a_{k'_n m'_n} a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle \delta_{k k'} \delta_{m m'} \pm \langle 0 | a_{k'_1 m'_1} \dots a_{k'_n m'_n} a_{k m} a_{k' m'}^\dagger a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle$$

onde os sinais superiores e inferiores se referem respectivamente aos casos de bosons e de férmions. O primeiro dos dois termos assim obtidos é, a menos apenas de fatores de normalização, o produto escalar $\langle \Psi' | \Psi \rangle$ que se anula se o conjunto dos $\{k_i m_i\}$ for diferente do conjunto dos $\{k'_i m'_i\}$. No caso em que esses dois conjuntos forem iguais seu valor é dado pela norma $\prod \nu_q!$. O segundo termo, por outro lado, corresponde ao produto escalar de dois vetores do tipo (7.47) *no setor de $n + 1$ partículas*. A condição para que ele não se anule é que o conjunto

dos operadores de criação seja (a menos de permutações) igual ao conjunto dos operadores de aniquilação. Essa condição só pode ser satisfeita em dois casos: a) o conjunto dos $\{k'_i m'_i\}$ é igual ao conjunto dos $\{k_i m_i\}$, e nesse caso deve-se ter também $k = k'$ e $m = m'$, ou b) esses dois conjuntos diferem *apenas* por um estado de uma partícula, $k_j m_j \neq k'_j m'_j$, e então $k = k'_j$, $m = m'_j$ e $k' = k_j$, $m' = m_j$. O valor das contribuições do segundo termo depende ainda da particular estatística que esteja sendo considerada. Para a estatística de Bose

$$\text{caso a)} \quad \langle 0 | a_{k_1 m_1} \dots a_{k_n m'_n} a_{k m} a_{k m}^\dagger a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = (\nu + 1) \prod_q \nu_q !$$

$$\text{caso b)} \quad \langle 0 | a_{k_1 m_1} \dots a_{k'_j m'_j} \dots a_{k_n m'_n} a_{k_j m_j} a_{k'_j m'_j}^\dagger a_{k_n m_n}^\dagger \dots a_{k_j m_j}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = \prod_q \sqrt{\nu_q ! \nu'_q !} \sqrt{\nu_j \nu'_j}$$

onde, no caso a), ν é a multiplicidade de km no conjunto $\{k_i m_i\}$ e, no caso b), ν_j e ν'_j são respectivamente as multiplicidades de ocupação dos estados $k_j m_j$ e $k'_j m'_j$ em $\{k_i m_i\}$ e $\{k'_i m'_i\}$. Para a estatística de Fermi, por outro lado, todas as multiplicidades de ocupação são limitadas a 1 e

$$\text{caso a)} \quad \langle 0 | a_{k_1 m_1} \dots a_{k_n m'_n} a_{k m} a_{k m}^\dagger a_{k_n m_n}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = \begin{cases} 1 & \text{se } \nu_j = 0 \\ 0 & \text{se } \nu_j = 1 \end{cases}$$

$$\text{caso b)} \quad \langle 0 | a_{k_1 m_1} \dots a_{k'_j m'_j} \dots a_{k_n m'_n} a_{k_j m_j} a_{k'_j m'_j}^\dagger a_{k_n m_n}^\dagger \dots a_{k_j m_j}^\dagger \dots a_{k_1 m_1}^\dagger | 0 \rangle = -1$$

onde o sinal negativo no último caso provém da diferença de ordenamento dos operadores de criação e aniquilação. Levando todos esses resultados à relação (7.53) o que se obtém é

$$\langle \Psi' | K | \Psi \rangle = \begin{cases} \text{Bose :} & \begin{aligned} a) & \sum_i \nu_i \langle k_i m_i | \frac{p^2}{2M} | k_i m_i \rangle \\ b) & \sqrt{\nu_j \nu'_j} \langle k'_j m'_j | \frac{p^2}{2M} | k_j m_j \rangle \end{aligned} \\ \text{Fermi :} & \begin{aligned} a) & \sum_i \nu_i \langle k_i m_i | \frac{p^2}{2M} | k_i m_i \rangle \\ b) & \langle k'_j m'_j | \frac{p^2}{2M} | k_j m_j \rangle \end{aligned} \end{cases}$$

o que reproduz os resultados obtidos em termos de permanentes e determinantes.

Um cálculo análogo para o operador (7.52) mostra a adequação dessa prescrição para a representação de operadores de dois corpos no espaço de Fock. A extensão da prescrição para operadores de tres ou mais corpos também pode ser obtida sem dificuldade. No caso de operadores de tres corpos, por exemplo, a prescrição correspondente a (7.52) é

$$W \rightarrow \hat{W} = \frac{1}{3!} \sum_{k_i, m_i} \langle k_4 m_4 k_5 m_5 k_6 m_6 | w | k_1 m_1 k_2 m_2 k_3 m_3 \rangle a_{k_4 m_4}^\dagger a_{k_5 m_5}^\dagger a_{k_6 m_6}^\dagger a_{k_3 m_3} a_{k_2 m_2} a_{k_1 m_1}.$$

Finalmente, utilizando a relação (7.34) é possível exprimir as variáveis dinâmicas no espaço de Fock em termos dos operadores de campo $\psi_m(\vec{r})$. A transformação da expressão (7.51) para a energia cinética dá

$$\begin{aligned} \hat{K} &= \int d^3 r \int d^3 r' \sum_{k_1 m_1 k_2 m_2} \phi_{k_1}(\vec{r}) \langle k_1 m_1 | \frac{p^2}{2M} | k_2 m_2 \rangle \phi_{k_2}^*(\vec{r}') \psi_{m_1}^\dagger(\vec{r}) \psi_{m_2}(\vec{r}') \\ &= \sum_{m_1} \int d^3 r \int d^3 r' \psi_{m_1}^\dagger(\vec{r}) \langle \vec{r} | \frac{p^2}{2M} | \vec{r}' \rangle \psi_{m_1}(\vec{r}') \\ &= \sum_{m_1} \int d^3 r \psi_{m_1}^\dagger(\vec{r}) \frac{-\hbar^2 \nabla^2}{2M} \psi_{m_1}(\vec{r}) \end{aligned}$$

onde foi usada a completeza das funções $\phi_k(\vec{r})$ e a independência de spin de $p^2/2M$ para escrever

$$\sum_{k_1 k_2} \phi_{k_1}(\vec{r}) \langle k_1 m_1 | \frac{p^2}{2M} | k_2 m_2 \rangle \phi_{k_2}^*(\vec{r}') = \langle \vec{r} m_1 | \frac{p^2}{2M} | \vec{r}' m_2 \rangle = \langle \vec{r} | \frac{p^2}{2M} | \vec{r}' \rangle \delta_{m_1 m_2}.$$

A eliminação da integração sobre \vec{r}' pode ser justificada tendo em vista que o elemento de matriz de $p^2/2M$ entre os vetores impróprios $|\vec{r}\rangle$ e $|\vec{r}'\rangle$ pode ser escrito sob a forma

$$\langle \vec{r} | \frac{p^2}{2M} | \vec{r}' \rangle = -\frac{\hbar^2}{2M} \delta''(\vec{r} - \vec{r}')$$

onde $\delta''(\vec{r} - \vec{r}')$ representa a segunda derivada da função delta de Dirac. Uma dupla integração por partes com relação a \vec{r}' leva então ao resultado indicado.

Um operador de dois corpos independente de spin e local se escreve, de forma análoga, como

$$\hat{V} = \sum_{m_1 m_2} \int d^3 r_1 \int d^3 r_2 v(\vec{r}_1, \vec{r}_2) \psi_{m_1}^\dagger(\vec{r}_1) \psi_{m_2}^\dagger(\vec{r}_2) \psi_{m_2}(\vec{r}_2) \psi_{m_1}(\vec{r}_1).$$

Uma propriedade importante (e facilmente verificável) de operadores como \hat{K} , \hat{V} e \hat{W} acima é que eles comutam com o operador número de férmions N , isto é $[\hat{K}, N] = [\hat{V}, N] = [\hat{W}, N] = 0$. Portanto um Hamiltoniano de muitos corpos que contém energia cinética e interações de dois corpos descritas por um potencial, $H = K + V$, também comuta com N , $[H, N] = 0$. Esta

última relação significa que para tais sistemas *o número de partículas é uma constante do movimento*, isto é, uma quantidade conservada.

Como um balanço final dos resultados desta seção e da seção anterior, é interessante notar que, enquanto o espaço produto de N espaços de uma partícula que funciona como espaço de fases para um sistema de N partículas contém sub-espaços associados a diferentes estatísticas (sub-espaço simétrico para bosons e anti-simétrico para férmions) para o mesmo número de partículas, o espaço de Fock contém sub-espaços associados a *diferentes números de partículas* para uma dada estatística. A diferentes números de partículas estão associados diferentes espaços produto, e a diferentes estatísticas estão associados, para um dado número de partículas, diferentes sub-espaços do mesmo espaço produto. Inversamente, a diferentes estatísticas estão associados diferentes espaços de Fock, e a diferentes números de partículas com uma dada estatística estão associados diferentes sub-espaços de um mesmo espaço de Fock.

Na medida em que *todas* as variáveis dinâmicas de um dado sistema comutem com N haverá, no espaço de Fock, uma regra de superseleção análoga à associada à simetria no caso de partículas idênticas, e na medida em que uma tal regra de superseleção exista, o tratamento em termos do espaço de Fock é exatamente equivalente à descrição em termos de espaços produto. No espaço de Fock, no entanto, há a possibilidade de que se considere operadores, não traduzíveis em termos de uma descrição envolvendo espaços produto, que não conservam o número de partículas. Um exemplo de operador desse tipo será encontrado no capítulo 8 com relação à interação da matéria com o campo de radiação eletromagnética. Nesse caso há operadores associados à emissão e à absorção de fótons pelas fontes do campo eletromagnético que implicam na não conservação do número de fótons no sistema.

7.3.3 Exemplos de aplicação: gás de Fermi livre e com forças de dois corpos.

1. Gás de Fermi livre. As propriedades, já discutidas anteriormente (v. página 366), do estado de menor energia de um gás de Fermi livre podem ser obtidas também através do uso das técnicas de segunda quantização, como uma primeira ilustração do uso dessas técnicas. Para esse fim, é conveniente construir operadores de criação e de aniquilação fermiônicos associados à base de autofunções do momento com condições de contorno periódicas no volume L^3 . Neste caso, a equação (7.34) assume a forma particular

$$a_{\vec{k},m} = \int_{L^3} d^3r \frac{e^{-i\vec{k}\cdot\vec{r}}}{\sqrt{L^3}} \psi_m(\vec{r})$$

onde agora o índice \vec{k} se refere ao autovalor do momento, e os operadores $a_{\vec{k},m}$ e $a_{\vec{k},m}^\dagger$, que respectivamente aniquilam e criam um férmion com momento $\hbar\vec{k}$ e com o estado de spin carac-

terizado por m , satisfazem as relações de anti-comutação (7.36). O hamiltoniano do sistema de férmions livres é um operador de um corpo cuja versão segundo-quantizada é

$$H = \sum_i \frac{p_i^2}{2M} \longrightarrow \sum_{\vec{k}, m} \sum_{\vec{k}', m'} \langle \vec{k}', m' | \frac{p^2}{2M} | \vec{k}, m \rangle a_{\vec{k}', m'}^\dagger a_{\vec{k}, m}.$$

Como os estados envolvidos no elemento de matriz são autoestados do momento \vec{p} , é imediato que

$$\langle \vec{k}', m' | \frac{p^2}{2M} | \vec{k}, m \rangle = \frac{\hbar^2 k^2}{2M} \delta_{\vec{k}', \vec{k}} \delta_{m', m}$$

de modo que a versão segundo-quantizada do hamiltoniano livre se reduz simplesmente a

$$H = \sum_{\vec{k}, m} \frac{\hbar^2 k^2}{2M} a_{\vec{k}, m}^\dagger a_{\vec{k}, m}.$$

Como dito anteriormente, os operadores $N_{\vec{k}m} = a_{\vec{k}, m}^\dagger a_{\vec{k}, m}$ que intervêm na soma correspondem ao número de férmions no estado $|\vec{k}, m\rangle$, de modo que esse hamiltoniano pode ser visto como uma soma das energias cinéticas de cada um dos estados da base ponderada com o número de férmions presentes nesse estado. O número total de férmions presentes no sistema corresponde, por outro lado, ao operador

$$N = \sum_{\vec{k}, m} N_{\vec{k}m} = \sum_{\vec{k}, m} a_{\vec{k}, m}^\dagger a_{\vec{k}, m}.$$

Vale a pena notar explicitamente que esse operador, bem como cada um dos operadores $N_{\vec{k}m}$, comutam com o hamiltoniano e também comutam entre si, de modo que o número total de férmions, o número de férmions em cada um dos estados de momento e de spin e a energia do gás de Fermi livre são observáveis compatíveis. Desse modo, os autovetores de H podem ser escolhidos de forma a serem autovetores simultâneos de cada um dos $N_{\vec{k}m}$. Estes são, no entanto operadores de projeção (isto é, $N_{\vec{k}m} = N_{\vec{k}m}^\dagger = N_{\vec{k}m}^2$), de modo que seus autovalores são 1 ou 0, correspondendo respectivamente à presença ou à ausência de um férmion no estado $|\vec{k}, m\rangle$. Com essas idéias como guia, é fácil ver que os autovetores simultâneos de H e de cada um dos $N_{\vec{k}m}$ são dados como um produto de operadores de criação agindo sobre o vácuo de Fock, isto é, são dados por

$$|\vec{k}_1 m_1, \dots, \vec{k}_n m_n\rangle = a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger |0\rangle. \quad (7.54)$$

Das propriedades de anti-comutação dos operadores de criação fermiônicos resulta imediatamente que todos os $a_{\vec{k}_i m_i}^\dagger$ que aparecem nesse produto devem ser distintos para que o estado resultante não tenha norma nula. Essa condição corresponde evidentemente às restrições impostas pelo princípio de exclusão de Pauli, implementadas no formalismo de segunda quantização através das propriedades algébricas dos operadores de criação e de aniquilação. As funções de onda correspondentes a esses estados de muitos férmions são na realidade determinantes de Slater de dimensão n construídos a partir das autofunções do momento (com condições de contorno periódicas) e dos autoestados de s_3 e devidamente normalizados. A ambigüidade de sinal envolvida na possibilidade de diferentes ordenamentos das linhas e colunas desses determinantes corresponde à ambigüidade de sinal associada a diferentes ordenamentos dos operadores de criação envolvidos no produto.

Os autovalores dos operadores N e H associados ao estado (7.54) podem ser imediatamente identificados como sendo n e

$$E = \sum_{\vec{k}, m} \frac{\hbar^2 k^2}{2M} n_{\vec{k}m} \quad (7.55)$$

onde

$$n_{\vec{k}m} = \begin{cases} 1 & \text{se } a_{\vec{k}m}^\dagger \text{ aparece em (7.54)} \\ 0 & \text{se } a_{\vec{k}m}^\dagger \text{ não aparece em (7.54)} \end{cases}.$$

Desse modo, o estado fundamental do sistema de n férmions corresponde ao estado da forma (7.54) em que os operadores de criação envolvem os menores valores possíveis de \vec{k} , tendo em conta a condição de que eles devem ser todos distintos, seja no valor de \vec{k} , seja no valor de m . O maior valor de $|\vec{k}|$ que intervém no produto de operadores de criação, nesse caso, é o momento de Fermi k_F . As relações obtidas anteriormente da densidade $\rho = n/L^3$ e da energia do estado fundamental por partícula E_0/n com o momento de Fermi resultam agora das expressões

$$\rho = \frac{1}{L^3} \sum_{k \leq k_F, m} 1 \quad \text{e} \quad \frac{E_0}{n} = \frac{1}{n} \sum_{k \leq k_F, m} \frac{\hbar^2 k^2}{2M}$$

que podem ser reduzidas transformando as somas em integrais da forma usual, com resultado idêntico ao obtido anteriormente.

Uma outra forma de obter a expressão (7.55) para a energia do sistema de férmions livres consiste em calcular o valor esperado

$$\begin{aligned}
E &= \langle \vec{k}_1 m_1, \dots, \vec{k}_n m_n | H | \vec{k}_1 m_1, \dots, \vec{k}_n m_n \rangle \\
&= \langle 0 | a_{\vec{k}_1 m_1} \dots a_{\vec{k}_n m_n} H a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger | 0 \rangle \\
&= \sum_{\vec{k}, m} \frac{\hbar^2 k^2}{2M} \langle 0 | a_{\vec{k}_1 m_1} \dots a_{\vec{k}_n m_n} a_{\vec{k} m}^\dagger a_{\vec{k} m} a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger | 0 \rangle
\end{aligned}$$

que tem a forma característica de uma soma de elementos de matriz ponderados com valores esperados de um produto de operadores de criação e de aniquilação no vácuo. Estes pesos podem ser calculados utilizando apenas as propriedades algébricas desses operadores, em particular a de que a ação direta sobre o vácuo de um operador de aniquilação qualquer dá um vetor de norma nula, juntamente com as relações de comutação. Uma forma de abreviar o cálculo neste caso específico consiste em repetir aqui o procedimento adotado anteriormente com relação à equação (7.53). De fato, das relações de anticomutação dos operadores de criação e de aniquilação resulta

$$a_{\vec{k} m}^\dagger a_{\vec{k} m} = 1 - a_{\vec{k} m} a_{\vec{k} m}^\dagger$$

de modo que

$$\begin{aligned}
&\langle 0 | a_{\vec{k}_1 m_1} \dots a_{\vec{k}_n m_n} a_{\vec{k} m}^\dagger a_{\vec{k} m} a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger | 0 \rangle = \\
&= \langle 0 | a_{\vec{k}_1 m_1} \dots a_{\vec{k}_n m_n} a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger | 0 \rangle - \langle 0 | a_{\vec{k}_1 m_1} \dots a_{\vec{k}_n m_n} a_{\vec{k} m}^\dagger a_{\vec{k} m} a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger | 0 \rangle = \\
&= 1 - \left(1 - \sum_{i=1}^n \delta_{\vec{k} m, \vec{k}_i m_i} \right).
\end{aligned}$$

De fato, o primeiro dos dois termos em que o peso fica assim decomposto é simplesmente a norma do estado considerado, enquanto o segundo termo é a norma de um estado antissimétrico que resulta da criação, sobre esse estado, de um férmion adicional no estado de uma partícula $|\vec{k}, m\rangle$. Se esse estado de uma partícula já não for ocupado no estado considerado, essa norma é também 1, mas ela será nula sempre que esse estado de uma partícula já estiver ocupado. Desse modo os pesos a serem utilizados no cálculo da energia serão iguais a 1 quando, no estado considerado, houver um férmion com números quânticos \vec{k}, m (pois nesse caso o parênteses contendo a soma é nulo) e zero quando não houver no estado considerado um férmion com esses números quânticos (caso em que o parênteses é igual a 1), reproduzindo portanto o resultado (7.55).

2. Gás de Fermi com forças de dois corpos. Um sistema de muitos férmions idênticos que interagem entre si através de forças que podem ser representadas por um potencial de dois

corpos, é na realidade um sistema (muito) mais complicado que o gás de Fermi livre tratado no item anterior, e tem a sua dinâmica caracterizada por um hamiltoniano da forma

$$H = \sum_i \frac{p_i^2}{2M} + \frac{1}{2} \sum_{i,j} v(i,j) \longrightarrow \sum_{\vec{k},m} \frac{\hbar^2 k^2}{2M} a_{\vec{k},m}^\dagger a_{\vec{k},m} + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} \langle \alpha\beta | v | \gamma\delta \rangle a_{\vec{k}_\alpha m_\alpha}^\dagger a_{\vec{k}_\beta m_\beta}^\dagger a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\gamma m_\gamma}$$

onde os operadores de criação e aniquilação são os mesmos que foram definidos anteriormente em termos das autofunções do momento com condições periódicas de contorno para os diferentes autovalores de s_3 , e a notação abreviada usada para o elemento de matriz do potencial de dois corpos corresponde mais explicitamente a

$$\langle \alpha\beta | v | \gamma\delta \rangle \equiv \langle \vec{k}_\alpha m_\alpha, \vec{k}_\beta m_\beta | v(1,2) | \vec{k}_\gamma m_\gamma, \vec{k}_\delta m_\delta \rangle$$

sendo que os estados rotulados por α e γ correspondem à partícula 1 enquanto os estados rotulados por β e δ correspondem à partícula 2. O problema de obter autovetores desse hamiltoniano (e mesmo o de obter apenas o seu estado fundamental, em casos nos quais ele exista) não pode ser resolvido, de modo que o tratamento desse tipo de sistema envolve tipicamente o uso de métodos perturbativos ou variacionais para obter aproximações reveladoras de propriedades tidas como relevantes.

O cálculo da correção perturbativa de ordem mais baixa da energia do estado fundamental do gás livre, devida à interação de dois corpos, pode servir como exemplo do uso de técnicas de segunda quantização. Essa correção é dada simplesmente como o valor médio do termo de interação de H no estado de menor energia do gás *livre*, obtido no item anterior. Desse modo

$$\Delta E_0^{(1)} = \langle E_0 | \frac{1}{2} \sum_{\alpha\beta\gamma\delta} \langle \alpha\beta | v | \gamma\delta \rangle a_{\vec{k}_\alpha m_\alpha}^\dagger a_{\vec{k}_\beta m_\beta}^\dagger a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\gamma m_\gamma} | E_0 \rangle \quad (7.56)$$

onde

$$|E_0\rangle = a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger |0\rangle$$

e os índices \vec{k}_i, m_i correspondem aos n estados de uma partícula com os menores valores possíveis do momento, estando satisfeita a condição de que eles sejam todos distintos. Substituindo essa expressão para o estado fundamental do gás livre em (7.56), a correção perturbativa aparece novamente sob a forma padrão de uma soma de elementos de matriz do potencial de dois corpos ponderada com valores esperados no vácuo de produtos de operadores de criação e de aniquilação. De fato,

$$\Delta E_0^{(1)} = \frac{1}{2} \sum_{\alpha\beta\gamma\delta} \langle \alpha\beta | v | \gamma\delta \rangle \langle 0 | a_{\vec{k}_1 m_1} \dots a_{\vec{k}_n m_n} a_{\vec{k}_\alpha m_\alpha}^\dagger a_{\vec{k}_\beta m_\beta}^\dagger a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\gamma m_\gamma} a_{\vec{k}_n m_n}^\dagger \dots a_{\vec{k}_1 m_1}^\dagger | 0 \rangle. \quad (7.57)$$

Uma forma de calcular os pesos a partir das propriedades dos operadores de criação e de aniquilação (cf. Problema 7.7) consiste em usar mais uma vez um procedimento análogo ao adotado com relação à equação (7.53). Os operadores de criação e de aniquilação provenientes do termo de dois corpos de H podem ser re-escritos como

$$\begin{aligned} a_{\vec{k}_\alpha m_\alpha}^\dagger a_{\vec{k}_\beta m_\beta}^\dagger a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\gamma m_\gamma} &= a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\gamma m_\gamma} a_{\vec{k}_\alpha m_\alpha}^\dagger a_{\vec{k}_\beta m_\beta}^\dagger \\ &- \delta_{\beta\delta} a_{\vec{k}_\gamma m_\gamma} a_{\vec{k}_\alpha m_\alpha}^\dagger - \delta_{\alpha\gamma} a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\beta m_\beta}^\dagger + \delta_{\alpha\delta} a_{\vec{k}_\gamma m_\gamma} a_{\vec{k}_\beta m_\beta}^\dagger + \delta_{\beta\gamma} a_{\vec{k}_\delta m_\delta} a_{\vec{k}_\alpha m_\alpha}^\dagger \\ &+ \delta_{\alpha\gamma} \delta_{\beta\delta} - \delta_{\alpha\delta} \delta_{\beta\gamma} \end{aligned}$$

onde os índices dos deltas de Kronecker foram escritos usando a mesma notação abreviada usada para os elementos de matriz de dois corpos. O termo com quatro operadores dá contribuição não nula quando

$$\begin{array}{ccc} \alpha = \gamma & & \alpha = \delta \\ \beta = \delta & \text{ou} & \beta = \gamma \end{array}$$

com $\alpha \neq \beta$ e com sinal negativo no segundo caso, devido à inversão na ordem de um par de operadores. Esses operadores devem corresponder ainda a estados de uma partícula não ocupados no estado fundamental não perturbado. De forma semelhante, os dois operadores dos quatro termos seguintes devem ser também iguais, correspondendo a um estado de uma partícula não ocupado no estado fundamental. Juntando esses resultados a correção de primeira ordem para a energia adquire a forma

$$\begin{aligned} \Delta E_0^{(1)} &= \frac{1}{2} \sum_{\alpha\beta}^\infty [\langle \alpha\beta | v | \alpha\beta \rangle - \langle \alpha\beta | v | \beta\alpha \rangle] - \\ &- \sum_{\alpha,\beta>n}^\infty [\langle \alpha\beta | v | \alpha\beta \rangle - \langle \alpha\beta | v | \beta\alpha \rangle] + \frac{1}{2} \sum_{\alpha,\beta>n}^\infty [\langle \alpha\beta | v | \alpha\beta \rangle - \langle \alpha\beta | v | \beta\alpha \rangle] = \\ &= \frac{1}{2} \sum_{\alpha\beta}^n [\langle \alpha\beta | v | \alpha\beta \rangle - \langle \alpha\beta | v | \beta\alpha \rangle] \end{aligned} \quad (7.58)$$

As restrições escritas nas somas como $\alpha, \beta > n$ e $\alpha, \beta < n$ se referem respectivamente a estados de uma partícula não ocupados e ocupados no estado fundamental não perturbado.

Uma forma alternativa de obter esse mesmo resultado é notar que, para que o valor esperado no vácuo do produto de operadores de criação e de aniquilação, *tal como escrito inicialmente em (7.57)*, seja diferente de zero é preciso que

$$\begin{array}{ccc} \alpha = \gamma < n & & \alpha = \delta < n \\ \beta = \delta < n & \text{ou} & \beta = \gamma < n \end{array}$$

com $\alpha \neq \beta$, e com sinal negativo no segundo caso, devido à inversão na ordem de um par de operadores.

O passo seguinte consiste em explicitar os elementos de matriz de dois corpos que aparecem na expressão (7.58). O primeiro deles é dado simplesmente por

$$\langle \alpha \beta | v | \alpha \beta \rangle = \frac{1}{L^6} \int_{L^3} d^3 r_1 \int_{L^3} d^3 r_2 e^{-i(\vec{k}_\alpha \cdot \vec{r}_1 + \vec{k}_\beta \cdot \vec{r}_2)} \langle m_\alpha m_\beta | v(1, 2) | m_\alpha m_\beta \rangle e^{i(\vec{k}_\alpha \cdot \vec{r}_1 + \vec{k}_\beta \cdot \vec{r}_2)}$$

onde os bra's e ket's que aparecem no integrando se referem aos estados de spin. O segundo elemento de matriz contém as contribuições *de troca*

$$\langle \alpha \beta | v | \beta \alpha \rangle = \frac{1}{L^6} \int_{L^3} d^3 r_1 \int_{L^3} d^3 r_2 e^{-i(\vec{k}_\alpha \cdot \vec{r}_1 + \vec{k}_\beta \cdot \vec{r}_2)} \langle m_\alpha m_\beta | v(1, 2) | m_\beta m_\alpha \rangle e^{i(\vec{k}_\beta \cdot \vec{r}_1 + \vec{k}_\alpha \cdot \vec{r}_2)}.$$

O cálculo dessas expressões depende evidentemente da forma específica do potencial de dois corpos $v(1, 2)$. No caso mais simples em que esse potencial é um potencial central e independente dos spins, da forma

$$v(1, 2) \longrightarrow v(|\vec{r}_1 - \vec{r}_2|) \otimes \hat{1}_S$$

esses elementos de matriz poder ser calculados re-escrevendo as ondas planas em termos das coordenadas relativa $\vec{r} = \vec{r}_1 - \vec{r}_2$ e do centro de massa, $\vec{R} = (\vec{r}_1 + \vec{r}_2)/2$ do par de partículas e notando que

$$\int_{L^3} d^3 r_1 \int_{L^3} d^3 r_2 = \int_{L^3} d^3 R \int_{L^3} d^3 r.$$

O elemento de matriz direto se reduz simplesmente a⁵

$$\langle \alpha \beta | v | \beta \alpha \rangle = \frac{1}{L^3} \int_{L^3} d^3 r v(r)$$

enquanto o termo de troca aparece como

$$\langle \alpha \beta | v | \beta \alpha \rangle = \delta_{m_\alpha m_\beta} \frac{1}{L^3} \int_{L^3} d^3 r e^{-i(\vec{k}_\alpha - \vec{k}_\beta) \cdot \vec{r}} v(r)$$

⁵Essa expressão mostra que, para que o resultado do cálculo perturbativo de primeira ordem seja pelo menos *finito*, é preciso exigir que o potencial central de dois corpos satisfaça à condição $\int d^3 r v(r) < \infty$.

de modo que a correção perturbativa de primeira ordem para a energia do estado fundamental, devida ao potencial de dois corpos, é dada por

$$\Delta E_0^{(1)} = \frac{1}{2L^3} \sum_{\alpha\beta}^n \left(\int_{L^3} d^3r v(r) - \delta_{m_\alpha m_\beta} \int_{L^3} d^3r e^{-i(\vec{k}_\alpha - \vec{k}_\beta) \cdot \vec{r}} v(r) \right). \quad (7.59)$$

Essa expressão depende da densidade do gás de Fermi através das somas, que envolvem todos os valores do momento até o momento de Fermi. Além disso, embora o potencial tenha sido tomado como dependente de $|\vec{r}_1 - \vec{r}_2|$ apenas, o termo de troca introduz uma dependência com os spins, bem como uma dependência com o momento relativo $\hbar(\vec{k}_\alpha - \vec{k}_\beta)$ do par de partículas envolvido na interação. O termo direto pode ser calculado imediatamente substituindo da forma usual as somas por integrais sobre os momentos, o que leva a

$$\begin{aligned} \frac{1}{2L^3} \sum_{\alpha\beta}^n \int_{L^3} d^3r v(r) &\rightarrow 4 \frac{L^3}{2(2\pi)^6} \int_{k_\alpha \leq k_F} d^3k_\alpha \int_{k_\beta \leq k_F} d^3k_\beta \int d^3r v(r) \\ &= 4 \frac{L^3}{2(2\pi)^6} \left(\frac{4\pi}{3} k_F^3 \right)^2 \int d^3r v(r) = \frac{N}{2} \rho \int d^3r v(r) \end{aligned}$$

onde o fator 4 provém das somas sobre os estados de spin (supondo que os férmions tenham spin 1/2),

$$N = 2 \frac{L^3}{(2\pi)^3} \left(\frac{4\pi}{3} k_F^3 \right)$$

é o número total de férmions no volume L^3 e $\rho = N/L^3$ é a densidade do gás de Fermi. O termo de troca, por outro lado, pode ser calculado como

$$\frac{1}{2L^3} \sum_{\alpha\beta}^n \delta_{m_\alpha m_\beta} \int_{L^3} d^3r e^{-i(\vec{k}_\alpha - \vec{k}_\beta) \cdot \vec{r}} v(r) \rightarrow \frac{L^3}{(2\pi)^6} \int_{k_\alpha \leq k_F} d^3k_\alpha \int_{k_\beta \leq k_F} d^3k_\beta \int d^3r e^{-i(\vec{k}_\alpha - \vec{k}_\beta) \cdot \vec{r}} v(r).$$

É conveniente fazer antes as integrações sobre os momentos usando os resultados

$$\int_{k \leq k_F} d^3k e^{i\vec{k} \cdot \vec{r}} = 4\pi \int_0^{k_F} dk k^2 j_0(kr) = \frac{4\pi}{3} k_F^3 \frac{j_1(k_F r)}{k_F r}$$

onde j_0 e j_1 são funções de Bessel esféricas (v. pag. 256 e Problema 7.10), com os quais o termo de troca fica reduzido à forma

$$\frac{N}{4} \rho \int d^3r \left(\frac{j_1(k_F r)}{k_F r} \right)^2 v(r).$$

A correção perturbativa de primeira ordem *por partícula* para a energia do estado fundamental do gás de Fermi é portanto, no caso de um potencial de dois corpos da forma $v(|\vec{r}_1 - \vec{r}_2|)$,

$$\frac{\Delta E_0^{(1)}}{N} = \frac{\rho}{2} \left[\int d^3r v(r) - \frac{1}{2} \int d^3r \left(\frac{j_1(k_F r)}{k_F r} \right)^2 v(r) \right].$$

O valor médio da energia *total* por partícula (isto é, incluindo a energia cinética) no estado fundamental não perturbado do gás de Fermi livre com densidade dada através do momento de Fermi k_F pode ser escrito agora como

$$\frac{\langle E \rangle}{N} = \frac{K}{N} + \frac{\Delta E_0^{(1)}}{N} = \frac{3}{5} \frac{\hbar^2 k_F^2}{2M} + \frac{\rho}{2} \left[\int d^3r v(r) - \frac{1}{2} \int d^3r \left(\frac{j_1(k_F r)}{k_F r} \right)^2 v(r) \right].$$

Essa expressão envolve, na realidade, o valor médio do hamiltoniano do gás de Fermi com interações de dois corpos no estado fundamental do gás de Fermi *livre* com a densidade correspondente ao valor utilizado de k_F , e pode ser também vista, portanto, como um *limite superior variacional* para a energia por partícula do estado fundamental *exato* do gás de Fermi com interações de dois corpos, *qualquer que seja o valor do momento de Fermi* k_F . O termo proveniente da energia cinética é sempre positivo e depende de k_F como k_F^2 , enquanto a dependência dos termos de energia potencial é dominada pela dependência com k_F^3 do termo direto (pois $j_1(kr)/kr$ tem um comportamento oscilatório e é em módulo sempre menor que 1) quando a densidade, e portanto também k_F , são suficientemente grandes. Portanto, quando o potencial de dois corpos é *puramente atrativo* no sentido que $v(|\vec{r}_1 - \vec{r}_2|) \leq 0$, o limite superior para a energia por partícula do estado fundamental vai a $-\infty$ para $k_F \rightarrow \infty$. Esse comportamento revela o colapso do sistema nesse caso, isto é, a ausência de um estado fundamental com densidade e energia por partícula finitas.

Problemas.

7.1 - Considere o estado de menor energia de um gás de férmions livres, de spin 1/2, em duas e em uma dimensão espacial, respectivamente. Calcule para cada um desses dois casos

- a) o momento de Fermi k_F e a energia de Fermi E_F como funções da densidade;
- b) a energia total por partícula do gás como função da energia de Fermi e como função da densidade do gás; e
- c) a derivada da energia total por partícula com relação à densidade. No caso tri-dimensional esta última quantidade se relaciona com a pressão do gás. Obtenha essa relação, e também obtenha e interprete as relações análogas em duas e uma dimensão.

7.2 - Correlações de Pauli. Dois elétrons ocupam os estados livres

$$\langle \vec{r}s | \vec{k}_1 \rangle = \frac{1}{\sqrt{L^3}} e^{i\vec{k}_1 \cdot \vec{r}} \chi_{\frac{1}{2}}(s) \quad \text{e} \quad \langle \vec{r}s | \vec{k}_2 \rangle = \frac{1}{\sqrt{L^3}} e^{i\vec{k}_2 \cdot \vec{r}} \chi_{\frac{1}{2}}(s).$$

Note que os dois estados envolvem a mesma função de onda de spin. Tendo em vista que os dois elétrons são férmions idênticos, calcule

- a) a densidade de probabilidade para que o centro de massa dos dois elétrons esteja num ponto dado R_0 , no interior do volume de quantização L^3 ;
- b) a densidade de probabilidade para que um elétron esteja a uma distância \vec{d} do outro;
- c) a densidade de probabilidade para que haja *um* elétron na posição \vec{r}_0 .

Sugestão: nos casos **a)** e **b)**, re-escreva a função de onda dos dois elétrons em termos das coordenadas relativa e de centro de massa

$$\vec{r} = \vec{r}_2 - \vec{r}_1, \quad \vec{R} = \frac{\vec{r}_1 + \vec{r}_2}{2}.$$

- d) Repita o mesmo problema substituindo os dois estados anteriores por

$$\langle \vec{r}s | \vec{k}_1 \rangle = \frac{1}{\sqrt{L^3}} e^{i\vec{k}_1 \cdot \vec{r}} \chi_{\frac{1}{2}}(s) \quad \text{e} \quad \langle \vec{r}s | \vec{k}_2 \rangle = \frac{1}{\sqrt{L^3}} e^{i\vec{k}_2 \cdot \vec{r}} \left(\cos \theta \chi_{\frac{1}{2}}(s) + e^{i\varphi} \sin \theta \chi_{-\frac{1}{2}}(s) \right).$$

7.3 - Três partículas de massa M se movem independentemente num potencial de oscilador harmônico isotrópico de frequência ω_0 . Dê, para um tal sistema,

- a) o Hamiltoniano;
- b) todas as suas autofunções e
- c) todos os seus autovalores e degenerescências

em cada um dos seguintes casos

- a) as partículas são bosons idênticos de spin zero;
- b) as partículas são férmions idênticos de spin 1/2;
- c) as partículas são bosons distinguíveis de spin zero e
- d) as partículas são férmions distinguíveis de spin 1/2.

As autofunções devem ser dadas em termos das autofunções de uma partícula no oscilador harmônico, $\phi_{n_1 n_2 n_3}(\vec{r})$ cujos autovalores são $\epsilon_{n_1 n_2 n_3} = \hbar \omega_0 (n_1 + n_2 + n_3 + 3/2)$.

7.4 - Supondo que dois nucleons (cada um dos quais é um próton ou um neutron) interagem apenas através de um potencial central, o momento angular orbital relativo \vec{L} e o spin total

$\vec{S} = \vec{s}_1 + \vec{s}_2$ são constantes do movimento, e é então conveniente classificar os estados de dois nucleons em usando os autovalores $\hbar^2 l(l+1)$ (ou simplesmente l) de L^2 e $\hbar^2 S(S+1)$ (ou simplesmente S) de S^2 . Dado que dois prótons ou dois neutrons são férmions idênticos de spin $1/2$, e que prótons e neutrons são distinguíveis, faça uma lista das combinações possíveis de l e S para

- a) dois prótons;
- b) dois neutrons;
- c) um próton e um neutron.

Sugestão: use o fato que estados com l par (ímpar) são simétricos (anti-simétricos) pela troca das coordenadas de posição das duas partículas. (Você sabe justificar isso?)

d) O deuteron é um estado ligado de um próton e de um neutron que tem $S = 1$. Dentro da suposição feita acima para a interação dessas duas partículas, o que se pode dizer do momento angular relativo nesse caso?

7.5 - Aproximação de Hartree-Fock para o átomo de lítio. O átomo de lítio tem três elétrons que, de acordo com a distribuição esperada dos níveis de partícula independente eletrônicos (v. figura 7.3) e consistentemente com o spin eletrônico observado desse átomo, deve corresponder a uma configuração de elétrons independentes na qual dois electrons preenchem o nível $nlj \equiv 1s\ 1/2$ e um elétron adicional ocupa o nível $2s\ 1/2$. A ausência de momento angular orbital ($1s, 2s \leftrightarrow l = 0$) faz com que a contribuição desses orbitais para o campo médio seja esfericamente simétrica. Aproximando o Hamiltoniano que descreve esse átomo como

$$H = \sum_{i=1}^3 \left(\frac{p^2}{2m_e} - \frac{3e^2}{r_i} \right) + \frac{1}{2} \sum_{i,j=1}^3 \frac{e^2}{|\vec{r}_i - \vec{r}_j|}$$

e escrevendo os estados $1s\ 1/2$ e $2s\ 1/2$ explicitamente em termos de amplitudes radiais $u_{nl}(r)$ como

$$\langle \vec{r}s | 10\ 1/2m \rangle = \frac{u_{10}(r)}{r\sqrt{4\pi}} \chi_m(s) \quad \text{e} \quad \langle \vec{r}s | 20\ 1/2m \rangle = \frac{u_{20}(r)}{r\sqrt{4\pi}} \chi_m(s)$$

a) escreva a função de onda anti-simetrizada para os três electrons que corresponde ao estado fundamental do átomo sob a forma de um determinante de Slater construído com esses estados. Suponha que o spin do elétron $2s\ 1/2$ esteja no estado com projeção $+\hbar/2$ no eixo 3. Essa suposição é relevante dinamicamente, dentro da aproximação adotada para H ?

b) Calcule o valor médio de H nessa função de onda anti-simetrizada.

c) Através da variação das amplitudes radiais $u_{n0}(r)$, $n = 1, 2$, obtenha as equações integro-diferenciais acopladas para essas duas amplitudes que correspondem à aproximação de Hartree-Fock. Quais as condições de contorno que devem ser satisfeitas pelas soluções? **Sugestões:** i)

A variação se dá num espaço vetorial complexo, de modo que as variações de u_{nl} e de u_{nl}^* podem ser consideradas independentes. **ii)** Para dar conta da normalização dos estados de um elétron submetidos à variação, introduza as condições subsidiárias $\langle n0\ 1/2m | n0\ 1/2m \rangle = 1$, $n = 1, 2$ através de multiplicadores de Lagrange.

d) Identifique os termos direto (local) e de troca (não local) do potencial de Hartree-Fock. Como você faria para tentar obter uma solução dessas equações?

7.6 - Repita o item **b)** do problema anterior re-escrevendo H na linguagem de segunda quantização e usando a representação apropriada dos estados anti-simetrizados de tres elétrons no espaço de Fock. Verifique que a forma resultante para o valor esperado de H é idêntica à que se obtém usando determinantes de Slater.

7.7 - Supondo o operador de campo *fermiônico* $\psi_m(\vec{r})$ expandido numa base ortonormal e discreta de funções $\{\phi_n(\vec{r})\}$, $n = 1, 2, \dots$ como

$$\psi_m(\vec{r}) = \sum_k \phi_k(\vec{r}) a_{km}$$

a) mostre, usando as relações de (anti-)comutação para os operadores a_{km} , a_{km}^\dagger , que os operadores de número

$$N_{km} = a_{km}^\dagger a_{km}$$

são *idempotentes*, isto é, $N_{km}^2 = N_{km}$, e que são portanto operadores de projeção.

b) Diante desse fato, quais são os seus possíveis autovalores? Relacione isso com o “princípio de exclusão” de Pauli e verifique que essas propriedades *não* se aplicam no caso bosônico.

c) Mostre que

$$e^{\sum_{km} \alpha_{km} a_{km}^\dagger a_{km}} = \prod_{km} e^{\alpha_{km} a_{km}^\dagger a_{km}} = \prod_{km} \left(e^{\alpha_{km} a_{km}^\dagger a_{km}} + a_{km} a_{km}^\dagger \right).$$

7.8 - Supondo o operador de campo $\psi_m(\vec{r})$ expandido, como no problema anterior, numa base ortonormal e discreta de funções $\{\phi_n(\vec{r})\}$, $n = 1, 2, \dots$, isto é

$$\psi_m(\vec{r}) = \sum_k \phi_k(\vec{r}) a_{km},$$

mostre que, no caso da estatística de Bose, valem as relações

$$[\psi_m(\vec{r}), a_{km'}^\dagger] = \phi_k(\vec{r}) \delta_{mm'} \quad \text{e} \quad [\psi_m(\vec{r}), a_{km'}] = 0$$

enquanto que, no caso da estatística de Fermi, valem

$$\{\psi_m(\vec{r}), a_{km'}^\dagger\} = \phi_k(\vec{r})\delta_{mm'} \quad \text{e} \quad \{\psi_m(\vec{r}), a_{km'}\} = 0.$$

7.9 - Aproximação de Hartree-Fock e excitações de partícula-buraco. O hamiltoniano segundo-quantizado que descreve um sistema de muitos férmions idênticos contém termos de um e de dois corpos e é escrito como

$$H = \sum_{\alpha\beta} \langle \alpha | h_0 | \beta \rangle a_\alpha^\dagger a_\beta + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} \langle \alpha\beta | v | \gamma\delta \rangle a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma.$$

com os operadores de criação e de aniquilação definidos em termos de uma base de referência de funções de onda de uma partícula $\{\phi_\alpha(\vec{r}, s)\}$.

Considerando o estado

$$|\Psi_N\rangle = a_N^\dagger a_{N-1}^\dagger \dots a_1^\dagger |0\rangle$$

a) calcule o valor esperado de H nesse estado.

b) Obtenha as equações de Hartree-Fock para esse sistema através da condição variacional $\delta\langle\Psi_N|H|\Psi_N\rangle = 0$ aplicada ao resultado anterior. **Sugestão:** os estados (ou funções de onda) de uma partícula são vetores de um espaço vetorial *complexo* e podem portanto ser variados independentemente. Introduza multiplicadores de Lagrange para ter em conta a condição de normalização para os estados de uma partícula.

Considere em seguida o estado

$$|\Psi_N^{(aA)}\rangle = a_a^\dagger a_A |\Psi_N\rangle$$

onde o índice maiúsculo denota um dos estados de uma partícula *ocupados* em $|\Psi_N\rangle$, e o índice minúsculo denota um estado *não ocupado* nesse mesmo estado. Estados como $|\Psi_N^{(aA)}\rangle$ são chamados “excitações de partícula-buraco” de $|\Psi_N\rangle$.

c) Calcule o elemento de matriz

$$M_{aA} \equiv \langle\Psi_N^{(aA)}|H|\Psi_N\rangle$$

que pode ser pensado como a amplitude de probabilidade para que H produza a excitação de partícula-buraco em $|\Psi_N\rangle$.

d) Mostre que, se $|\Psi_N\rangle$ é o estado de Hartree-Fock e a base $\{\phi_\alpha(\vec{r}, s)\}$ é a base de Hartree-Fock, então $M_{aA} \equiv 0$, isto é, o estado de Hartree-Fock é tal que a amplitude de excitação de partícula-buraco pelo hamiltoniano é zero.

e) Obtenha uma amplitude de transição diferente de zero devida a H a partir do estado de Hartree-Fock que mostre que esse estado não é, em geral, um autoestado de H .

7.10 - Usando a expansão em ondas parciais (5.18) da onda plana mostre que

$$\int_{k \leq k_F} d^3 k e^{\pm i \vec{k} \cdot \vec{r}} = 4\pi \int_0^{k_F} dk k^2 j_0(kr).$$

Mostre em seguida que o resultado da integração sobre k pode ser expresso como

$$4\pi \int_0^{k_F} dk k^2 j_0(kr) = \frac{4\pi}{3} k_F^3 \frac{j_1(k_F r)}{k_F r}.$$

7.11 - a) Exprima em termos da densidade a energia do estado fundamental de um gás de férmions de spin 1/2 não interagentes mas constrangidos a se moverem em uma dimensão espacial apenas. Relacione a densidade com o momento de Fermi desse sistema.

b) Calcule a correção perturbativa de primeira ordem para a energia calculada no item **a)** devida a uma interação entre os férmions representada por um potencial de dois corpos que é independente dos spins e do tipo $v(1, 2) = v(|x_1 - x_2|)$. **Sugestão:** podem ser seguidos os passos utilizados nas pags. 402 e seguintes, com as alterações cabíveis devido à mudança de dimensionalidade.

c) Utilizando os resultados de **a)** e **b)** escreva um limite superior variacional para a energia por partícula no estado fundamental do gás unidimensional com interações. Utilizando o caso específico de um potencial de dois corpos gaussiano e puramente atrativo, isto é

$$v(|x_1 - x_2|) = -V_0 \exp\left(-\frac{(x_1 - x_2)^2}{b^2}\right), \quad V_0 > 0$$

obtenha uma expressão para o limite superior para a energia por partícula como função do momento de Fermi. Mostre que, diferentemente do que acontece no caso tri-dimensional, essa expressão tem um mínimo. Determine o momento de Fermi e a densidade correspondentes a esse mínimo.

Referências gerais.

Um trabalho legível sobre a questão da identidade de partículas e sua relação com a simetria de vetores de estado no contexto da mecânica quântica é

* A. M. L. Messiah e O. W. Greenberg, *Phys. Rev.* **B136**, 248, 1964.

Sistemas de muitas partículas idênticas são tratados em praticamente todos os textos de mecânica quântica. Discussões de sistemas atômicos com muitos elétrons, inclusive as aproximações de Thomas-Fermi, Hartree e Hartree-Fock podem ser encontradas em

* **A. Messiah**, op. cit., Capítulo XIV.

* **L. D. Landau e E. M. Lifshitz**, op. cit., Capítulo X.

O capítulo IX desta última referência contém um tratamento das técnicas de segunda quantização, aplicadas no contexto da mecânica quântica não relativística. Um tratamento mais detalhado dessas técnicas se encontra também em

* **A. L. Fetter e J. D. Walecka**, *Quantum Theory of Many-Particle Systems*, McGraw-Hill Book Co, 1971, Capítulo 1.

Capítulo 8

Interação da matéria com a radiação.

As interações eletromagnéticas desempenham um papel essencial na estrutura da matéria usual. Os modelos estudados para o átomo de hidrogênio e também para átomos mais complexos dão exemplos claros desse fato. As cargas e correntes associadas aos seus constituintes são ainda, por outro lado, as fontes do campo de radiação eletromagnética, que portanto está também essencialmente acoplado à matéria. O propósito deste capítulo é apresentar um esquema perturbativo simples para tratar a interação da matéria com o campo de radiação. Para isso a descrição clássica do campo de radiação livre será brevemente revista com vistas à sua quantização. Uma discussão do acoplamento do campo de radiação quantizado com sistemas materiais carregados levará então ao esquema perturbativo mencionado, que permite tratar problemas de espalhamento, emissão radiativa, etc.

8.1 O campo de radiação livre.

O campo de radiação livre, isto é, as propriedades eletromagnéticas do vácuo na ausência de fontes (densidade de carga ρ e densidade de corrente \vec{j} nulas), pode ser representado de forma particularmente compacta e conveniente em termos apenas de um potencial vetor $\vec{A}(\vec{r}, t)$ relacionado com os campos elétrico $\vec{E}(\vec{r}, t)$ e magnético $\vec{B}(\vec{r}, t)$ através das relações (em unidades Gaussianas)

$$\vec{E}(\vec{r}, t) = -\frac{1}{c} \frac{\partial \vec{A}(\vec{r}, t)}{\partial t} \quad \text{e} \quad \vec{B}(\vec{r}, t) = \nabla \times \vec{A}(\vec{r}, t).$$

Nessa representação o potencial escalar é identicamente nulo, $\Phi(\vec{r}, t) \equiv 0$, e a dinâmica do campo, regida pelas equações de Maxwell sem cargas ou correntes, pode ser reduzida à equação de ondas

$$\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \vec{A}(\vec{r}, t) = 0$$

com a condição de “padrão” (ou ‘gauge’)

$$\nabla \cdot \vec{A}(\vec{r}, t) = 0,$$

que define o chamado “padrão de Coulomb”. Como os campos \vec{E} e \vec{B} são quantidades necessariamente *reais* e são obtidas do potencial vetor simplesmente por diferenciação, resulta que, nessa representação, $\vec{A}(\vec{r}, t)$ deve ser também real, isto é

$$\vec{A}^*(\vec{r}, t) = \vec{A}(\vec{r}, t).$$

A solução geral da equação de ondas para esse campo, com condições de contorno periódicas num volume L^3 , pode ser escrita como superposição de soluções “elementares” $\vec{A}_{\vec{k}}(\vec{r}, t)$ que têm a forma de ondas planas escritas como

$$\vec{A}_{\vec{k}}(\vec{r}, t) = \sqrt{\frac{4\pi c^2}{L^3}} \vec{\epsilon}_{\vec{k}} q_{\vec{k}}(t) e^{i\vec{k} \cdot \vec{r}} \quad (8.1)$$

onde os coeficientes incluídos na raiz quadrada são introduzidos para fixar convenientemente unidades e escalas das amplitudes $q_{\vec{k}}(t)$. O caráter vetorial de \vec{A} é determinado pelo vetor de polarização $\vec{\epsilon}_{\vec{k}}$. A relação geral entre o potencial vetor e o campo elétrico mostra, de fato, que o campo elétrico $\vec{E}_{\vec{k}}$ associado à solução elementar é *anti-paralelo* a $\vec{\epsilon}_{\vec{k}}$.

O fato de que os $\vec{A}_{\vec{k}}(\vec{r}, t)$ devem satisfazer tanto a equação de ondas como também a condição de padrão introduz várias restrições para os diferentes objetos que aparecem nessas soluções elementares. A condição de padrão $\nabla \cdot \vec{A}_{\vec{k}}(\vec{r}, t) = 0$ de fato conduz imediatamente a

$$\vec{\epsilon}_{\vec{k}} \cdot \vec{k} = 0$$

o que significa que o vetor de polarização deve ser perpendicular ao vetor de onda \vec{k} . A condição de padrão aparece dessa forma como uma *condição de transversalidade* para o campo, que permite associar dois estados linearmente independentes de polarização, $\vec{\epsilon}_{\vec{k}, \alpha}$, $\alpha = 1, 2$ a cada vetor de onda \vec{k} . A figura 8.1 mostra uma forma conveniente de escolher esses estados. Ela consiste em 1. dado um vetor de onda \vec{k} , tomar como vetores básicos de polarização dois vetores unitários perpendiculares a \vec{k} e também perpendiculares entre si e 2. associar *os mesmos* vetores básicos de polarização a vetores de onda de mesmo módulo mas de sentidos opostos, \vec{k} e $-\vec{k}$. Desse modo



Figura 8.1: Convenção para escolha de vetores de polarização para os vetores de onda \vec{k} e $-\vec{k}$.

$$\vec{\epsilon}_{\vec{k}\alpha} = \vec{\epsilon}_{-\vec{k}\alpha}, \quad \alpha = 1, 2.$$

A equação de ondas, por outro lado, define a dependência temporal das amplitudes $q_{\vec{k}}(t)$ das soluções elementares, escritas agora tendo em conta as possíveis diferentes polarizações como

$$\vec{A}_{\vec{k}\alpha}(\vec{r}, t) = \sqrt{\frac{4\pi c^2}{L^3}} \vec{\epsilon}_{\vec{k}\alpha} q_{\vec{k}\alpha}(t) e^{i\vec{k}\cdot\vec{r}}.$$

Da equação

$$\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \vec{A}_{\vec{k}\alpha}(\vec{r}, t) = 0$$

resulta, de fato, que

$$-k^2 q_{\vec{k}\alpha}(t) - \frac{1}{c^2} \frac{d^2 q_{\vec{k}\alpha}(t)}{dt^2} = 0 \quad \text{ou} \quad \frac{d^2 q_{\vec{k}\alpha}(t)}{dt^2} + \omega_k^2 q_{\vec{k}\alpha}(t) = 0 \quad (8.2)$$

com a frequência ω_k definida como $\omega_k = kc$. Essa equação mostra que as amplitudes das soluções elementares oscilam harmonicamente com frequência ω_k .

É importante notar ainda que, não obstante a condição de realidade que deve ser satisfeita pelo potencial vetor, as amplitudes $q_{\vec{k}\alpha}(t)$ podem em geral ser complexas. De fato, uma solução geral da equação de ondas pode ser escrita como

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}, \alpha} \vec{A}_{\vec{k}\alpha}(\vec{r}, t)$$

e o que é exigido pela condição de realidade é que

$$\sum_{\vec{k}, \alpha} \vec{A}_{\vec{k}\alpha}(\vec{r}, t) = \sum_{\vec{k}, \alpha} \vec{A}_{\vec{k}\alpha}^*(\vec{r}, t)$$

ou seja

$$\sum_{\vec{k}, \alpha} \vec{\epsilon}_{\vec{k}\alpha} q_{\vec{k}\alpha}(t) e^{i\vec{k}\cdot\vec{r}} = \sum_{\vec{k}, \alpha} \vec{\epsilon}_{\vec{k}\alpha} q_{\vec{k}\alpha}^*(t) e^{-i\vec{k}\cdot\vec{r}} = \sum_{\vec{k}, \alpha} \vec{\epsilon}_{-\vec{k}\alpha} q_{-\vec{k}\alpha}^*(t) e^{i\vec{k}\cdot\vec{r}}.$$

Da independência linear das ondas planas no volume de periodicidade L^3 resulta portanto que a realidade do potencial vetor exige que se tenha

$$\sum_{\alpha} \vec{\epsilon}_{\vec{k}\alpha} q_{\vec{k}\alpha}(t) = \sum_{\alpha} \vec{\epsilon}_{-\vec{k}\alpha} q_{-\vec{k}\alpha}^*(t).$$

Fazendo uso da convenção mostrada na figura 8.1 para a escolha dos vetores básicos de polarização, essa condição se reduz a

$$q_{\vec{k}\alpha}(t) = q_{-\vec{k}\alpha}^*(t). \quad (8.3)$$

8.1.1 Quantização do campo de radiação livre.

O tratamento precedente permitiu reduzir o campo de radiação livre a uma coleção infinita de osciladores harmônicos complexos, associados às soluções elementares $\vec{A}_{\vec{k}\alpha}$. A quantização desse campo pode portanto ser feita simplesmente quantizando essa coleção de osciladores harmônicos. No caso de um oscilador *real*, cuja dinâmica clássica é dada em termos de uma equação de Newton

$$\frac{d^2x}{dt^2} + \omega^2 x = 0$$

isso pode ser feito definindo inicialmente um momento p canonicamente conjugado a x como

$$p \equiv \frac{dx}{dt}$$

o que corresponde a dizer que a dinâmica do oscilador é governada pela hamiltoniana

$$h = \frac{p^2}{2} + \frac{\omega^2}{2} x^2.$$

A quantização se faz então reinterpretando a amplitude x e o momento p como operadores hermiteanos de um espaço de Hilbert que satisfazem a relação de comutação canônica

$$[x, p] = i\hbar.$$

Com isso, a equação de Newton original se torna a equação dinâmica do oscilador quântico na descrição de Heisenberg. Os operadores de criação e de aniquilação a e a^\dagger do oscilador serão dados nessa descrição por

$$x(t) = \sqrt{\frac{\hbar}{2\omega}} [ae^{-i\omega t} + a^\dagger e^{i\omega t}] \quad \text{e} \quad p(t) = \frac{dx(t)}{dt} = i\sqrt{\frac{\hbar\omega}{2}} [a^\dagger e^{i\omega t} - ae^{-i\omega t}].$$

Eles são independentes do tempo e satisfazem a relação de comutação usual $[a, a^\dagger] = 1$. Os operadores independentes do tempo x e p da descrição de Schrödinger, por outro lado, podem ser identificados com os operadores de Heisenberg em $t = 0$.

Esse mesmo esquema pode ser facilmente adaptado para a quantização dos osciladores complexos que correspondem às soluções elementares para o campo de radiação livre. De fato, a amplitude complexa que obedece à equação de movimento (8.2) pode ser decomposta em duas amplitudes reais separando suas partes real e imaginária

$$q_{\vec{k}\alpha} \equiv \frac{1}{\sqrt{2}} (q_{\vec{k}\alpha}^{(R)} + iq_{\vec{k}\alpha}^{(I)})$$

cada uma das quais satisfaz uma equação de movimento idêntica à (8.2) mas *real*, e que portanto pode ser quantizada da mesma forma que o oscilador real x . Os momentos canonicamente conjugados a $q_{\vec{k}\alpha}^{(R)}$ e $q_{\vec{k}\alpha}^{(I)}$ são definidos como

$$p_{\vec{k}\alpha}^{(R)} \equiv \frac{dq_{\vec{k}\alpha}^{(R)}}{dt} \quad \text{e} \quad p_{\vec{k}\alpha}^{(I)} \equiv \frac{dq_{\vec{k}\alpha}^{(I)}}{dt}$$

e em seguida re-interpretados como operadores hermiteanos satisfazendo as relações de comutação

$$\begin{aligned} [q_{\vec{k}\alpha}^{(R)}, p_{\vec{k}'\alpha'}^{(R)}] &= [q_{\vec{k}\alpha}^{(I)}, p_{\vec{k}'\alpha'}^{(I)}] = i\hbar\delta_{\vec{k}\vec{k}'}\delta_{\alpha\alpha'}, \\ [q_{\vec{k}\alpha}^{(R)}, p_{\vec{k}'\alpha'}^{(I)}] &= [q_{\vec{k}\alpha}^{(I)}, p_{\vec{k}'\alpha'}^{(R)}] = 0, \quad [q_{\vec{k}\alpha}^{(R)}, q_{\vec{k}'\alpha'}^{(I)}] = [p_{\vec{k}\alpha}^{(R)}, p_{\vec{k}'\alpha'}^{(I)}] = 0 \end{aligned}$$

com o que as equações de movimento são re-interpretadas como equações de Heisenberg para as amplitudes quantizadas do campo. É fácil verificar explicitamente que o mesmo resultado pode ser obtido definindo um momento *complexo* canonicamente conjugado à amplitude complexa $q_{\vec{k}\alpha}$ como

$$p_{\vec{k}\alpha} \equiv \frac{1}{\sqrt{2}} \left[\frac{dq_{\vec{k}\alpha}^{(R)}}{dt} - i \frac{dq_{\vec{k}\alpha}^{(I)}}{dt} \right] = \frac{dq_{\vec{k}\alpha}^*}{dt} = \frac{dq_{-\vec{k}\alpha}}{dt},$$

onde, na última passagem, foi utilizada a condição de realidade (8.3), e quantizando através das relações de comutação

$$[q_{\vec{k}\alpha}, p_{\vec{k}'\alpha'}] = i\hbar\delta_{\vec{k}\vec{k}'}\delta_{\alpha\alpha'}, \quad [q_{\vec{k}\alpha}, q_{\vec{k}'\alpha'}] = 0, \quad [p_{\vec{k}\alpha}, p_{\vec{k}'\alpha'}] = 0.$$

Esta última forma de quantização é também conveniente por permitir definir operadores de criação e de aniquilação associados às soluções elementares através das relações

$$q_{\vec{k}\alpha}(t) = \sqrt{\frac{\hbar}{2\omega_k}} [a_{\vec{k}\alpha} e^{-i\omega_k t} + a_{-\vec{k}\alpha}^\dagger e^{i\omega_k t}] \quad \text{e} \quad p_{\vec{k}\alpha}(t) = i\sqrt{\frac{\hbar\omega_k}{2}} [a_{\vec{k}\alpha}^\dagger e^{i\omega_k t} - a_{-\vec{k}\alpha} e^{-i\omega_k t}] \quad (8.4)$$

que satisfazem as condições de realidade (hermiticidade) do campo $q_{\vec{k}\alpha} = q_{-\vec{k}\alpha}^\dagger$, $p_{\vec{k}\alpha} = p_{-\vec{k}\alpha}^\dagger$ e das quais seguem as relações de comutação

$$[a_{\vec{k}\alpha}, a_{\vec{k}'\alpha'}^\dagger] = \delta_{\vec{k}\vec{k}'}\delta_{\alpha\alpha'} \quad \text{e} \quad [a_{\vec{k}\alpha}, a_{\vec{k}'\alpha'}] = 0.$$

A quantização do campo de radiação livre, com algumas simplificações técnicas mas de outro modo seguindo um esquema semelhante ao utilizado acima, foi considerada já em 1925 por Heisenberg, Born e Jordan¹ e representa uma escalada importante na expansão dos domínios da teoria quântica. De fato, as idéias dessa teoria foram implementadas pela primeira vez no contexto da mecânica de partículas, na qual as variáveis dinâmicas são tipicamente variáveis canônicas de posição e momento. A quantização consistiu em reinterpretar essas variáveis dinâmicas em termos de operadores lineares agindo sobre amplitudes de probabilidade, subvertendo dessa forma a cinemática clássica completamente. No contexto de uma teoria clássica de campos, as variáveis dinâmicas são os valores assumidos pelo(s) campo(s) (como funções do tempo) em cada posição do espaço, e a quantização consiste em reinterpretar *esses valores* em termos de operadores lineares agindo sobre amplitudes de probabilidade. A lição que se pode tirar desses fatos é a de que o conteúdo essencial da teoria quântica admite ser implementado em diferentes contextos, e não apenas no contexto mecânico no qual foi pela primeira vez formulado. Em outras palavras, existe uma “teoria quântica” com abrangência maior que a da mecânica quântica propriamente dita.

Para simplificar a notação é útil introduzir as abreviações

$$\lambda \equiv \{\vec{k}, \alpha\}, \quad -\lambda \equiv \{-\vec{k}, \alpha\}, \quad \text{e} \quad \omega_k \equiv \omega_\lambda = \omega_{-\lambda}$$

¹W. Heisenberg, M. Born e P. Jordan, *Zeit. f. Physik* **35** 557 (1925). Este trabalho está incluído no livro de van der Waerden (*op. cit.*).

com as quais as amplitudes $q_{\vec{k}\alpha}$ passam a ser escritas como

$$q_{\lambda}(t) = \sqrt{\frac{\hbar}{2\omega_{\lambda}}} \left[a_{\lambda} e^{-i\omega_{\lambda}t} + a_{-\lambda}^{\dagger} e^{i\omega_{\lambda}t} \right]$$

e o potencial vetor quantizado $\vec{A}(\vec{r}, t)$ é dado por

$$\begin{aligned} \vec{A}(\vec{r}, t) &= \sum_{\lambda} \sqrt{\frac{4\pi c^2}{L^3}} \vec{\epsilon}_{\lambda} q_{\lambda} e^{i\vec{k} \cdot \vec{r}} = \\ &= \sqrt{\frac{1}{L^3}} \sum_{\lambda} \sqrt{\frac{2\pi \hbar c}{k}} \vec{\epsilon}_{\lambda} \left[a_{\lambda} e^{-i\omega_{\lambda}t} + a_{-\lambda}^{\dagger} e^{i\omega_{\lambda}t} \right] e^{i\vec{k} \cdot \vec{r}} = \\ &= \sqrt{\frac{1}{L^3}} \sum_{\lambda} \sqrt{\frac{2\pi \hbar c}{k}} \vec{\epsilon}_{\lambda} \left[a_{\lambda} e^{i(\vec{k} \cdot \vec{r} - \omega_{\lambda}t)} + a_{\lambda}^{\dagger} e^{-i(\vec{k} \cdot \vec{r} - \omega_{\lambda}t)} \right]. \end{aligned} \quad (8.5)$$

Para obter a última expressão a convenção relativa aos vetores de polarização foi utilizada para substituir o índice mudo $-\lambda$ por λ . Vale a pena notar mais uma vez que esta expressão corresponde à descrição de Heisenberg do campo quantizado, e que o operador de Schrödinger para o potencial vetor pode ser definido como $\vec{A}(\vec{r}, t=0)$, o que equivale a escolher $t=0$ como o instante em que os operadores das duas representações coincidem. Como consequência da condição de realidade sobre as amplitudes q_{λ} esta expressão para $\vec{A}(\vec{r}, t)$ é hermiteana.

8.1.2 Energia, momento e momento angular do campo livre.

1. Energia. O fato de ter sido possível reduzir o campo de radiação livre a uma coleção (infinita) de osciladores harmônicos independentes sugere imediatamente que o Hamiltoniano do campo livre possa ser escrito como a soma dos Hamiltonianos dos diferentes osciladores, e ainda que os autovetores de energia do campo possam ser realizados em termos de produtos de autovetores dos osciladores individuais. Isso leva a um Hamiltoniano da forma

$$H_{\text{rad}} = \sum_{\lambda} \hbar \omega_{\lambda} \left(a_{\lambda}^{\dagger} a_{\lambda} + \frac{1}{2} \right) \quad (8.6)$$

e aos autovetores e autovalores

$$|\{n_{\lambda}\}\rangle \equiv |n_1\rangle \otimes |n_2\rangle \otimes \dots \otimes |n_{\lambda}\rangle \otimes \dots \quad \Longleftrightarrow \quad E_{\{n_{\lambda}\}} = \sum_{\lambda} \hbar \omega_{\lambda} \left(n_{\lambda} + \frac{1}{2} \right). \quad (8.7)$$

Os n_λ são números inteiros que indicam o número de excitações presentes no oscilador correspondente $\lambda \equiv \{\vec{k}, \vec{\epsilon}\}$, isto é

$$a_\lambda^\dagger a_\lambda |n_\lambda\rangle = n_\lambda |n_\lambda\rangle \quad \text{e ainda} \quad a_\lambda^\dagger a_\lambda |n_1, n_2, \dots, n_\lambda, \dots\rangle = n_\lambda |n_1, n_2, \dots, n_\lambda, \dots\rangle. \quad (8.8)$$

A contribuição de cada excitação para o autovalor de energia do estado do campo é portanto $\hbar\omega_\lambda$, o que sugere a sua interpretação em termos de *fótons* com vetor de onda \vec{k} e polarização $\vec{\epsilon}$. Todos os autovalores contém uma soma divergente comum $\sum_\lambda \hbar\omega_\lambda/2$, que pode ser reconhecida como o autovalor do estado no qual todos os $n_\lambda = 0$, isto é, do estado sem nenhum fóton. Para evitar a inconveniência dessa constante infinita correspondente à energia do estado de zero fótons (o “vácuo”), é conveniente e comum redefinir o zero da energia do campo como a energia do vácuo, o que equivale a eliminar as contribuições de ponto zero de todos os modos do campo, redefinindo o Hamiltoniano e os seus autovalores simplesmente como

$$H_{\text{rad}} \longrightarrow \sum_\lambda \hbar\omega_\lambda a_\lambda^\dagger a_\lambda \quad \text{e} \quad E_{\{n_\lambda\}} \longrightarrow \sum_\lambda \hbar\omega_\lambda n_\lambda.$$

Uma forma de verificar que essas escolhas são de fato apropriadas consiste em calcular a expressão *clássica* para a energia eletromagnética

$$U_{\text{rad}} = \frac{1}{8\pi} \int d^3r \left(\vec{E}^2 + \vec{B}^2 \right) = \frac{1}{8\pi} \int d^3r \left[\frac{1}{c^2} \left(\frac{\partial \vec{A}}{\partial t} \right)^2 + (\nabla \times \vec{A})^2 \right]$$

utilizando a expressão quantizada (8.5) para o potencial vetor $\vec{A}(\vec{r}, t)$. A derivada com relação ao tempo que corresponde ao campo elétrico é

$$\frac{\partial \vec{A}}{\partial t} = \sqrt{\frac{4\pi c^2}{L^3}} \sum_\lambda \vec{\epsilon}_\lambda \frac{dq_\lambda}{dt} e^{i\vec{k} \cdot \vec{r}} = \sqrt{\frac{4\pi c^2}{L^3}} \sum_\lambda \vec{\epsilon}_\lambda p_{-\lambda} e^{i\vec{k} \cdot \vec{r}}$$

onde a condição de realidade de \vec{A} e a definição do momento canônico p_λ foram utilizadas para escrever

$$\frac{dq_\lambda}{dt} = p_\lambda^\dagger = p_{-\lambda}.$$

Vale a pena notar ainda que a convenção da figura 8.1 também foi utilizada para obter a última expressão para $\partial \vec{A} / \partial t$. O termo correspondente ao campo magnético pode ser transformado usando as identidades

$$(\nabla \times \vec{A}) \cdot (\nabla \times \vec{A}) \equiv \nabla \cdot [\vec{A} \times (\nabla \times \vec{A})] + \vec{A} \cdot \nabla \times \nabla \times \vec{A}$$

e

$$\nabla \times \nabla \times \vec{A} \equiv \nabla (\nabla \cdot \vec{A}) - \nabla^2 \vec{A}.$$

O termo $\nabla \cdot [\vec{A} \times (\nabla \times \vec{A})]$ envolve um divergente e sua integral de volume pode ser reduzida a uma integral de superfície que se anula devido às condições de contorno periódicas impostas sobre \vec{A} . A condição de transversalidade implica, por outro lado no anulamento do termo que contém $\nabla \cdot \vec{A}$, de modo que a contribuição do campo magnético para a energia do campo se reduz a $-\vec{A} \cdot \nabla^2 \vec{A}$. O laplaciano do potencial vetor é calculado facilmente como

$$\nabla^2 \vec{A} = -\sqrt{\frac{4\pi c^2}{L^3}} \sum_{\lambda} \vec{\epsilon}_{\lambda} k^2 q_{\lambda} e^{i\vec{k} \cdot \vec{r}}$$

de modo que a energia do campo é dada por

$$U_{\text{rad}} = \frac{1}{8\pi} \frac{4\pi c^2}{L^3} \sum_{\lambda\lambda'} \vec{\epsilon}_{\lambda} \cdot \vec{\epsilon}_{\lambda'} \left(\frac{1}{c^2} p_{-\lambda} p_{-\lambda'} + k^2 q_{\lambda} q_{\lambda'} \right) \int d^3r e^{i(\vec{k}+\vec{k}') \cdot \vec{r}}.$$

As condições de contorno periódicas impostas sobre as ondas planas dão agora

$$\int d^3r e^{i(\vec{k}+\vec{k}') \cdot \vec{r}} = V \delta_{\vec{k}, \vec{k}'},$$

e, usando a convenção da figura 8.1 para os vetores de polarização,

$$\vec{\epsilon}_{\vec{k}\alpha} \cdot \vec{\epsilon}_{-\vec{k}\alpha'} = \delta_{\alpha\alpha'}.$$

Portanto

$$U_{\text{rad}} = \frac{1}{2} \sum_{\lambda} (p_{-\lambda} p_{\lambda} + \omega_{\lambda}^2 q_{-\lambda} q_{\lambda}) = \frac{1}{2} \sum_{\lambda} (p_{\lambda}^{\dagger} p_{\lambda} + \omega_{\lambda}^2 q_{\lambda}^{\dagger} q_{\lambda})$$

o que corresponde a uma soma de Hamiltonianos de osciladores (complexos) independentes. Re-escrevendo os p_{λ} e os q_{λ} em termos dos operadores a_{λ} e a_{λ}^{\dagger} resulta finalmente

$$U_{\text{rad}} = \sum_{\lambda} \frac{\hbar\omega_{\lambda}}{2} (a_{\lambda}^{\dagger} a_{\lambda} + a_{\lambda} a_{\lambda}^{\dagger}) = \sum_{\lambda} \hbar\omega_{\lambda} \left(a_{\lambda}^{\dagger} a_{\lambda} + \frac{1}{2} \right) = H_{\text{rad}}.$$

Esse resultado mostra, em particular, que U_{rad} resulta ser uma quantidade *independente do tempo*, embora seus ingredientes sejam expressos em termos da versão de Heisenberg, e portanto dependente do tempo, de $\vec{A}(\vec{r}, t)$. A razão para isso é que dependência temporal dos operadores de Heisenberg associados ao campo é justamente a que é gerada por $H_{\text{rad}} = U$. Assim, por exemplo, a dependência temporal de $q_\lambda(t)$ satisfaz, como se pode verificar facilmente,

$$i\hbar \frac{dq_\lambda(t)}{dt} = [q_\lambda(t), H_{\text{rad}}] = [q_\lambda(t), U].$$

O fato de que U_{rad} resulta ser independente do tempo é portanto apenas um caso particular do resultado geral de que um hamiltoniano (visto como gerador da evolução temporal do sistema em questão) que não depende *explicitamente* do tempo, é independente do tempo (isto é, uma constante do movimento).

2. Momento. O momento \vec{P}_{rad} associado ao campo de radiação livre pode ser identificado de uma forma semelhante à utilizada no caso da energia total U_{rad} , partindo da expressão clássica

$$\vec{P}_{\text{cl}} = \frac{1}{4\pi c} \int d^3r \vec{E}(\vec{r}, t) \times \vec{B}(\vec{r}, t)$$

e calculando-a em termo das expressões quantizadas para os campos

$$\vec{E}(\vec{r}, t) = -\frac{1}{c} \sqrt{\frac{4\pi c^2}{L^3}} \sum_{\lambda} \vec{\epsilon}_{\lambda} p_{-\lambda}(t) e^{i\vec{k} \cdot \vec{r}} \quad \text{e} \quad \vec{B}(\vec{r}, t) = \sqrt{\frac{4\pi c^2}{L^3}} \sum_{\lambda} i\vec{k} \times \vec{\epsilon}_{\lambda} q_{\lambda}(t) e^{i\vec{k} \cdot \vec{r}}.$$

Uma dificuldade que aparece aqui é a de que as expressões quânticas para $\vec{E}(\vec{r}, t)$ e $\vec{B}(\vec{r}, t)$ não comutam, o que compromete a hermiticidade do operador que resulta da substituição direta dessas expressões em \vec{P}_{cl} . Essa dificuldade pode ser contornada usando para o momento total a expressão simetrizada

$$\begin{aligned} \vec{P}_{\text{rad}} &= \frac{1}{8\pi c} \int d^3r \left(\vec{E}(\vec{r}, t) \times \vec{B}(\vec{r}, t) - \vec{B}(\vec{r}, t) \times \vec{E}(\vec{r}, t) \right) \\ &= \frac{i}{8\pi c^2} \frac{4\pi c^2}{L^3} \sum_{\lambda, \lambda'} \vec{\epsilon}_{\lambda} \times (\vec{k}' \times \vec{\epsilon} - \lambda') p_{-\lambda} q_{\lambda'} \int d^3r e^{i(\vec{k}=\vec{k}') \cdot \vec{r}} + \text{h.c.} \\ &= \frac{i}{2} \sum_{\alpha, \alpha'} \sum_{\vec{k}} \vec{\epsilon}_{\vec{k}\alpha} \times (\vec{k} \times \vec{\epsilon}_{-\vec{k}\alpha'}) p_{-\vec{k}\alpha} q_{-\vec{k}\alpha'} + \text{h.c.} . \end{aligned}$$

Nessas expressões h.c. significa o hermiteano conjugado do termo precedente. Deve ser notado o sinal *negativo* do termo $\vec{B}(\vec{r}, t) \times \vec{E}(\vec{r}, t)$, que resulta da propriedade de antissimetria do produto

vetorial. O duplo produto vetorial que aparece na última forma dessa expressão pode ser escrito como

$$\vec{\epsilon}_{\vec{k}\alpha} \times (\vec{k} \times \vec{\epsilon}_{-\vec{k}\alpha'}) = (\vec{\epsilon}_{\vec{k}\alpha} \cdot \vec{\epsilon}_{-\vec{k}\alpha'}) \vec{k} - \vec{\epsilon}_{\vec{k}\alpha} (\vec{k} \cdot \vec{\epsilon}_{-\vec{k}\alpha'})$$

sendo que o último termo se anula devido ao caráter transversal do campo. Desse modo resulta

$$\vec{P}_{\text{rad}} = \frac{i}{2} \sum_{\lambda} \vec{k} (p_{-\lambda} q_{-\lambda} - p_{-\lambda}^{\dagger} q_{-\lambda}^{\dagger}) = \frac{i}{2} \sum_{\lambda} \vec{k} (p_{-\lambda} q_{-\lambda} - p_{\lambda} q_{\lambda}).$$

Os p_{λ} e q_{λ} podem ser escritos em termos dos operadores a_{λ} e a_{λ}^{\dagger} usando as relações (8.4), o que leva finalmente à expressão para o momento total do campo de radiação

$$\vec{P}_{\text{rad}} = \sum_{\lambda} \hbar \vec{k} \left(a_{\lambda}^{\dagger} a_{\lambda} + \frac{1}{2} \right) = \sum_{\lambda} \hbar \vec{k} a_{\lambda}^{\dagger} a_{\lambda}. \quad (8.9)$$

Como os operadores $n_{\lambda} = a_{\lambda}^{\dagger} a_{\lambda}$ correspondem ao número de fótons em cada modo λ , a soma $\sum_{\lambda} \hbar \vec{k}/2$ corresponde ao momento total do vácuo (estado de zero fótons), e foi tomada como sendo igual a zero. Para cada termo $\lambda \equiv \{\vec{k}, \alpha\}$ essa soma contém de fato um termo igual e oposto, correspondente a $-\lambda \equiv \{-\vec{k}, \alpha\}$.

A expressão final (8.9) mostra que a cada fóton no modo $\lambda \equiv \{\vec{k}, \alpha\}$ está associada uma contribuição $\hbar \vec{k}$ para o momento total. Disso decorre, em particular, que a energia correspondente a esse fóton está relacionada com o módulo do momento através de

$$E_{\lambda} = \hbar \omega_{\lambda} = |\hbar \vec{k}| c$$

que caracteriza a relação entre energia e momento para objetos relativísticos de massa zero.

3. Componente intrínseca do momento angular. A expressão para o momento total do campo de radiação estudada no item anterior pode ser vista como definindo uma densidade de momento \vec{G}_{rad} dada por

$$\vec{G}_{\text{rad}}(\vec{r}, t) = \frac{1}{8\pi c} \left(\vec{E}(\vec{r}, t) \times \vec{B}(\vec{r}, t) - \vec{B}(\vec{r}, t) \times \vec{E}(\vec{r}, t) \right)$$

a partir da qual se pode escrever uma densidade de momento angular, relativamente à origem a que são referidos os vetores de posição \vec{r} ,

$$\vec{r} \times \vec{G}_{\text{rad}} = \frac{1}{8\pi c} \left[\vec{r} \times \left(\vec{E}(\vec{r}, t) \times \vec{B}(\vec{r}, t) \right) - \vec{r} \times \left(\vec{B}(\vec{r}, t) \times \vec{E}(\vec{r}, t) \right) \right].$$

Essas expressões são obtidas diretamente de expressões clássicas correspondentes, tendo em conta a não comutatividade dos campos quantizados $\vec{E}(\vec{r}, t)$ e $\vec{B}(\vec{r}, t)$. Uma propriedade importante desta última expressão para a densidade de momento angular é que ela pode ser decomposta em duas partes, uma das quais *não depende explicitamente de \vec{r}* . Essa parte representa uma contribuição das propriedades intrínsecas do campo para o momento angular, e corresponde portanto ao spin campo. De fato, o duplo produto vetorial pode ser transformado como

$$\vec{r} \times (\vec{E} \times \vec{B}) = \vec{E} (\vec{r} \cdot \vec{B}) - (\vec{r} \cdot \vec{E}) \vec{B} = \vec{E} (\vec{r} \cdot \nabla \times \vec{A}) - (\vec{r} \cdot \vec{E}) \nabla \times \vec{A}.$$

Nessas expressões a ordem dos campos \vec{E} e \vec{B} foi mantida em cada um dos termos, o que é essencial devido à sua não comutatividade. O segundo termo pode ainda ser transformado como

$$- (\vec{r} \cdot \vec{E}) \nabla \times \vec{A} = [\nabla (\vec{r} \cdot \vec{E})] \times \vec{A} - \nabla \times [(\vec{r} \cdot \vec{E}) \vec{A}] = [(\vec{r} \cdot \nabla) \vec{E}] \times \vec{A} + \vec{E} \times \vec{A} - \nabla \times [(\vec{r} \cdot \vec{E}) \vec{A}].$$

Não é difícil verificar diretamente que a contribuição final do último termo depende explicitamente de \vec{r} , apesar da presença do operador diferencial $\nabla \times$, de modo que, juntando os resultados anteriores,

$$\vec{r} \times (\vec{E} \times \vec{B}) = \vec{E} (\vec{r} \cdot \nabla \times \vec{A}) + [(\vec{r} \cdot \nabla) \vec{E}] \times \vec{A} - \nabla \times [(\vec{r} \cdot \vec{E}) \vec{A}] + \vec{E} \times \vec{A}$$

e o último termo é o único que dá contribuições que não dependem explicitamente de \vec{r} . O termo em que as posições de \vec{E} e \vec{B} são trocadas leva à hermiteana conjugada dessa quantidade, levando finalmente à forma para a componente intrínseca do momento angular do campo

$$\vec{I}_{\text{rad}} = \frac{1}{8\pi c} \int d^3r [\vec{E}(\vec{r}, t) \times \vec{A}(\vec{r}, t) + \text{h.c.}] .$$

O operador \vec{I}_{rad} pode ser expresso sem dificuldade em termos da expansão do campo em ondas planas (8.5), seguindo as mesmas linhas dos itens anteriores. De fato,

$$\begin{aligned} \int d^3r [\vec{E}(\vec{r}, t) \times \vec{A}(\vec{r}, t)] &= -\frac{1}{c} \frac{4\pi c^2}{L^3} \sum_{\lambda, \lambda'} (\vec{\epsilon}_{\lambda} \times \vec{\epsilon}_{\lambda'}) p_{-\lambda} q_{\lambda'} L^3 \delta_{\vec{k}, \vec{k}'} \\ &= -4\pi c \sum_{\alpha, \alpha'} \sum_{\vec{k}} (\vec{\epsilon}_{\vec{k}\alpha} \times \vec{\epsilon}_{-\vec{k}\alpha'}) p_{-\vec{k}\alpha} q_{-\vec{k}\alpha'}. \end{aligned}$$

Esta última expressão contém o produto vetorial de dois vetores de polarização e, ao contrário do que ocorre no caso do cálculo da energia e do momento do campo, só não se anula quando

$\alpha \neq \alpha'$. Explicitando a soma sobre polarizações, e fazendo uso da anti-simetria do produto vetorial resulta

$$\int d^3r \left[\vec{E}(\vec{r}, t) \times \vec{A}(\vec{r}, t) \right] = -4\pi c \sum_{\vec{k}} \left(\vec{\epsilon}_{\vec{k}1} \times \vec{\epsilon}_{\vec{k}2} \right) \left(p_{\vec{k}1} q_{\vec{k}2} - p_{\vec{k}2} q_{\vec{k}1} \right).$$

O passo final consiste em exprimir os $p_{\vec{k}\alpha}$ e $q_{\vec{k}\alpha}$ em termos de operadores de criação e de aniquilação, isto é

$$\begin{aligned} p_{\vec{k}1} q_{\vec{k}2} - p_{\vec{k}1} q_{\vec{k}2} &= -\frac{i\hbar}{c} \left[\left(a_{\vec{k}1}^\dagger e^{i\omega t} - a_{-\vec{k}1} e^{-i\omega t} \right) \left(a_{\vec{k}2} e^{-i\omega t} + a_{-\vec{k}2}^\dagger e^{i\omega t} \right) - \right. \\ &\quad \left. - \left(a_{\vec{k}2}^\dagger e^{i\omega t} - a_{-\vec{k}2} e^{-i\omega t} \right) \left(a_{\vec{k}1} e^{-i\omega t} + a_{-\vec{k}1}^\dagger e^{i\omega t} \right) \right]. \end{aligned}$$

Os termos dependentes do tempo (contendo dois operadores de criação ou dois operadores de aniquilação) aparecem nessa expressão aos pares e com sinais opostos. Daí segue portanto que

$$\int d^3r \left[\vec{E}(\vec{r}, t) \times \vec{A}(\vec{r}, t) \right] = 4\pi i c \sum_{\vec{k}} \left(\vec{\epsilon}_{\vec{k}1} \times \vec{\epsilon}_{\vec{k}2} \right) \left(a_{\vec{k}2}^\dagger a_{\vec{k}1} - a_{\vec{k}1}^\dagger a_{\vec{k}2} \right).$$

Devido ao caráter anti-simétrico do produto vetorial esse operador é de fato hermiteano, o que torna a adição do hermiteano conjugado supérflua, neste caso. Escolhendo os vetores de polarização de modo que se tenha

$$\vec{\epsilon}_{\vec{k}1} \times \vec{\epsilon}_{\vec{k}2} = \vec{\epsilon}_{-\vec{k}1} \times \vec{\epsilon}_{-\vec{k}2} = \hat{k} \equiv \frac{\vec{k}}{|\vec{k}|}$$

resulta finalmente que a componente intrínseca do momento angular do campo é uma quantidade independente do tempo (e portanto também uma constante do movimento) dada por

$$\vec{I}_{\text{rad}} = i\hbar \sum_{\vec{k}} \hat{k} \left(a_{\vec{k}2}^\dagger a_{\vec{k}1} - a_{\vec{k}1}^\dagger a_{\vec{k}2} \right). \quad (8.10)$$

Esse resultado mostra imediatamente que \vec{I} não é *diagonal* nos estados de polarização descritos pelos vetores $\vec{\epsilon}_{\vec{k}\alpha}$. De fato, os termos de \vec{I} aniquilam fótons de uma polarização linear dada e criam fótons linearmente polarizados numa direção perpendicular à da polarização do fóton aniquilado. Esse caráter não diagonal na polarização pode porém ser eliminado através da transformação

$$a_{\vec{k}1} = \frac{1}{\sqrt{2}} (b_{\vec{k}1} + b_{\vec{k}2}), \quad a_{\vec{k}2} = \frac{i}{\sqrt{2}} (b_{\vec{k}1} - b_{\vec{k}2})$$

que pode ser facilmente invertida, dando

$$b_{\vec{k}1} = \frac{1}{\sqrt{2}} (a_{\vec{k}1} - ia_{\vec{k}2}), \quad b_{\vec{k}2} = \frac{1}{\sqrt{2}} (a_{\vec{k}1} + ia_{\vec{k}2}).$$

Essa é uma transformação canônica, no sentido que as relações de comutação satisfeitas pelos novos operadores $b_\lambda, b_\lambda^\dagger$ são idênticas às que são satisfeitas pelos operadores $a_\lambda, a_\lambda^\dagger$, isto é

$$[b_\lambda, b_{\lambda'}^\dagger] = \delta_{\lambda\lambda'}, \quad [b_\lambda, b_{\lambda'}] = 0.$$

Em termos desses novos operadores a expressão (8.10) assume a forma

$$\vec{I}_{\text{rad}} = \hbar \sum_{\vec{k}} \hat{k} (b_{\vec{k}1}^\dagger b_{\vec{k}1} - b_{\vec{k}2}^\dagger b_{\vec{k}2}). \quad (8.11)$$

A expansão em ondas planas do potencial vetor $\vec{A}(\vec{r}, t)$ também pode ser expressa em termos de b_λ e b_λ^\dagger como

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}} \sqrt{\frac{4\pi\hbar c}{2kL^3}} \left[(\vec{\epsilon}_{\vec{k}+} b_{\vec{k}1} + \vec{\epsilon}_{\vec{k}-} b_{\vec{k}2}) e^{i(\vec{k}\cdot\vec{r} - \omega_k t)} + \text{h.c.} \right]$$

onde foram introduzidos os novos vetores de polarização

$$\vec{\epsilon}_{\vec{k}\pm} \equiv \frac{1}{\sqrt{2}} (\vec{\epsilon}_{\vec{k}1} \pm i\vec{\epsilon}_{\vec{k}2}).$$

Esses vetores de polarização complexos correspondem aos estados de polarização circular “à esquerda” (sinal +) e “à direita” (sinal −) da óptica ondulatória, respectivamente. A expressão de $\vec{A}(\vec{r}, t)$ mostra então que o operador $b_{\vec{k}1}^\dagger$ cria um fóton de momento $\hbar\vec{k}$ circularmente polarizado à direita, enquanto $b_{\vec{k}2}^\dagger$ cria um fóton circularmente polarizado à esquerda. De acordo com (8.11), cada fóton de momento $\hbar\vec{k}$ circularmente polarizado à esquerda contribui com $\hbar\hat{k}$ para \vec{I} , enquanto cada fóton de momento $\hbar\vec{k}$ circularmente polarizado à direita contribui $-\hbar\hat{k}$. A quantidade $\vec{I} \cdot \hat{k}/\hbar$ para um estado de um fóton é chamada *helicidade* do fóton, e portanto os estados de polarização circular à esquerda (direita) correspondem a estados de helicidade positiva, +1 (negativa, −1).

Como alternativa aos estados de Fock $|\{n_\lambda\}\rangle$ com fótons linearmente polarizados, eq. (8.7), é possível também usar estados de Fock construídos com fótons com polarização circular, com o que os operadores $a_\lambda, a_\lambda^\dagger$ são substituídos por $b_\lambda, b_\lambda^\dagger$ em (8.8). As propriedades da energia total e do momento total do campo com relação à polarização permitem escrever os operadores correspondentes simplesmente como

$$H_{\text{rad}} = \sum_{\lambda} \hbar\omega_{\lambda} b_{\lambda}^\dagger b_{\lambda} \quad \text{e} \quad \vec{P}_{\text{rad}} = \sum_{\lambda} \hbar\vec{k} b_{\lambda}^\dagger b_{\lambda}.$$

8.2 Interação com um sistema de partículas carregadas.

O acoplamento da versão descrita acima do campo de radiação quantizado com um sistema de partículas carregadas que funcionarão como fontes para o campo pode ser feito usando o procedimento já utilizado na seção 6.4. Formalmente, o acoplamento é implementado através da substituição

$$\vec{p}_i \longrightarrow \vec{p}_i - \frac{e_i}{c} \vec{A}(\vec{r}_i) \quad (8.12)$$

onde o operador \vec{p}_i é o momento canonicamente conjugado à posição \vec{r}_i da i -ésima partícula do sistema cuja carga é e_i e $\vec{A}(\vec{r})$ é o potencial vetor que descreve o campo de radiação quantizado na descrição de Schrödinger, obtida fazendo $t = 0$ na expressão (8.5), que corresponde à descrição de Heisenberg. O argumento de $\vec{A}(\vec{r})$ é identificado, no acoplamento, com o operador de posição \vec{r}_i da partícula considerada. Isso corresponde ao fato de que, classicamente, a interação entre o campo eletromagnético e uma carga puntiforme é dada em termos do valor do campo na posição da carga.

Uma situação bastante geral, que pode ser tomada como exemplo desse procedimento, é a de um sistema de muitas partículas, sujeitas a um potencial externo de um corpo $V(\vec{r}_i)$ (onde \vec{r}_i é o operador de posição da i -ésima partícula), e que interagem entre si através de um potencial de dois corpos $v(\vec{r}_i, \vec{r}_j)$. O hamiltoniano correspondente a esse sistema de partículas tem a forma

$$H_{\text{part}} = \sum_i \left(\frac{p_i^2}{2m_i} + V(\vec{r}_i) \right) + \frac{1}{2} \sum_{i,j} v(\vec{r}_i, \vec{r}_j)$$

e, em presença do campo de radiação, o hamiltoniano total a ser considerado é

$$H = H_{\text{rad}} + H_{\text{part}} + H_{\text{int}} \quad (8.13)$$

onde $H_{\text{part}} + H_{\text{int}}$ resultam da substituição (8.12) feita em H_{part} , que dá

$$\begin{aligned} H_{\text{part}} &\longrightarrow H_{\text{part}} + H_{\text{int}} = \sum_i \left[\frac{1}{2m_i} \left(\vec{p}_i - \frac{e_i}{c} \vec{A}(\vec{r}_i) \right)^2 + V(\vec{r}_i) \right] + \frac{1}{2} \sum_{i,j} v(\vec{r}_i, \vec{r}_j) \\ &= H_{\text{part}} - \sum_i \frac{e_i}{2m_i c} \left(\vec{p}_i \cdot \vec{A}(\vec{r}_i) + \vec{A}(\vec{r}_i) \cdot \vec{p}_i \right) + \sum_i \frac{e_i^2}{2m_i c^2} \vec{A}^2(\vec{r}_i). \end{aligned}$$

Como o potencial vetor correspondente ao padrão de Coulomb tem divergência zero, ele na realidade comuta com os momentos \vec{p}_i canonicamente conjugados às posições \vec{r}_i , de modo que ao hamiltoniano de interação pode ser escrito como

$$H_{\text{int}} = - \sum_i \frac{e_i}{m_i c} \vec{p}_i \cdot \vec{A}(\vec{r}_i) + \sum_i \frac{e_i^2}{2m_i c^2} \vec{A}^2(\vec{r}_i) \quad (8.14)$$

onde

$$\vec{A}(\vec{r}) = \sqrt{\frac{4\pi c^2}{L^3}} \sum_{\lambda} \vec{\epsilon}_{\lambda} q_{\lambda} e^{i\vec{k} \cdot \vec{r}} \quad \text{com} \quad q_{\lambda} = \sqrt{\frac{\hbar}{2\omega_{\lambda}}} (a_{\lambda} + a_{-\lambda}^{\dagger}). \quad (8.15)$$

A interação coulombiana entre as partículas carregadas é, nessa formulação, representada pelo potencial de dois corpos

$$v_c(\vec{r}_i, \vec{r}_j) = \frac{e_i e_j}{|\vec{r}_i - \vec{r}_j|}$$

que está supostamente incluído no termo de dois corpos de H_{part} .

Uma propriedade importante do hamiltoniano (8.13) é a de que ele *não é compatível* com o operador correspondente ao número total de fótons

$$N = \sum_{\lambda} a_{\lambda}^{\dagger} a_{\lambda}$$

isto é, $[H, N] \neq 0$. Isso significa que, em geral, os estados estacionários de H não serão autovetores de N , isto é, não conterão um número definido de fótons. Reciprocamente, estados com um número definido de fótons não serão, em geral, estacionários. Em particular, é fácil ver que estados da forma

$$|E^{(0)}; 0\rangle \equiv |\phi_E^{(0)}\rangle \otimes |0\rangle$$

onde $|\phi_E^{(0)}\rangle$ é um estado estacionário de H_{part} e $|0\rangle$ é o “vácuo” (estado com zero fótons) de H_{rad} , isto é

$$H_{\text{part}}|\phi_E^{(0)}\rangle = E|\phi_E^{(0)}\rangle \quad \text{e} \quad N|0\rangle = H_{\text{rad}}|E^{(0)}; 0\rangle = 0,$$

não são estacionários, pois

$$H|E^{(0)}; 0\rangle = (H_{\text{rad}} + H_{\text{part}} + H_{\text{int}})|E^{(0)}; 0\rangle = E|E^{(0)}; 0\rangle + H_{\text{int}}|E^{(0)}; 0\rangle$$

e o último termo contém componentes com zero, um e dois fótons. As componentes com um fóton resultam dos termos de H_{int} que são lineares em $\vec{A}(\vec{r}_i)$ e contém um operador de criação $a_{-\lambda}^{\dagger}$ (v. equações (8.14) e (8.15)), enquanto os termos com zero e dois fótons provém do termo quadrático em $\vec{A}(\vec{r}_i)$, que pode ser escrito explicitamente como

$$\sum_i \frac{e_1^2}{2m_i c^2} \vec{A}^2(\vec{r}_i) = \frac{4\pi c^2}{L^3} \sum_i \frac{e_1^2}{2m_i c^2} \sum_{\lambda, \lambda'} \vec{\epsilon}_\lambda \cdot \vec{\epsilon}_{\lambda'} e^{i(\vec{k} + \vec{k}') \cdot \vec{r}_i} \sqrt{\frac{\hbar^2}{4\omega_\lambda \omega_{\lambda'}}} (a_\lambda + a_{-\lambda}^\dagger) (a_{\lambda'} + a_{-\lambda'}^\dagger).$$

Essa expressão contém termos de quatro tipos, envolvendo respectivamente os operadores $a_\lambda a_{\lambda'}$, $a_{-\lambda}^\dagger a_{\lambda'}$, $a_\lambda a_{-\lambda'}^\dagger$ e $a_{-\lambda}^\dagger a_{-\lambda'}^\dagger$. Os dois primeiros aniquilam o vácuo $|0\rangle$, por conterem operadores de aniquilação agindo diretamente sobre esse estado. A ação do terceiro sobre $|0\rangle$ dá como resultado

$$a_\lambda a_{-\lambda'}^\dagger |0\rangle = ([a_\lambda, a_{-\lambda'}^\dagger] + a_{-\lambda'}^\dagger a_\lambda) |0\rangle = \delta_{\lambda, -\lambda'} |0\rangle,$$

onde foram usadas as relações de comutação satisfeitas pelos operadores de criação e de aniquilação. A “sobrevivência” do vácuo pode ser interpretada como resultante da criação de um fóton $-\lambda'$ seguida da aniquilação *desse mesmo fóton*, como indicado explicitamente pelo delta de Kronecker. Finalmente, da ação do quarto termo sobre o vácuo resulta um estado de dois fótons, que dá lugar à componente de dois fótons de $H_{\text{int}}|E^{(0)}; 0\rangle$.

Esse resultado mostra que os estados estacionários de H_{part} combinados com o vácuo de H_{rad} *deixam de ser estacionários* quando são levados em conta os efeitos dos termos de acoplamento H_{int} , devido à possibilidade de emissão de fótons, introduzida através desses termos. De uma forma inteiramente análoga é possível ver também que os estados estacionários de H_{part} combinados com estados de n (um ou mais) fótons de H_{rad} também não são estacionários, pois a ação de H_{int} sobre esses estados dá lugar a componentes não apenas com n fótons, mas também com $n - 2$ (para $n \geq 2$), $n \pm 1$ e $n + 2$ fótons. Nesses casos, além da emissão de fótons, pode ocorrer também a *absorção* de fótons através de H_{int} .

A dinâmica de sistemas descritos por hamiltonianos do tipo (8.13) não pode, em geral, ser tratada exatamente. Um procedimento típico seguido nesse contexto envolve a inclusão de efeitos devidos a H_{int} através de técnicas de teoria de perturbações, tomando como ponto de partida os estados estacionários de $H_0 = H_{\text{part}} + H_{\text{rad}}$. Ao contrário do que acontece com $H = H_0 + H_{\text{int}}$, H_0 é compatível não apenas com o operador correspondente ao número total de fótons N , mas também com todo o conjunto de operadores $N_\lambda = a_\lambda^\dagger a_\lambda$, de modo que os seus autovetores podem ser tomados como autovetores simultâneos de H_{part} e dos N_λ . Uma base conveniente de autoestados de H_0 pode de fato ser escrita usando o fato de que H_{part} e H_{rad} são compatíveis. Ela é a base produto formada por vetores de estado do tipo

$$|E_n^{(0)}; \{n_\lambda\}\rangle \equiv |\phi_{E_n}^{(0)}\rangle \otimes |\{n_\lambda\}\rangle; \quad H_0 |E_n^{(0)}; \{n_\lambda\}\rangle = \left(E_n + \sum_\lambda \hbar \omega_\lambda n_\lambda\right) |E_n^{(0)}; \{n_\lambda\}\rangle$$

onde os $|\phi_{E_n}^{(0)}\rangle$ são os estados estacionários de H_{part} e $|\{n_\lambda\}\rangle$ são os autoestados (8.7) de H_{rad} . O uso de técnicas de teoria de perturbação dependente do tempo permite calcular dessa forma

vários processos de interesse, entre os quais taxas de emissão, absorção e espalhamento de fótons. Alguns exemplos disso serão tratados nas seções seguintes.

A observação feita na seção 6.4 com relação à interação de propriedades eletromagnéticas associadas a spins com campos eletromagnéticos externos cabe igualmente no contexto da obtenção de H_{int} através da substituição (8.12). De fato, esse hamiltoniano de interação tão pouco inclui efeitos de interação com o campo de radiação de graus de propriedades eletromagnéticas associadas a graus liberdade sem análogo clássico, como é o caso dos momentos magnéticos associados aos spins. Para tratar esses efeitos é preciso adicionar explicitamente novos termos a H_{int} . Um tratamento que evita esse procedimento, de certa forma “ad hoc”, envolve o uso da equação de Dirac para tratar partículas de spin 1/2, e será discutido no capítulo 11.

8.2.1 Espalhamento Thomson.

Uma primeira aplicação do esquema geral montado na seção anterior consiste no espalhamento de fótons por uma partícula carregada (espalhamento Compton). Um caso particular especialmente simples ocorre no caso em que a frequência dos fótons incidentes é muito baixa, no sentido que o recuo da partícula espalhadora é desprezível, e o espalhamento é consequentemente elástico (espalhamento Thompson).

O processo geral a ser considerado envolve portanto transições que podem ser representadas como

$$\gamma + P \longrightarrow \gamma' + P$$

onde γ e γ' representam o fóton incidente e emergente, respectivamente, e P representa a partícula de massa M e carga Q . O hamiltoniano que descreve esse processo pode ser escrito como

$$H = H_{\text{rad}} + \frac{p^2}{2M} + H_{\text{int}}$$

onde o segundo termo corresponde à energia cinética da partícula e

$$H_{\text{int}} = -\frac{Q}{Mc} \vec{p} \cdot \vec{A} + \frac{Q^2}{2Mc^2} \vec{A}^2.$$

Tratando o problema no “sistema de laboratório”, no qual a partícula P tem inicialmente momento zero, as energias e momentos iniciais e finais são dados como no quadro abaixo.

$$\begin{array}{ccccccc}
\gamma & + & P & \longrightarrow & \gamma' & + & P \\
\hline
\text{Energia :} & E_i = \hbar k c & + & M c^2 & E_f = \hbar k' c & + & M c^2 + \frac{p'^2}{2M} \\
\text{Momento :} & \hbar \vec{k} & & 0 & \hbar \vec{k}' & & \vec{p}'
\end{array}$$

A energia final da partícula foi escrita na aproximação não relativística. Em particular, no limite do espalhamento Thompson, deve-se ter

$$\frac{p'^2}{2M} \ll \hbar k' c \simeq \hbar k c$$

ou seja, $|\vec{k}'| \simeq |\vec{k}|$, o que corresponde ao espalhamento elástico do fóton.

A seção de choque para esse processo de espalhamento é calculada, perturbativamente, através de uma aplicação direta da “regra áurea” para obter a probabilidade de transição por unidade de tempo entre o estado inicial e os estados finais $\{f\}$ apropriados, através do termo de interação H_{int} . Incluindo explicitamente termos de primeira e de segunda ordem em H_{int} , a expressão para a probabilidade de transição por unidade de tempo é

$$W = \frac{2\pi}{\hbar} \sum_{\{f\}} \left| \langle f | H_{\text{int}} | i \rangle + \sum_I \frac{\langle f | H_{\text{int}} | I \rangle \langle I | H_{\text{int}} | i \rangle}{E_i - E_I + i\eta} + \dots \right|^2 \delta(E_f - E_i)$$

Como o número de fótons no estado inicial e nos estados finais considerados é o mesmo, sendo o potencial vetor linear nos operadores de criação e de aniquilação de fótons, a_λ^\dagger e a_λ , é claro que o elemento de matriz

$$-\frac{Q}{Mc} \langle f | \vec{p} \cdot \vec{A} | i \rangle$$

é nulo para este processo. O termo de H_{int} envolvendo \vec{A}^2 , por outro lado, contém um termo proporcional a $a_\lambda^\dagger a_{\lambda'}$ e outro proporcional a $a_{\lambda'} a_\lambda^\dagger$, além de termos que criam e que aniquilam dois fótons. Os dois primeiros contribuem para o processo de espalhamento através do elemento de matriz

$$\frac{Q^2}{2Mc^2} \langle f | \vec{A}^2 | i \rangle. \quad (8.16)$$

No tratamento perturbativo de processos eletromagnéticos envolvendo cargas elétricas que são múltiplos pequenos da carga elementar, um parâmetro de expansão adequado é a constante

de estrutura fina $\alpha = e^2/\hbar c \simeq 1/137$, onde e representa a carga elementar. Supondo que essa restrição se aplique à carga Q da partícula espalhadora, e notando que a probabilidade de transição depende do quadrado do elemento de matriz de transição, resulta que a contribuição do elemento de matriz (8.16) dá contribuições de ordem α^2 para a probabilidade de transição. Para manter a consistência da expansão em potências de α é necessário no entanto considerar contribuições de segunda ordem no termo envolvendo $\vec{p} \cdot \vec{A}$, que são da mesma ordem em α que as contribuições de primeira ordem no termo envolvendo \vec{A}^2 . As amplitudes de segunda ordem relevantes nesse contexto são dadas por

$$\frac{Q^2}{M^2 c^2} \frac{\langle f | \vec{p} \cdot \vec{A} | I \rangle \langle I | \vec{p} \cdot \vec{A} | i \rangle}{E_i - E_I + i\eta}$$

onde $|I\rangle$ é um estado intermediário apropriado. Em geral, amplitudes deste tipo não são nulas quando $|I\rangle$ contém zero fótons ou dois fótons, dado que os estados inicial e final contém um fóton. Na particular situação considerada aqui, no entanto, o elemento de matriz que contém o estado inicial $|i\rangle$ é nulo, pois esse estado é um autoestado do momento \vec{p} com autovalor zero.

A referência explícita ao referencial de laboratório para o anulamento das amplitudes de segunda ordem acima pode à primeira vista parecer estranha, com o argumento de que é sempre possível utilizar o referencial em que a partícula espalhadora esteja inicialmente em repouso para calcular essas amplitudes, e que portanto elas devem se anular “sempre”. O argumento é de fato correto, e foi na realidade utilizado no cálculo apresentado, mas a conclusão tirada dele não é correta se “sempre” inclui o uso de outro referencial qualquer. É preciso ter em conta que a descrição do campo eletromagnético no padrão de Coulomb não é “manifestamente covariante”, estando ligada a um referencial definido, no sentido que, por exemplo, a condição de divergência nula que caracteriza esse padrão assume uma forma distinta quando se faz uma mudança de referencial (transformação de Lorentz) para outro em movimento com relação ao primeiro. O procedimento de quantização do campo de radiação adotado aqui faz uso explícito das condições que definem o padrão de Coulomb, e portanto deve ser entendido sempre em relação a um referencial definido. Desse modo, a escolha de um referencial apropriado deve *preceder* o processo de quantização. Desse modo, no cálculo apresentado, o potencial vetor é definido em termos do padrão de Coulomb *construído para o referencial em que a partícula espalhadora está inicialmente em repouso*. Os fótons resultantes da quantização utilizando o padrão de Coulomb são *diferentes* para escolhas também diferentes do referencial, e a amplitude que se anula é a que corresponde à criação ou aniquilação de fótons do referencial de repouso inicial da partícula através do termo $\vec{p} \cdot \vec{A}$.

O elemento de matriz relevante para o processo de espalhamento considerado em ordem perturbativa mais baixa é portanto dado por (8.16). Ele pode ser calculado escrevendo explicitamente os estados inicial e final como

$$|i\rangle \equiv |\vec{k}\vec{\epsilon}; \vec{p}_i = 0\rangle = a_{\vec{k}\vec{\epsilon}}^\dagger |0\rangle \frac{1}{\sqrt{L^3}}$$

$$|f\rangle \equiv |\vec{k}'\vec{\epsilon}'; \vec{p}_f = \vec{p}'\rangle = a_{\vec{k}'\vec{\epsilon}'}^\dagger |0\rangle \frac{1}{\sqrt{L^3}} e^{\frac{i}{\hbar} \vec{p}' \cdot \vec{r}}$$

onde $|0\rangle$ é o vácuo de fótons, e os estados da partícula foram escritos como autofunções do momento, usando condições de contorno periódicas no volume L^3 . Os vetores $\vec{\epsilon}$ e $\vec{\epsilon}'$ representam os estados de polarização do fóton inicial e do fóton final, respectivamente. Desse modo

$$\frac{Q^2}{2Mc^2} \langle f | \vec{A}^2 | i \rangle = \frac{Q^2}{2Mc^2} \frac{4\pi c^2}{L^3} \sum_{\lambda, \lambda'} \vec{\epsilon}_\lambda \cdot \vec{\epsilon}_{\lambda'} \left(\frac{1}{L^3} \int d^3r e^{-\frac{i}{\hbar} \vec{p}' \cdot \vec{r}} e^{i(\vec{k}_\lambda + \vec{k}_{\lambda'}) \cdot \vec{r}} \right) \langle 0 | a_{\vec{k}'\vec{\epsilon}'} q_\lambda q_{\lambda'} a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle.$$

A integral sobre \vec{r} pode ser feita imediatamente, dando uma função delta de Kronecker que impõe a conservação do momento total

$$\frac{1}{L^3} \int d^3r e^{-\frac{i}{\hbar} \vec{p}' \cdot \vec{r}} e^{i(\vec{k}_\lambda + \vec{k}_{\lambda'}) \cdot \vec{r}} = \delta_K \left(\vec{k}_\lambda + \vec{k}_{\lambda'} - \frac{\vec{p}'}{\hbar} \right)$$

enquanto, usando a definição

$$q_\lambda = \sqrt{\frac{\hbar}{2k_\lambda c}} (a_\lambda + a_{-\lambda}^\dagger),$$

o elemento de matriz no espaço de Fock dos fótons fica

$$\begin{aligned} \langle 0 | a_{\vec{k}'\vec{\epsilon}'} q_\lambda q_{\lambda'} a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle &= \frac{\hbar}{2c\sqrt{k_\lambda k_{\lambda'}}} \left[\langle 0 | a_{\vec{k}'\vec{\epsilon}'} a_\lambda a_{-\lambda'}^\dagger a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle + \langle 0 | a_{\vec{k}'\vec{\epsilon}'} a_{-\lambda}^\dagger a_{\lambda'} a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle \right. \\ &\quad \left. + \langle 0 | a_{\vec{k}'\vec{\epsilon}'} a_\lambda a_{\lambda'} a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle + \langle 0 | a_{\vec{k}'\vec{\epsilon}'} a_{-\lambda}^\dagger a_{-\lambda'}^\dagger a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle \right]. \end{aligned}$$

As duas últimas contribuições são claramente nulas, sendo valores esperados no vácuo de um número desigual de operadores de criação e de aniquilação. As duas primeiras são facilmente calculadas usando as relações de comutação desses operadores, dando

$$\langle 0 | a_{\vec{k}'\vec{\epsilon}'} q_\lambda q_{\lambda'} a_{\vec{k}\vec{\epsilon}}^\dagger | 0 \rangle = \frac{\hbar}{2c\sqrt{k_\lambda k_{\lambda'}}} \left(\delta_{\lambda, \vec{k}\vec{\epsilon}} \delta_{\lambda', -\vec{k}'\vec{\epsilon}'} + \delta_{\lambda, -\vec{k}'\vec{\epsilon}'} \delta_{\lambda', \vec{k}\vec{\epsilon}} \right).$$

Colecionando os vários resultados parciais a forma final da amplitude de transição para o processo de espalhamento (8.16) aparece como

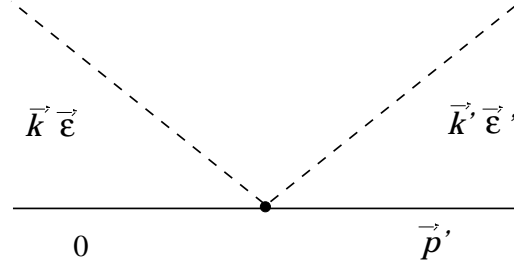


Figura 8.2: Diagrama “gaivota” para o espalhamento de fótons por uma partícula carregada.

$$\begin{aligned}
 M_g \equiv \frac{Q^2}{2Mc^2} \langle f | \vec{A}^2 | i \rangle &= \frac{Q^2}{2Mc^2} \frac{4\pi c^2}{L^3} 2\vec{\epsilon} \cdot \vec{\epsilon}' \delta_K \left(\vec{k} - \vec{k}' - \frac{\vec{p}'}{\hbar} \right) \frac{\hbar}{2c\sqrt{k k'}} \\
 &= \frac{2\pi\hbar c}{\sqrt{k k'}} \frac{Q^2}{Mc^2} \vec{\epsilon} \cdot \vec{\epsilon}' \frac{1}{L^3} \delta_K \left(\vec{k} - \vec{k}' - \frac{\vec{p}'}{\hbar} \right).
 \end{aligned}$$

Esse resultado pode ser interpretado como correspondendo à absorção do fóton inicial $\vec{k}, \vec{\epsilon}$ por um dos fatores \vec{A} com a emissão concomitante do fóton final $\vec{k}', \vec{\epsilon}'$ pelo outro fator \vec{A} . Como ambos esses fatores envolvem o mesmo valor de \vec{r} , esses processos de absorção e de emissão ocorrem “no mesmo ponto”, e a amplitude calculada pode ser representada pelo diagrama da figura 8.2 que, eventualmente por sua aparência, é às vezes chamado diagrama “gaivota” (sea gull). As duas partes em que é dividida a linha reta desse diagrama representam os estados inicial e final da partícula, as linhas tracejadas representam os fótons inicial e final e o vértice representa a ação do termo de interação envolvendo \vec{A}^2 .

O elemento de matriz “gaivota” M_g deve agora ser usado na regra áurea para calcular a probabilidade de transição por unidade de tempo como

$$W = \frac{2\pi}{\hbar} \sum_{\{f\}} \delta(E_f - E_i) |M_g|^2.$$

A especificação do conjunto de estados finais a serem incluídos na soma depende da especificação do processo cuja probabilidade se queira calcular. Para o cálculo da seção de choque diferencial de espalhamento do fóton incidente com polarização $\vec{\epsilon}$ para o elemento de ângulo sólido $d\Omega$ na direção final $\hat{k}' = \vec{k}'/k'$ sendo ainda a polarização do fóton final escolhida como $\vec{\epsilon}'$, a soma deve incluir todos os valores possíveis do módulo $k' = |\vec{k}'|$ dentro do elemento de ângulo sólido considerado, bem como todos os valores possíveis do momento final da partícula, \vec{p}' . O que se obtém desse modo é

$$dW = \frac{2\pi}{\hbar} \frac{1}{L^6} \sum_{k'} \sum_{\vec{p}'} \delta \left(\hbar k c - \hbar k' c - \frac{p'^2}{2M} \right) \delta_K \left(\vec{k} - \vec{k}' - \frac{\vec{p}'}{\hbar} \right) \frac{4\pi^2 \hbar^2 c^2}{k k'} \frac{Q^4}{M^2 c^4} (\vec{\epsilon}' \cdot \vec{\epsilon}')^2. \quad (8.17)$$

Essa expressão incorpora as condições de conservação de energia e momento na forma sugestiva das “propriedades corpusculares da luz”, que primeiro atraiu a atenção sobre o efeito Compton. Ela se simplifica no limite de fótons de baixa energia (espalhamento Thompson), no qual a energia de recuo da partícula $p'^2/2M$ é desprezível para os valores de \vec{p}' compatíveis com a conservação do momento. Isso permite desprezar o termo de energia cinética da partícula no delta de Dirac de conservação de energia, e fazer a soma sobre \vec{p}' usando o delta de Kronecker correspondente à conservação de momento, o que leva a

$$dW_{\text{Th}} = \frac{2\pi}{\hbar} \frac{1}{L^6} \sum_{k'} \delta(\hbar k c - \hbar k' c) \frac{4\pi^2 \hbar^2 Q^4}{M^2 c^2 k k'} (\vec{\epsilon}' \cdot \vec{\epsilon}')^2, \quad \frac{p'^2}{2M} \ll \hbar k c, \quad p' = \hbar |\vec{k} - \vec{k}'|. \quad (8.18)$$

Os passos seguintes na redução das expressões (8.17) ou (8.18) seguem os padrões usuais: substituição da(s) soma(s) sobre o(s) espectro(s) discreto(s) do(s) momento(s) resultante(s) das condições de contorno periódicas por integrais e mudança da variável de integração para o argumento da função delta de Dirac. Aplicando, por exemplo, esse procedimento à expressão (8.18), resulta

$$dW_{\text{Th}} = \frac{2\pi}{\hbar} \frac{1}{L^6} \frac{L^3}{(2\pi)^3} d\Omega \int_0^\infty dk' k'^2 \frac{1}{\hbar c} \delta(k - k') \frac{4\pi^2 \hbar^2 Q^4}{M^2 c^2 k k'} (\vec{\epsilon}' \cdot \vec{\epsilon}')^2 = \frac{d\Omega}{L^3} \frac{Q^4}{M^2 c^3} (\vec{\epsilon}' \cdot \vec{\epsilon}')^2$$

enquanto a seção de choque diferencial correspondente é

$$\frac{d\sigma_{\text{Th}}}{d\Omega} = \frac{1}{\text{fluxo incidente}} \frac{dW_{\text{Th}}}{d\Omega} = \frac{L^3}{c} \frac{dW_{\text{Th}}}{d\Omega} = \left(\frac{Q^2}{Mc^2} \right)^2 (\vec{\epsilon}' \cdot \vec{\epsilon}')^2. \quad (8.19)$$

O fluxo incidente de fótons foi calculado como o produto da densidade de fótons no estado inicial, $1/L^3$, pela velocidade associada aos fótons, c . A seção de choque Thompson diferencial, Eq. (8.19), resulta ser uma quantidade “clássica”, no sentido que nela não intervêm a constante de Planck \hbar . Na forma em que foi escrita, ela se refere ao caso do espalhamento de fótons *polarizados*, com polarização $\vec{\epsilon}$, dando fótons espalhados com polarização também definida $\vec{\epsilon}'$.

A dependência angular da seção de choque diferencial provém, na realidade, unicamente do produto escalar das polarizações inicial e final. A geometria envolvida nesse produto escalar pode ser melhor entendida fazendo referência à figura 8.3. O plano dessa figura coincide com o plano de espalhamento, isto é, com o plano definido pelos vetores de onda inicial e final, \vec{k} e

\vec{k}' . Para cada um desses vetores de onda, os dois vetores de polarização básicos são definidos como estando contidos nesse mesmo plano ($\vec{\epsilon}_1, \vec{\epsilon}_1'$) e perpendiculares a ele ($\vec{\epsilon}_2, \vec{\epsilon}_2' = \vec{\epsilon}_2$), respectivamente. Para um fóton incidente linearmente polarizado, o vetor de polarização inicial $\vec{\epsilon}$ estará em geral fora do plano de espalhamento, podendo ser expresso em termos do ângulo φ que ele faz com esse plano como

$$\vec{\epsilon} = \vec{\epsilon}_1 \cos \varphi + \vec{\epsilon}_2 \sin \varphi.$$

Analogamente, o vetor de polarização linear $\vec{\epsilon}'$ observado após o espalhamento (não incluído na figura) pode ser expandido como

$$\vec{\epsilon}' = \vec{\epsilon}_1' \cos \varphi' + \vec{\epsilon}_2' \sin \varphi' = \vec{\epsilon}_1' \cos \varphi' + \vec{\epsilon}_2 \sin \varphi'$$

onde φ' é o ângulo existente entre esse vetor e o plano de espalhamento. Tendo em conta que o ângulo entre os vetores básicos de polarização $\vec{\epsilon}_1$ e $\vec{\epsilon}_1'$ é o ângulo de espalhamento θ ,

$$\vec{\epsilon} \cdot \vec{\epsilon}' = \cos \varphi \cos \varphi' \vec{\epsilon}_1 \cdot \vec{\epsilon}_1' + \sin \varphi \sin \varphi' = \cos \varphi \cos \varphi' \cos \theta + \sin \varphi \sin \varphi'.$$

Quando o fóton incidente é linearmente polarizado numa direção $\vec{\epsilon}$ mas o detector do fóton espalhado não é sensível à polarização, a seção de choque relevante é a seção de choque calculada acima somada sobre uma base ortogonal qualquer de estados de polarização finais. Essa seção de choque *inclusiva* pode ser calculada, por exemplo, como

$$\frac{d\sigma_{\text{Th}}^{(\vec{\epsilon} \rightarrow \text{qqr. pol.})}}{d\Omega} = \sum_{i=1}^2 \frac{d\sigma_{\text{Th}}}{d\Omega} = \left(\frac{Q^2}{Mc^2} \right)^2 \sum_{i=1}^2 (\vec{\epsilon} \cdot \vec{\epsilon}_i')^2$$

Usando o fato de que (v. figura 8.3)

$$\vec{\epsilon} \cdot \vec{\epsilon}_1' = \cos \varphi \cos \theta \quad \text{e} \quad \vec{\epsilon} \cdot \vec{\epsilon}_2' = \sin \varphi$$

é fácil ver que

$$\frac{d\sigma_{\text{Th}}^{(\vec{\epsilon} \rightarrow \text{qqr. pol.})}}{d\Omega} = \left(\frac{Q^2}{Mc^2} \right)^2 (1 - \cos^2 \varphi \sin^2 \theta). \quad (8.20)$$

Um feixe incidente não polarizado, por outro lado, deve ser visto como uma mistura de dois feixes polarizados em direções ortogonais e com pesos iguais. A seção de choque para um feixe incidente de fótons não polarizado, no caso de um detector insensível à polarização final, corresponde à média de duas expressões como a (8.20) para duas polarizações iniciais ortogonais quaisquer. Escolhendo essas polarizações iniciais como $\vec{\epsilon}_1$ e $\vec{\epsilon}_2$ é fácil verificar que

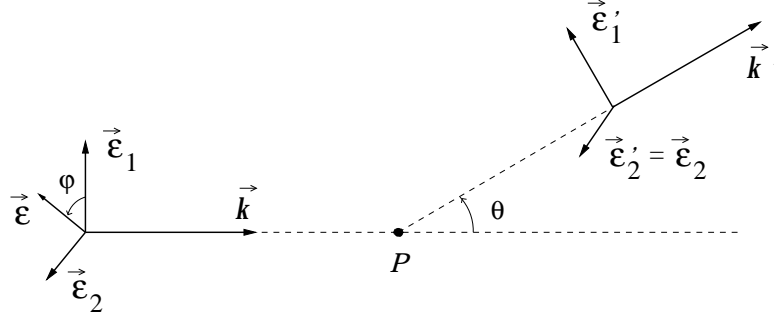


Figura 8.3: Vetores de polarização para o espalhamento Thompson.

$$\frac{d\sigma_{\text{Th}}^{(\text{n\~ao pol.} \rightarrow \text{qqr. pol.})}}{d\Omega} = \frac{1}{2} \left(\frac{Q^2}{Mc^2} \right)^2 (1 + \cos^2 \theta). \quad (8.21)$$

O fator $1/2$ nesta última expressão provém da média tomada sobre as duas polarizações iniciais. Como pode ser notado, a seção de choque (8.20), correspondente a fótons incidentes linearmente polarizados mas sem seleção da polarização final, depende não só do ângulo de espalhamento θ mas também do ângulo φ , entre a direção da polarização dos fótons e o plano de espalhamento. Uma outra forma de interpretar esta última dependência é ilustrada na figura 8.4. Aqui o plano da figura corresponde não mais ao plano de espalhamento mas ao “plano de incidência”, definido pelo momento do fóton incidente $\hbar\vec{k}$ e pelo seu vetor de polarização $\vec{\epsilon}$. O vetor de polarização quebra a simetria axial em torno da direção de incidência, resultando na dependência da seção de choque diferencial com φ , que agora pode ser reconhecido como o ângulo azimutal em que é observado o fóton emergente. No caso da seção de choque envolvendo fótons incidentes não polarizados a simetria em torno da direção de incidência é restabelecida, e a seção de choque diferencial depende apenas do ângulo de espalhamento θ .

Finalmente, a seção de choque *total* para o espalhamento Thompson, no caso de um feixe não polarizado e sem discriminação da polarização final, se obtém integrando a última expressão sobre todo o ângulo sólido. Portanto

$$\sigma_{\text{Th}} = \int \frac{d\sigma_{\text{Th}}^{(\text{n\~ao pol.} \rightarrow \text{qqr. pol.})}}{d\Omega} d\Omega = \frac{8\pi}{3} \left(\frac{Q^2}{Mc^2} \right)^2.$$

O valor dessa seção de choque envolve o comprimento

$$r_0 \equiv \frac{Q^2}{Mc^2}$$

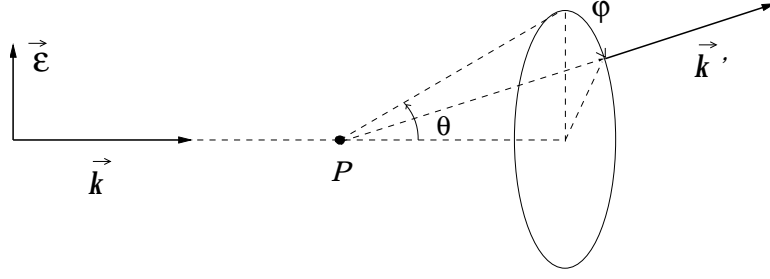


Figura 8.4: Diagrama ilustrando a dependência azimutal da seção de choque diferencial (8.20), em que os fótons incidentes são polarizados linearmente e a polarização final não é selecionada.

que dá uma medida da carga da partícula espalhadora, no caso em que sua massa (ou o valor da relação Q/M entre a carga e a massa) seja conhecida. O valor desse parâmetro para um elétron, em que Q e M são respectivamente a carga e a massa do elétron, é chamado *raio clássico do elétron*, e tem o valor $r_0 = 2.81785 \times 10^{-13}$ cm.

Convém observar explicitamente que a seção de choque inclusiva, sem observação de polarização final, e a seção de choque envolvendo feixe incidente não polarizado foram calculadas acima em termos de somas ou médias de *seções de choque* em que todas as polarizações são consideradas. Em outras palavras, as somas ou médias foram calculadas *no nível das probabilidades*, e não no nível das amplitudes quânticas referentes aos diferentes processos. Somas ou médias desse tipo são usualmente chamadas “incoerentes” para distingui-las da soma (dita “coerente”) de amplitudes quânticas. As probabilidades resultantes de somas coerentes diferem da soma incoerente das probabilidades correspondentes às várias amplitudes por conterem termos cruzados, ou de interferência, adicionais.

As somas incoerentes envolvidas no cálculo de seções de choque (ou probabilidades de transição) inclusivas podem ser entendidas no contexto da interpretação estatística usual da Mecânica Quântica, tal como exposta na seção 2.2.3. De fato, no caso de um sistema cuja dinâmica é regida por um hamiltoniano escrito como $H = H_0 + H_1$, a probabilidade de transição de um estado estacionário $|E_i\rangle$ de H_0 (no qual o sistema se encontra em $t = 0$) para um estado estado final $|E_f\rangle$, também estado estacionário de H_0 , no tempo t , é dada em geral e de forma *exata* pela expressão formal

$$T_{i \rightarrow f}(t) = |\langle E_f | U_I(t, 0) | E_i \rangle|^2 = \langle E_i | U_I^\dagger(t, 0) P_{E_f} U_I(t, 0) | E_i \rangle$$

onde $U_I(t, 0)$ é o operador de evolução na representação de interação, que satisfaz a equação (cf. Eq. (2.30))

$$i\hbar \frac{d}{dt} U_I(t, 0) = H_1^{(I)}(t) U_I(t, 0), \quad U_I(0, 0) = \hat{1}, \quad H_1^{(I)}(t) \equiv U_0^\dagger(t, 0) H_1 U_0(t, 0)$$

com $U_0(t, 0) = e^{-\frac{i}{\hbar}H_0t}$, e P_{E_f} é o operador de projeção

$$P_{E_f} \equiv |E_f\rangle\langle E_f|.$$

Caso o autovalor E_f de H_0 seja degenerado, a especificação unívoca do estado final considerado pode ser feita especificando também os autovalores de um conjunto completo de observáveis compatíveis com H_0 . Denotando esses parâmetros adicionais coletivamente por λ , isso corresponde a calcular uma probabilidade de transição $T_{i \rightarrow f, \lambda}(t)$ substituindo P_{E_f} pelo operador de projeção $P_{E_f, \lambda} \equiv |E_f \lambda\rangle\langle E_f \lambda|$. Além disso é possível considerar *probabilidades de transição inclusivas*, extendendo esse operador de projeção como no caso da equação (2.15), isto é

$$P_{E_f, \lambda} \rightarrow P_{E_f, \{\lambda\}} \equiv \sum_{\lambda} P_{E_f, \lambda} = \sum_{\lambda} |E_f \lambda\rangle\langle E_f \lambda|$$

onde foi utilizado o fato de que $\langle E_f \lambda | E_f \lambda' \rangle = \delta_{\lambda \lambda'}$, resultante de terem esses estados sido especificados pelos autovalores de um conjunto completo de observáveis *compatíveis*. A probabilidade de transição inclusiva resultante é então

$$T_{i \rightarrow f, \{\lambda\}}(t) = \langle E_i | U_I^\dagger(t, 0) P_{E_f, \{\lambda\}} U_I(t, 0) | E_i \rangle = \sum_{\lambda} T_{i \rightarrow f, \lambda}(t) \quad (8.22)$$

isto é, a soma incoerente das probabilidades de transição para cada um dos estados $|E_f \lambda\rangle$.

A ortogonalidade dos estados envolvidos no cálculo da probabilidade inclusiva, decorrente de serem eles autovetores simultâneos de um conjunto completo de observáveis compatíveis, foi essencial para que se obtenha esse resultado. Por outro lado, se

$$|E_f, \gamma\rangle = \sum_{\lambda} c_{\lambda} |E_f \lambda\rangle$$

é um autoestado de um observável g , compatível com H_0 mas não com os observáveis utilizados para definir os estados $|E_f \lambda\rangle$, a probabilidade de transição para esse estado será dada por

$$T_{i \rightarrow f, \gamma}(t) = |\langle E_f, \gamma | U_I^\dagger(t, 0) | E_i \rangle|^2 = \left| \sum_{\lambda} c_{\lambda}^* \langle E_f, \gamma | U_I^\dagger(t, 0) | E_i \rangle \right|^2$$

que envolve uma soma *coerente* das amplitudes associadas aos estados $|E_f \lambda\rangle$ com os coeficientes c_{λ} . Essa probabilidade não pode, portanto ser obtida a partir das probabilidades não inclusivas $T_{i \rightarrow f, \lambda}(t)$ apenas, pois nela intervêm também termos de interferência entre as diferentes amplitudes. Isso mostra que o que se pode calcular variando a extensão da soma incoerente sobre estados finais são na realidade probabilidades de transição mais

ou menos inclusivas *com relação aos autovalores do conjunto completo de observáveis compatíveis utilizado para especificar completamente cada um dos estados finais*, apenas.

A média sobre estados iniciais envolvida, por exemplo, no cálculo de seções de choque para feixes não polarizados, pode ser tratada formalmente de modo idêntico a êsse, embora sua interpretação seja algo diferente. De fato, essa média envolve, em particular, uma soma incoerente sobre diferentes estados *iniciais* especificados por um conjunto completo de observáveis compatíveis com H_0 , no caso em que o autovalor E_i seja por sua vez degenerado, o que confere um caráter “inclusivo” ao estado inicial, já que a probabilidade correspondente pode ser escrita, em analogia com a equação (8.22),

$$T_{i,\{\mu\} \rightarrow f} = \frac{1}{n_\mu} \langle E_f | U_I(t, 0) P_{E_i, \{\mu\}} U_I^\dagger(t, 0) | E_f \rangle$$

sendo $P_{E_i, \{\mu\}}$ o operador de projeção inclusivo sobre os estados iniciais especificados por E_i e por um conjunto completo de números quânticos adicionais associados a observáveis compatíveis, denotados coletivamente por μ , e sendo n_μ o número de estados incluídos no projetor. Na versão inclusiva, a representação do estado inicial é portanto estendida de

$$P_{E_i, \mu} \equiv |E_i \mu\rangle \langle E_i \mu|$$

para

$$\frac{1}{n_\mu} P_{E_i, \{\mu\}} = \frac{1}{n_\mu} \sum_\mu |E_i \mu\rangle \langle E_i \mu|.$$

Essas expressões representam o estado inicial através de sua *matriz densidade*, definida como na página 147. No primeiro caso, o estado inicial pode igualmente ser caracterizado pelo vetor de estado $|E_i \mu\rangle$ e corresponde por isso ao que se chama um estado “puro”, ao passo que no segundo caso a representação através de um único vetor de estado não existe, como no caso encontrado anteriormente na página 279. Este tipo de estado é aqui interpretado como descrevendo uma coleção de sistemas idênticos cujos estados iniciais são preparados em estados (por exemplo, de polarização) diferentes, mas com pesos iguais. A constante comum $1/n_\mu$ representa portanto a probabilidade “a priori” de que o estado inicial de um dos sistemas seja um dado estado $|E_i \mu\rangle$. Em geral, as probabilidades associadas a diferentes estados iniciais na coleção de sistemas considerada não precisam ser iguais. Estados iniciais “não puros” menos particulares podem portanto ser representados por matrizes densidade do tipo $\sum_\mu |E_i \mu\rangle p_\mu \langle E_i \mu|$ onde as probabilidades “a priori” são dadas pelos números positivos p_μ que devem satisfazer à condição de normalização $\sum_\mu p_\mu = 1$. Quando μ rotula diferentes estados de polarização, estados deste último tipo descrevem feixes *parcialmente polarizados*.

As probabilidades de transição calculadas em teoria de perturbações como Wt , onde W é a probabilidade de transição por unidade de tempo dada pela regra áurea com opções

apropriadas com relação à soma sobre estados finais e média sobre estados iniciais, nada mais são que versões aproximadas de expressões como as consideradas acima, às quais se aplicam portanto todas as considerações feitas com relação a essas somas e médias. O alcance dessas considerações não é, no entanto, limitado a cálculos de tipo perturbativo. No caso específico das seções de choque Thompson calculadas, os rótulos utilizados para identificar univocamente estados iniciais e finais, correspondentes a λ e μ acima, são, nos dois casos, o vetor de onda \vec{k} (ou o momento $\hbar\vec{k}$) do fóton e seu estado de polarização. Dessa forma, a matriz densidade que descreve o feixe incidente polarizado de fótons é $|\vec{k}\vec{\epsilon}\rangle\langle\vec{k}\vec{\epsilon}|$; o feixe incidente não polarizado corresponde, por outro lado, à matriz densidade

$$\text{feixe não polarizado} \longrightarrow \frac{1}{2} \left(|\vec{k}\vec{\epsilon}_1\rangle\langle\vec{k}\vec{\epsilon}_1| + |\vec{k}\vec{\epsilon}_2\rangle\langle\vec{k}\vec{\epsilon}_2| \right)$$

onde $\vec{\epsilon}_1$ e $\vec{\epsilon}_2$ representam um par qualquer de estados de polarização ortogonais do fóton de momento $\hbar\vec{k}$. No caso de um feixe polarizado numa direção $\vec{\epsilon}$ diferente desses estados de base escolhidos, a seção de choque será dada em termos de uma soma *coerente* de amplitudes que tenham sido calculadas com os estados iniciais de polarização $\vec{\epsilon}_1$ e $\vec{\epsilon}_2$, com coeficientes α_1 e α_2 tais que $\vec{\epsilon} = \alpha_1\vec{\epsilon}_1 + \alpha_2\vec{\epsilon}_2$.

8.2.2 Radiação de freiamento (Bremsstrahlung).

A amplitude para um processo em que um elétron interage com o campo de radiação dando lugar à criação de um fóton é em geral diferente de zero, na ordem mais baixa de teoria de perturbação, através do termo envolvendo $\vec{p} \cdot \vec{A}$. De fato, o elemento de matriz desse termo entre um estado em que o elétron tem momento \vec{p} e no qual não há fótons, e outro estado em que o elétron tem momento \vec{p}' e há um fóton com momento $\hbar\vec{k}$ e polarização $\vec{\epsilon}$ pode ser facilmente calculado como sendo

$$\langle \vec{p}'; \vec{k}\vec{\epsilon} | - \frac{e}{m_e c} \vec{p} \cdot \vec{A} | \vec{p}; 0 \rangle = - \frac{e}{m_e c} \sqrt{\frac{4\pi\hbar c}{2kL^3}} \vec{\epsilon} \cdot \vec{p} \delta_K \left(\vec{k} + \frac{\vec{p}'}{\hbar} - \frac{\vec{p}}{\hbar} \right).$$

Utilizando as mesmas convenções da figura 8.2, esse elemento de matriz pode ser representado pelo diagrama da figura 8.5. O delta de Kronecker que aparece como último fator representa a conservação do momento total

$$\hbar\vec{k} + \vec{p}' = \vec{p},$$

e isso implica no anulamento da probabilidade de transição por unidade de tempo calculada através da regra áurea, pois essa relação é incompatível com a condição de conservação de energia

$$\hbar kc + \frac{p'^2}{2m_e} = \frac{p^2}{2m_e}.$$

Desse modo, embora estando associada a uma amplitude não nula, a emissão de um fóton por um eléctron que interage apenas com o campo de radiação é excluída pelo anulamento do espaço de fase para esse processo.

A emissão do fóton se torna possível, no entanto, quando o eléctron interage também com outra partícula. Nesse caso a outra partícula pode também trocar momento com o eléctron, do que resulta um espaço de fase não nulo para a emissão do fóton. O mesmo ocorre no caso em que o eléctron interage com um campo externo, que pode ser visto como o limite do caso anterior quando a massa da outra partícula é muito maior que a massa do eléctron, de modo que o seu recuo pode ser desprezado. Devido ao papel essencial desempenhado pela outra partícula (ou pelo campo externo) nesse processo de emissão, ele é conhecido como radiação de freiamento (Bremsstrahlung), e pode ser representado como

$$e + Z \longrightarrow e' + Z' + \gamma.$$

A partícula adicional é representada por Z . Sua massa M pode eventualmente ser muito maior que a massa do eléctron. Tipicamente ela pode ser um núcleo de carga Ze que interage com o eléctron através de seu campo coulombiano. Os parâmetros cinemáticos relevantes para o cálculo perturbativo do processo de radiação de freiamento estão resumidos no quadro abaixo. Nesse quadro, por simplicidade o eléctron e a partícula Z são tratados em termos de uma cinemática não relativística, o que é permissível quando $p^2/2m_e \ll m_e c^2 = 510 \text{ KeV}$ e $M > m_e$. O spin das partículas é ignorado.

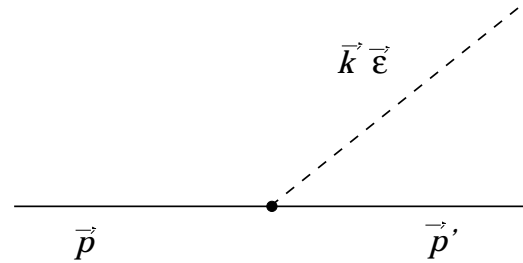


Figura 8.5: Diagrama que representa o processo de emissão de um fóton por um eléctron de momento \vec{p} .

	e	$+$	Z	\longrightarrow	e'	$+$	Z'	$+$	γ
<hr/>									
Energia :	$E_i = \frac{p^2}{2m_e}$	$+$	0		$E_f = \frac{p'^2}{2m_e}$	$+$	$\frac{P^2}{2M}$	$+$	$\hbar ck$
Momento :	\vec{p}		0		\vec{p}'		\vec{P}		$\hbar \vec{k}$
<hr/>									

Uma outra forma de encarar esse processo, a qual é relevante para o desenvolvimento que segue, é considerá-lo como um espalhamento inelástico do elétron pela partícula Z , acompanhado da emissão de um fóton. Ainda por simplicidade, a massa de Z será tomada como sendo muito maior que m_e . Nessas condições, a energia de recuo $P^2/2M$ é desprezível em comparação com os outros termos de E_f para todos os valores de \vec{P} compatíveis com a conservação de momento. Isso permite ignorar a energia cinética de Z e descrever a dinâmica do processo através do hamiltoniano

$$H = H_{\text{rad}} + \frac{p^2}{2m_e} - V_Z(\vec{r}) + \frac{e}{m_e c} \vec{p} \cdot \vec{A} + \frac{e^2}{2m_e c^2} \vec{A}^2. \quad (8.23)$$

O potencial $V_Z(\vec{r})$ descreve a interação de Z com o elétron, e será tomado como o potencial coulombiano blindado

$$V_Z(\vec{r}) = -\frac{Ze^2}{r} e^{-\mu r}$$

com vistas ao seu tratamento na aproximação de Born, fazendo o parâmetro μ tender finalmente a zero (cf. tratamento da página 208). Na realidade, seria possível partir do hamiltoniano “não perturbado”

$$H'_0 = H_{\text{rad}} + \frac{p^2}{2m_e} - V_Z(\vec{r})$$

e tratar perturbativamente apenas os termos de interação com o campo de radiação. Para evitar a complicação técnica do uso das funções coulombianas, que são as autofunções de H'_0 , a opção feita a seguir parte do hamiltoniano não perturbado livre

$$H_0 = H_{\text{rad}} + \frac{p^2}{2m_e}$$

e trata perturbativamente também a interação $V_z(\vec{r})$, isto é

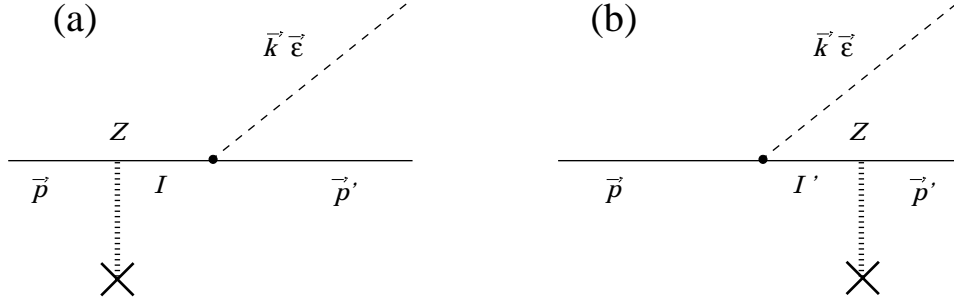


Figura 8.6: Diagramas que representam as amplitudes de segunda ordem que contribuem para a radiação de freiamto.

$$H_{\text{int}} = -\frac{Ze^2}{r}e^{-\mu r} + \frac{e}{2m_e c}\vec{p} \cdot \vec{A} + \frac{e^2}{2m_e c^2}\vec{A}^2.$$

Tendo em vista o carácter essencial do potencial $V_Z(\vec{r})$ para o processo de emissão de radiação, é fácil ver que as amplitudes de ordem mais baixa que contribuem para ele são de segunda ordem, e envolvem uma interação do tipo $\vec{p} \cdot \vec{A}$ além da interação do elétron com Z . Elas podem ser classificadas em dois tipos segundo a natureza dos estados intermediários envolvidos, e correspondem aos dois diagramas da figura 8.6. Em (a) a interação com Z , representada pela linha hachureada vertical, precede a interação $\vec{p} \cdot \vec{A}$ responsável pela emissão do fóton. Conseqüentemente, os estados intermediários I contém o elétron num estado de momento \vec{p}_I e zero fótons. Em (b), por outro lado, a criação do fóton precede a interação com Z , e os estados intermediários I' contém o fóton com momento $\hbar\vec{k}$ e polarização $\vec{\epsilon}$ além do elétron com momento $\vec{p}_{I'}$.

A probabilidade de transição por unidade de tempo é dada pela regra áurea como

$$W = \frac{2\pi}{\hbar} \sum_{\{f\}} \delta(E_f - E_i) \frac{Z^2 e^6}{m_e^2 c^2} \left| \sum_I \frac{\langle f | \vec{p} \cdot \vec{A} | I \rangle \langle I | \frac{e^{-\mu r}}{r} | i \rangle}{E_i - E_I + i\eta} + \sum_{I'} \frac{\langle f | \frac{e^{-\mu r}}{r} | I' \rangle \langle I' | \vec{p} \cdot \vec{A} | i \rangle}{E_i - E_{I'} + i\eta} \right|^2. \quad (8.24)$$

Os elementos de matriz que aparecem nessa expressão podem ser calculados facilmente. Os estados envolvidos são, explicitamente,

$$|i\rangle = |0\rangle \frac{1}{\sqrt{L^3}} e^{i\frac{\vec{p}}{\hbar} \cdot \vec{r}}, \quad |I\rangle = |0\rangle \frac{1}{\sqrt{L^3}} e^{i\frac{\vec{p}_I}{\hbar} \cdot \vec{r}}, \quad |I'\rangle = |\vec{k}\vec{\epsilon}\rangle \frac{1}{\sqrt{L^3}} e^{i\frac{\vec{p}_{I'}}{\hbar} \cdot \vec{r}} \quad \text{e} \quad |f\rangle = |\vec{k}\vec{\epsilon}\rangle \frac{1}{\sqrt{L^3}} e^{i\frac{\vec{p}'}{\hbar} \cdot \vec{r}}.$$

Portanto

$$\langle I' | \vec{p} \cdot \vec{A} | i \rangle = \sqrt{\frac{4\pi\hbar c}{2kL^3}} \vec{\epsilon} \cdot \vec{p} \delta_K \left(\vec{k} + \frac{\vec{p}_{I'}}{\hbar} - \frac{\vec{p}}{\hbar} \right)$$

e

$$\langle f | \vec{p} \cdot \vec{A} | I \rangle = \sqrt{\frac{4\pi\hbar c}{2kL^3}} \vec{\epsilon} \cdot \vec{p}_I \delta_K \left(\vec{k} + \frac{\vec{p}'}{\hbar} - \frac{\vec{p}_I}{\hbar} \right).$$

Por outro lado

$$\lim_{\mu \rightarrow 0} \langle I | \frac{e^{-\mu r}}{r} | i \rangle = \lim_{\mu \rightarrow 0} \frac{1}{L^3} \int_{L^3} d^3r e^{\frac{i}{\hbar}(\vec{p}_I - \vec{p}) \cdot \vec{r}} \frac{e^{-\mu r}}{r} = \frac{4\pi}{L^3} \frac{\hbar^2}{(\vec{p}_I - \vec{p})^2}$$

e, de uma forma inteiramente análoga,

$$\lim_{\mu \rightarrow 0} \langle f | \frac{e^{-\mu r}}{r} | I' \rangle = \frac{4\pi}{L^3} \frac{\hbar^2}{(\vec{p}' - \vec{p}_{I'})^2}.$$

Usando esses resultados juntamente com os denominadores de energia

$$E_i - E_I + i\eta = \frac{p^2}{2m_e} - \frac{p_I^2}{2m_e} + i\eta \quad \text{e} \quad E_i - E_{I'} + i\eta = \frac{p^2}{2m_e} - \frac{p_{I'}^2}{2m_e} - \hbar kc + i\eta$$

as amplitudes de segunda ordem que aparecem na regra áurea, correspondentes respectivamente aos processos (a) e (b) da figura 8.6, podem ser escritas sob a forma

$$\sum_I \frac{\langle f | \vec{p} \cdot \vec{A} | I \rangle \langle I | \frac{e^{-\mu r}}{r} | i \rangle}{E_i - E_I + i\eta} = \frac{4\pi}{L^{9/2}} \sqrt{\frac{4\pi\hbar c}{2k}} \sum_I \vec{\epsilon} \cdot \vec{p}_I \frac{\hbar^2}{(\vec{p}_I - \vec{p})^2} \frac{2m_e}{p^2 - p_I^2 + i\eta} \delta_K \left(\vec{k} + \frac{\vec{p}'}{\hbar} - \frac{\vec{p}_I}{\hbar} \right)$$

e

$$\begin{aligned} \sum_{I'} \frac{\langle f | \frac{e^{-\mu r}}{r} | I' \rangle \langle I' | \vec{p} \cdot \vec{A} | i \rangle}{E_i - E_{I'} + i\eta} &= \frac{4\pi}{L^{9/2}} \sqrt{\frac{4\pi\hbar c}{2k}} \vec{\epsilon} \cdot \vec{p} \times \\ &\times \sum_{I'} \frac{\hbar^2}{(\vec{p}' - \vec{p}_{I'})^2} \frac{2m_e}{p^2 - p_{I'}^2 - 2m_e\hbar kc + i\eta} \delta_K \left(\vec{k} + \frac{\vec{p}_{I'}}{\hbar} - \frac{\vec{p}}{\hbar} \right). \end{aligned}$$

Como os deltas de Kronecker efetivamente reduzem as somas sobre estados intermediários a um único termo, a probabilidade de transição por unidade de tempo (8.24) fica reduzida a

$$W = \frac{(4\pi)^4}{L^9} \frac{Z^2 \hbar^4 e^6}{kc} \sum_{\{f\}} \delta(\hbar kc + \frac{p'^2}{2m_e} - \frac{p^2}{2m_e}) |M|^2 \quad (8.25)$$

com

$$M = \frac{1}{(\hbar \vec{k} + \vec{p}' - \vec{p})^2} \left(\frac{\vec{\epsilon} \cdot (\hbar \vec{k} + \vec{p}')}{p^2 - (\hbar \vec{k} + \vec{p}')^2 + i\eta} + \frac{\vec{\epsilon} \cdot \vec{p}}{p^2 - (\hbar \vec{k} - \vec{p})^2 - 2m_e \hbar kc + i\eta} \right).$$

Desse modo o objeto $|M|^2$ que aparece em (8.25) é dado na realidade como a soma de três termos com dependências diferentes dos parâmetros cinemáticos finais \vec{p}' , \vec{k} e $\vec{\epsilon}$.

A expressão (8.25) para W pode ser utilizada para obter várias quantidades físicas de interesse através de diferentes restrições impostas sobre a soma sobre estados finais. De fato, medindo a probabilidade de transição por unidade de tempo em unidades do fluxo incidente de elétrons, $v/L^3 = p/m_e L^3$, o que se obtém é uma seção de choque

$$d\sigma = \frac{m_e L^3}{p} \frac{(4\pi)^4}{L^9} \frac{Z^2 \hbar^4 e^6}{kc} \sum_{\{\vec{p}'\}} \sum_{\{\vec{k}\}} \sum_{\{\vec{\epsilon}\}} \delta(\hbar kc + \frac{p'^2}{2m_e} - \frac{p^2}{2m_e}) |M|^2$$

da qual é possível obter, restringindo ou não de forma apropriada as somas sobre \vec{p}' , \vec{k} e $\vec{\epsilon}$, a distribuição angular de fótons com ou sem seleção da polarização, a distribuição angular dos elétrons emergentes, seções de choque duplamente diferenciais que dão a correlação angular entre fótons e elétrons emergentes, etc. Enquanto as quantidades mais inclusivas levam a resultados de visualização mais simples, no sentido de dependerem de um número menor de variáveis cinemáticas finais, o seu cálculo envolve por outro lado de um número maior de integrações para as quais pode ser necessário (ou pelo menos mais conveniente) recorrer a técnicas numéricas. Uma dessas quantidades inclusivas é a seção de choque $d\sigma/dk$, que dá a distribuição em energia dos fótons emitidos, independentemente de sua polarização e da sua direção de emissão. A sua obtenção envolve uma integração tri-dimensional sobre o momento final do elétron e uma integração bi-dimensional sobre os ângulos do vetor de onda do fóton, \hat{k} , além da soma sobre os estados de polarização dos fótons. Transformando, da forma usual, as somas sobre \vec{p}' e \vec{k} que aparecem na última expressão acima em integrais se obtém

$$\frac{d\sigma}{dk} = \frac{(4\pi)^4}{L^6} \frac{Z^2 \hbar^4 e^6 m_e}{p k c} \frac{L^6}{(2\pi)^6} \int \frac{d^3 p'}{\hbar^3} \int d^2 \hat{k} \, k^2 \, \delta(\hbar kc + \frac{p'^2}{2m_e} - \frac{p^2}{2m_e}) \sum_{\{\vec{\epsilon}\}} |M|^2.$$

A integração sobre o módulo de \vec{p}' pode ser feita imediatamente usando a função delta de Dirac de conservação de energia, o que leva a

$$\frac{d\sigma}{dk} = \frac{4Z^2 e^6 m_e^3 \hbar k}{\pi^2 p} \sqrt{\left(\frac{p}{m_e c}\right)^2 - \frac{2\hbar k}{m_e c}} \int d^2 \hat{p}' \int d^2 \hat{k} \sum_{\vec{\epsilon}} |M|_{p'^2=p^2+2m_e \hbar k c}^2$$

onde a notação introduzida no último fator indica que $|M|^2$ deve ser calculado com a substituição

$$\vec{p}' \longrightarrow \hat{p}' \sqrt{p^2 + 2m_e \hbar k c}.$$

Este último resultado envolve ainda quatro integrais angulares, além da soma sobre polarizações. A dependência do integrando com as variáveis angulares a serem integradas aparece através de termos contendo $\vec{p} \cdot \vec{p}'$, $\vec{p} \cdot \hat{k}$ e $\hat{k} \cdot \vec{p}'$, e a integração deve ser feita numericamente. O comportamento qualitativo de $d\sigma/dk$ pode no entanto ser inferido do fato de que é possível verificar que a função de k definida por

$$X(k) \equiv m_e^4 c^4 \hbar^2 k^2 \int d^2 \hat{p}' \int d^2 \hat{k} \sum_{\vec{\epsilon}} |M|_{p'^2=p^2+2m_e \hbar k c}^2$$

se aproxima de um valor *finito* quando $k \rightarrow 0$, de modo que o resultado das integrais com a soma sobre as polarizações se comporta como k^{-2} nesse limite. Os fatores $m_e^4 c^4 \hbar^2$ foram introduzidos com o objetivo de tornar a função $X(k)$ adimensional. Usando essa definição, a seção de choque $d\sigma/dk$ pode ser escrita

$$\begin{aligned} \frac{d\sigma}{dk} &= \frac{4Z^2 e^6 m_e^3 \hbar}{\pi^2 p} \frac{1}{m_e^4 c^4 \hbar^2} \sqrt{\left(\frac{p}{m_e c}\right)^2 - \frac{2\hbar k}{m_e c}} \frac{X(k)}{k} \\ &= \frac{4Z^2}{\pi^2} \left(\frac{e^2}{\hbar c}\right) \left(\frac{e^2}{m_e c^2}\right)^2 \left(\frac{m_e c}{p}\right) \sqrt{\left(\frac{p}{m_e c}\right)^2 - \frac{2\hbar k}{m_e c}} \frac{X(k)}{k} \\ &= \frac{4Z^2}{\pi^2} \alpha r_0^2 \sqrt{1 - \frac{\hbar k c}{\frac{p^2}{2m_e}}} \frac{X(k)}{k} \end{aligned}$$

onde $\alpha = e^2/\hbar c$ é a constante de estrutura fina e $r_0 = e^2/m_e c^2$ é o raio clássico do elétron. Essa seção de choque portanto se anula quando a energia $\hbar k c$ do fóton se torna igual à energia cinética do elétron incidente (pois nesse limite o fator envolvendo a raiz quadrada se anula) e diverge como k^{-1} quando $k \rightarrow 0$.

A divergência de $d\sigma/dk$ para $k \rightarrow 0$ implica que a seção de choque total a que está sujeito o feixe incidente de elétrons, que é dada por

$$\sigma = \int_0^{\frac{p^2}{2m_e\hbar c}} \frac{d\sigma}{dk} dk, \quad (8.26)$$

diverge logaritmicamente no limite inferior da integração sobre k . Essa divergência, chamada divergência, ou “catástrofe” infra-vermelha (por ocorrer no limite em que $k \rightarrow 0$), não é no entanto real, mas é um artefato espúrio do cálculo perturbativo da probabilidade de emissão de um fóton, como argumentado já em 1937 por Bloch e Nordsieck². De fato, para que o resultado do cálculo perturbativo que leva a (8.26) possa ser relevante, é preciso que a probabilidade de emissão de dois fótons seja muito menor que a probabilidade de emissão de apenas um fóton, a probabilidade de emissão de três fótons muito menor que a de dois, etc. Uma estimativa das razões dessas probabilidades é no entanto $\alpha(p/m_e c)^2 \log(p^2/2m_e\hbar kc)$, que é a ordem de grandeza da razão da seção de choque com emissão de um fóton de vetor de onda k pela seção de choque correspondente sem a emissão de nenhum fóton. Essa quantidade deixa de ser pequena quando $k \rightarrow 0$. O modo padrão de consertar o tratamento perturbativo nesse limite consiste em calcular a seção de choque de espalhamento *sem emissão de fótons* até a mesma ordem perturbativa $Z^2 e^6$, o que envolve os processos representados na figura 8.7. O primeiro desses processos é simplesmente a amplitude de Born, proporcional a Ze^2 , correspondente à interação Z , enquanto o segundo, que leva a uma amplitude proporcional a Ze^4 , envolve a emissão de um fóton através do termo em $\vec{p} \cdot \vec{A}$ antes da interação Z , e a posterior reabsorção do mesmo fóton através de uma nova interação $\vec{p} \cdot \vec{A}$. O termo de Born de Z dá uma contribuição para a seção de choque proporcional a $Z^2 e^4$, mas o *termo cruzado* das duas amplitudes da figura 8.7 dão uma contribuição proporcional a $Z^2 e^6$. A contribuição para a seção de choque sem emissão de fótons que envolve o quadrado da segunda amplitude é proporcional a $Z^2 e^8$ e deve ser ignorada por ser de ordem superior à considerada no cálculo da seção de choque com emissão de fótons. A contribuição do termo cruzado é também divergente para $k \rightarrow 0$ de forma que a soma das seções de choque com e sem a emissão de um fóton não mais diverge. A soma das duas seções de choque é entendida como a “seção de choque observável”, no sentido que um processo com a emissão de um fóton de energia muito baixa é efetivamente indistinguível dos processos sem emissão de fótons representados figura 8.7. É possível mostrar também que cancelamentos deste tipo ocorrem em todas as ordens de teoria de perturbações.

²F. Bloch e A. Nordsieck, *Phys. Rev.* **52**, 54 (1937). Esse trabalho está incluído na coletânea *Selected Papers on Quantum Electrodynamics*, Julian Schwinger (ed.), Dover, N.Y., 1958, pag. 129. Tratamentos progressivamente mais modernos dessa questão (utilizando uma formulação covariante) se encontram, por exemplo, em F. Mandl e G. Shaw, *Quantum Field Theory*, John Wiley & Sons (1986) Seção 9.7 e em S. Weinberg, *The Quantum Theory of Fields*, Vol. I, Cambridge University Press (1995), Cap. 13.

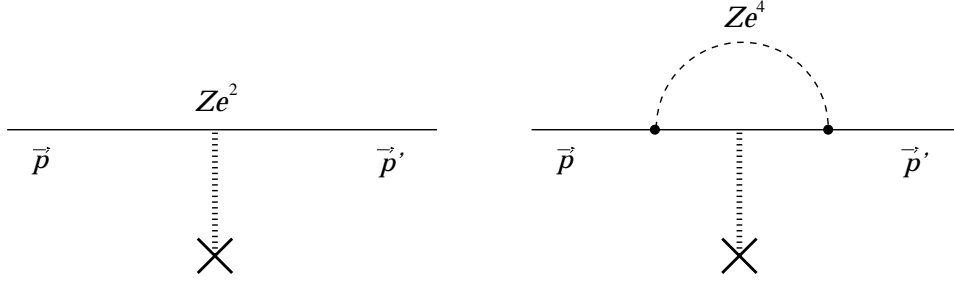


Figura 8.7: Processos sem emissão de fótons que cancelam a divergência de (8.26).

8.2.3 Transições radiativas em sistemas ligados.

Nas duas seções anteriores os sistemas carregados cuja interação com o campo de radiação foi tratada em termos de teoria de perturbações são sistemas “simples”, descritos como constituídos de uma única partícula. No caso da radiação de freiamento foi necessário, a rigor, introduzir na descrição também a interação dessa partícula “ativa” com uma outra partícula, afim de identificar processos cujas probabilidades não se anulem por razões cinemáticas, isto é, ligadas essencialmente a argumentos de espaço de fases. A simplicidade do tratamento foi porém mantida notando, em primeiro lugar, a equivalência entre a interação com a partícula adicional e a interação com um potencial externo, no caso em que essa partícula adicional tenha massa suficientemente grande quando comparada com a massa da partícula ativa; e, em segundo lugar, tratando os efeitos do potencial externo equivalente também de forma perturbativa, o que reduz a dinâmica não perturbada da partícula ativa a uma dinâmica de partícula livre.

O tratamento *não perturbativo* da partícula ativa com o potencial externo que representa a segunda partícula pesada na questão da radiação de freiamento, cuja possibilidade foi mencionada mas descartada na seção anterior, tipifica o problema a ser considerado agora. Em particular, no caso em que a partícula ativa é um elétron e a partícula pesada é um núcleo atômico de dimensões desprezíveis e carga positiva Ze , a dinâmica não perturbada da partícula ativa deixa de ser a dinâmica de uma partícula livre e se torna a dinâmica de um átomo hidrogenóide, com seu espectro parcialmente discreto e parcialmente contínuo. Transições perturbativas entre dois estados do espectro contínuo com emissão de fótons descrevem então a radiação de freiamento que acompanha o espalhamento coulombiano dos elétrons. O tratamento da seção anterior é na realidade um tratamento simplificado deste problema, no qual a interação coulombiana é tratada na aproximação de Born, apenas. O tratamento não perturbativo do potencial externo leva no entanto à possibilidade de processos que não podem ser descritos dessa forma, por envolverem estados do espectro *discreto* do átomo hidrogenóide, para os quais a aproximação de Born não é relevante. Esses processos novos incluem a *captura radiativa* do elétron, que

pode ser vista como um processo de radiação de freiamento em que o elétron, inicialmente em um estado não perturbado do espectro contínuo, sofre uma transição radiativa para um estado final ligado, pertencente à parte discreta do espectro não perturbado, e ainda as *transições radiativas entre diferentes estados do espectro discreto*. Da mesma forma que no caso da captura radiativa, é possível estabelecer relações de afinidade também destes processos com o da radiação de freiamento, embora envolvendo agora apenas estados ligados do elétron no potencial coulombiano. Como foi discutido no capítulo 1, este foi um dos problemas-chave para a elaboração da Mecânica Quântica em sua forma atual por Heisenberg e Schrödinger há agora três quartos de século, e é o que será tratado a seguir.

Para isso será considerado novamente o hamiltoniano (8.23) usado para tratar o processo de radiação de freiamento

$$H = H_{\text{rad}} + \frac{p^2}{2m_e} - \frac{Ze^2}{r} + \frac{e}{m_e c} \vec{p} \cdot \vec{A} + \frac{e^2}{2m_e c^2} \vec{A}^2.$$

onde agora, no entanto, apenas os dois últimos termos

$$H_{\text{int}} = \frac{e}{m_e c} \vec{p} \cdot \vec{A} + \frac{e^2}{2m_e c^2} \vec{A}^2$$

serão tratados perturbativamente. Como pode ser notado, para escrever esse hamiltoniano foram ignorados efeitos da massa finita do núcleo bem como os spins das partículas. Desse modo as funções de onda que representam os estados estacionários não perturbados do elétron são autofunções do hamiltoniano que descreve o átomo hidrogenóide

$$H_p = \frac{p^2}{2m_e} - \frac{Ze^2}{r}$$

cujos espectro discreto foi obtido na seção 5.2. As autofunções do espectro discreto foram escritas na equação (5.37), e serão aqui abreviadas como

$$\psi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\hat{r}), \quad l < n, \quad n \geq 1, \quad n, l \text{ inteiros}$$

onde as funções radiais $R_{nl}(r)$ são normalizadas de modo que

$$\int_0^\infty dr \, r^2 R_{nl}^2(r) = 1$$

e os $Y_{lm}(\hat{r})$ são harmônicas esféricas dos ângulos do vetor \vec{r} . Os autovalores de H_p associados a essas autofunções dependem apenas do número quântico n e são dados por

$$E_n = -\frac{Z^2 e^4 m_e}{2\hbar^2 n^2}.$$

A escala da extensão espacial dessas autofunções pode ser caracterizada pelo primeiro raio de Bohr para o átomo hidrogenóide

$$a_0^{(Z)} = \frac{a_0}{Z} = \frac{1}{Z} \frac{\hbar^2}{m_e e^2}.$$

Uma transição radiativa entre dois estados ligados com emissão de um fóton envolve portanto um estado inicial $|0\rangle\psi_{nlm}(\vec{r}) \equiv |nlm; 0\rangle$, onde $|0\rangle$ é o vácuo de fótons, e um estado final $|\vec{k}\vec{\epsilon}\rangle\psi_{n'l'm'}(\vec{r}) \equiv |n'l'm'; \vec{k}\vec{\epsilon}\rangle$, onde $|\vec{k}\vec{\epsilon}\rangle$ é um estado contendo um fóton de momento $\hbar\vec{k}$ e polarização $\vec{\epsilon}$ (v. figura 8.8). As energias não perturbadas associadas ao estado inicial e ao estado final são então, respectivamente

$$E_i = E_n \quad \text{e} \quad E_f = \hbar kc + E_{n'}$$

e a probabilidade de transição por unidade de tempo devida aos termos de interação do átomo com a radiação é dada pela regra áurea, em ordem mais baixa de H_{int} , como

$$W = \frac{2\pi}{\hbar} \sum_{\{f\}} \left| \langle n'l'm'; \vec{k}\vec{\epsilon} | H_{\text{int}} | nlm; 0 \rangle \right|^2 \delta(E_f - E_i). \quad (8.27)$$

Como o estado final contém um fóton mais que o estado inicial, apenas os termos lineares no potencial vetor contribuem para os elementos de matriz que aparecem nessa expressão. De fato, o termo envolvendo \vec{A}^2 pode apenas, seja não alterar o número de fótons, seja criar ou aniquilar *dois* fótons. Desse modo

$$\langle n'l'm'; \vec{k}\vec{\epsilon} | H_{\text{int}} | nlm; 0 \rangle = -\frac{e}{m_e c} \langle n'l'm'; \vec{k}\vec{\epsilon} | \vec{p} \cdot \vec{A} | nlm; 0 \rangle =$$

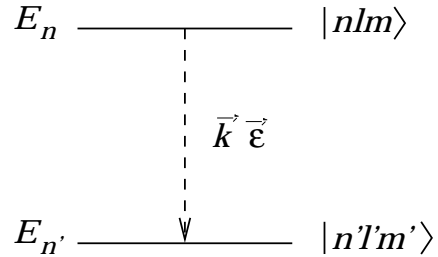


Figura 8.8: Transição do estado ligado $|nlm\rangle$ para o estado ligado $|n'l'm'\rangle$ com emissão de um fóton de momento $\hbar\vec{k}$ e polarização $\vec{\epsilon}$.

$$\begin{aligned}
&= -\frac{e}{m_e c} \sqrt{\frac{4\pi\hbar c}{2kL^3}} \int d^3r \psi_{n'l'm'}^*(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{p} \psi_{nlm}(\vec{r}) = \\
&= -e \sqrt{\frac{4\pi\hbar c}{2kL^3}} \vec{\epsilon} \cdot \vec{J}_{nlm, n'l'm'}(\vec{k})
\end{aligned} \tag{8.28}$$

com

$$\vec{J}_{nlm, n'l'm'}(\vec{k}) \equiv \int d^3r e^{-i\vec{k}\cdot\vec{r}} \psi_{n'l'm'}^*(\vec{r}) \frac{\vec{p}}{m_e c} \psi_{nlm}(\vec{r}). \tag{8.29}$$

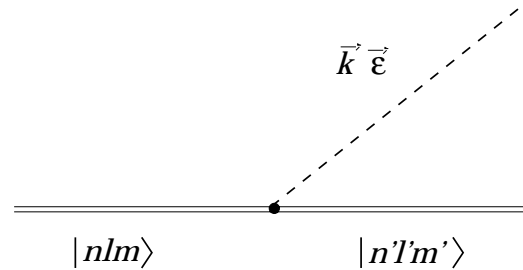
Note que o denominador $m_e c$ incluído em (8.29) torna esse objeto adimensional. O elemento de matriz (8.28) pode ser representado como na figura 8.9, onde as linhas duplas representam os estados ligados inicial $|nlm\rangle$ e final $|n'l'm'\rangle$ do elétron, respectivamente. Ele corresponde a um processo que difere daquele representado na figura 8.5 apenas pela natureza dos estados inicial e final do elétron. Essa diferença tem, no entanto, conseqüências radicais, pois como agora os estados inicial e final *não* são autovetores do momento \vec{p} , a integração sobre \vec{r} não mais conduz ao delta de Kronecker de conservação do momento total obtido naquele caso, mas à quantidade $\vec{J}_{nlm, n'l'm'}(\vec{k})$, que pode ser vista como a componente de Fourier \vec{k} da *densidade de corrente de transição*

$$\vec{J}_{nlm, n'l'm'}(\vec{r}) \equiv \psi_{n'l'm'}^*(\vec{r}) \frac{\vec{p}}{m_e c} \psi_{nlm}(\vec{r}).$$

Em conseqüência disso, o conflito cinemático que leva ao anulamento do espaço de fases para a emissão de um fóton no caso do processo representado pela figura 8.5 deixa de existir, e a probabilidade de transição por unidade de tempo (8.27) é em geral diferente de zero.

Como nos outros casos tratados até aqui, diferentes quantidades de interesse podem ser obtidas de (8.27) adotando diferentes tipos de restrição para a soma sobre estados finais. Em particular, a probabilidade de transição total se obtém somando sobre todos os momentos e polarizações do fóton emitido. Transformando a soma sobre momentos em uma integral da

Figura 8.9: Diagrama que representa a transição de um elétron atômico do estado $|nlm\rangle$ para o estado $|n'l'm'\rangle$, acompanhada da emissão de um fóton de momento $\hbar\vec{k}$ e polarização $\vec{\epsilon}$.



forma usual e explicitando as energias inicial e final no delta de Dirac da regra áurea, isso leva a

$$W = \frac{2\pi}{\hbar} \frac{L^3}{(2\pi)^3} \int d^3k \frac{4\pi e^2 \hbar}{2kcL^3} \delta(E_{n'} + \hbar kc - E_n) \left| \vec{\epsilon} \cdot \vec{J}_{nlm, n'l'm'}(\vec{k}) \right|^2.$$

A distribuição angular da radiação emitida com uma dada polarização, por outro lado, é obtida fixando a polarização desejada e a direção de emissão do fóton $\hat{k} \equiv \vec{k}/k$. A integração sobre o módulo de \vec{k} é portanto a única que deve ser feita nesse caso. Ela pode ser feita imediatamente utilizando a função delta de conservação de energia, com o resultado

$$\begin{aligned} \frac{dW_{\vec{k}\vec{\epsilon}}}{d\Omega_k} &= \frac{1}{2\pi} \left(\frac{e^2}{\hbar c} \right) \frac{E_n - E_{n'}}{\hbar} \left| \vec{\epsilon} \cdot \vec{J}_{nlm, n'l'm'} \left(\frac{E_n - E_{n'}}{\hbar c} \hat{k} \right) \right|^2 = \\ &= \frac{\alpha}{2\pi} \frac{\Delta E}{\hbar} \left| \vec{\epsilon} \cdot \vec{J}_{nlm, n'l'm'} \left(\frac{\Delta E}{\hbar c} \hat{k} \right) \right|^2 \end{aligned} \quad (8.30)$$

onde α é a constante de estrutura fina e $\Delta E = E_n - E_{n'}$ é a diferença de energia entre os dois estados ligados envolvidos na transição (v. figura 8.8). Esse resultado mostra explicitamente que a energia do fóton emitido é dada por $\hbar kc = \hbar \omega = E_n - E_{n'}$. As transições radiativas entre estados do espectro discreto do átomo leva portanto a um “espectro de linhas” de fótons com frequências bem definidas, determinadas pelo espaçamento em energia dos estados do espectro discreto do átomo.

É importante lembrar que relação $\hbar \omega = E_n - E_{n'}$ que fixa a frequência dos fótons emitidos na transição $|nlm\rangle \rightarrow |n'l'm'\rangle$ foi obtida usando a aproximação que consiste em ignorar a energia de recuo do núcleo, tratado a rigor simplesmente como um potencial externo. Quando isso não é feito, essa energia de recuo deve ser incluída como parte da energia final, o que faz com que na realidade $\hbar \omega$ seja ligeiramente *menor* que a diferença de energia entre os estados ligados envolvidos. A energia δE perdida quando o núcleo de massa M recua com momento P pode ser estimada simplesmente usando $P \simeq \hbar k$, o que dá

$$\delta E = \frac{P^2}{2M} \simeq \frac{(\hbar kc)^2}{2Mc^2}$$

que, para $\hbar kc \simeq \Delta E$ da ordem de eV, corresponde a $\delta E \sim 10^{-9}$ eV para M da ordem da massa do próton. Esse valor é comparável com a incerteza na energia ΔE devida à vida média finita dos estados atômicos, a qual pode ser estimada como $\hbar/\tau \simeq \hbar W$ onde W é a probabilidade de transição por unidade de tempo calculada acima, que resulta da ordem de 10^7 s^{-1} ($\tau \sim 10^{-7} \text{ s}$).

Resta, finalmente, tratar do cálculo de (8.29), que constitui o ingrediente dinâmico central da probabilidade da transição radiativa no átomo, dada pela expressão (8.30). Como os estados inicial e final têm aqui um *momento angular* bem definido, uma expansão útil consiste em escrever a onda plana que aparece no integrando também em termos de autofunções do momento angular. Isso será tratado de forma mais completa na seção seguinte, incluindo na expansão o vetor de polarização que aparece também em (8.30). Um procedimento mais simples, que é no entanto suficiente em muitos casos, consiste em notar que os valores do argumento $\vec{k} \cdot \vec{r}$ da exponencial que contribuem significativamente para o valor da integral são limitados pela extensão espacial dos estados ligados envolvidos. Esta pode ser estimada em termos do raio médio $\langle r \rangle_{nl}$ para o estado de números quânticos $\{n, l, m\}$ que foi obtido na seção 5.2.1 como

$$\langle r \rangle_{nl} = \frac{a_0}{2Z} [3n^2 - l(l+1)],$$

onde a_0 é o primeiro raio de Bohr. Em particular, para os estados circulares, para os quais $l = n - 1$, $\langle r \rangle_{n, l=n-1} = (2n^2 + n)a_0/2Z$. Por outro lado, a diferença de energia $E_n - E_{n-1}$ define o módulo do vetor de onda do fóton emitido nessa transição como

$$k = \frac{E_n - E_{n-1}}{\hbar c} = \frac{Z^2 R}{\hbar c} \frac{2n - 1}{n^2(n - 1)^2}$$

de modo que, para transições desse tipo os maiores valores de kr são da ordem de

$$k \langle r \rangle_{nl} \sim \frac{Z a_0 R}{2\hbar c} \frac{(2n - 1)(2n^2 + n)}{n^2(n - 1)^2} = \frac{Z\alpha}{4} \frac{4n^2 - 1}{n(n - 1)^2}$$

onde $\alpha = 1/137$ é a constante de estrutura fina. Para a transição $n = 2 \rightarrow n = 1$ essa quantidade é portanto da ordem de $Z\alpha$, e *diminui* à medida que n aumenta. Isso significa que, para transições desse tipo (ou para qualquer outra transição para a qual também se tenha $k \langle r \rangle_{nl} \ll 1$) e valores de Z pequenos com relação a 137, o comprimento de onda da radiação emitida é muito maior que a extensão espacial dos estados atômicos envolvidos, sugerindo que a exponencial no integrando de (8.29) possa ser simplesmente substituída pelo primeiro termo de sua expansão em série de potências, isto é, $e^{-ikr} \simeq 1$. Por razões que se tornarão claras em seguida, essa aproximação é chamada a *aproximação de dipolo elétrico*³, na qual a densidade de corrente de transição é dada simplesmente por

³Mais precisamente, essa aproximação corresponde ao *limite de grandes comprimentos de onda* da amplitude de dipolo elétrico, que pode ser extraída independentemente dessa aproximação, como mostrado na seção seguinte. Também na seção seguinte serão identificados na expansão da onda plana em autofunções do momento angular os outros multipolos elétricos e magnéticos. É usual designar a transição de 2^L -polo elétrico por EL , bem como a de 2^L -polo magnético por ML . Desse modo o rótulo convencional para a componente de dipolo elétrico é $E1$.

$$\vec{J}_{nlm, n'l'm'}^{E1} = \int d^3r \psi_{n'l'm'}^*(\vec{r}) \frac{\vec{p}}{m_e c} \psi_{nlm}(\vec{r}).$$

De fato, essa expressão pode ser re-escrita notando que

$$\frac{\vec{p}}{m_e c} = \frac{i}{\hbar c} [H_p, \vec{r}]$$

com o que, dado que $\psi_{nlm}(\vec{r})$ e $\psi_{n'l'm'}(\vec{r})$ são autofunções de H_p , ela adquire a forma

$$\begin{aligned} \vec{J}_{nlm, n'l'm'}^{E1} &= \frac{i}{\hbar c} \int d^3r \psi_{n'l'm'}^*(\vec{r}) [H_p, \vec{r}] \psi_{nlm}(\vec{r}) = \\ &= i \frac{E_{n'} - E_n}{\hbar c} \int d^3r \psi_{n'l'm'}^*(\vec{r}) \vec{r} \psi_{nlm}(\vec{r}) \equiv -i \frac{\Delta E}{\hbar c} \vec{d}_{nlm, n'l'm'} \end{aligned}$$

onde $\vec{d}_{nlm, n'l'm'}$ é (a menos da carga do elétron, e) o *momento de dipolo elétrico de transição* (cf. eq. (6.28)). A probabilidade por unidade de tempo da transição radiativa é dada portanto, na aproximação de dipolo elétrico, por

$$\frac{dW_{k\vec{\epsilon}}^{E1}}{d\Omega_k} = \frac{\alpha}{2\pi} \frac{\Delta E^3}{\hbar^3 c^2} \left| \vec{\epsilon} \cdot \vec{d}_{nlm, n'l'm'} \right|^2.$$

Esta expressão é essencialmente a que resulta do tratamento de Heisenberg do problema da intensidade das linhas espectrais do átomo de hidrogênio (v. Capítulo 1, páginas 17 e 23). Note que a probabilidade de transição por unidade de tempo W se relaciona com a expressão da página 17 por

$$W = \frac{\delta\epsilon}{\Delta E}$$

onde $\delta\epsilon$ é a energia radiada por unidade de tempo e $\Delta E = \hbar k c = \hbar \omega$ é a energia da transição. É de certa forma curioso que o passo crucial para a invenção da mecânica quântica teve como ponto de apoio um tratamento perturbativo do problema da radiação, na aproximação de dipolo elétrico. Longe de dificultarem o passo, essas limitações efetivamente expuseram para um tratamento mais adequado as variáveis de posição, diante da inadequação de sua interpretação clássica.

Existem situações em que a probabilidade de transição na aproximação de dipolo elétrico dW^{E1} se anula identicamente, devido a propriedades de simetria dos estados inicial e final

envolvidos no cálculo do momento de dipolo de transição. O conjunto de condições que esses estados devem satisfazer para que transições de dipolo elétrico sejam possíveis constitui as “regras de seleção” para esse tipo de transição.

Uma primeira regra de seleção envolve a paridade dos estados inicial e final. De fato, as funções de onda ψ_{nlm} têm paridades bem definidas, isto é

$$\psi_{nlm}(-\vec{r}) = (-1)^l \psi_{nlm}(\vec{r})$$

e, em consequência disso, a integral

$$\vec{d}_{nlm, n'l'm'} = \int d^3r \psi_{n'l'm'}^*(\vec{r}) \vec{r} \psi_{nlm}(\vec{r})$$

se anula identicamente sempre que as duas funções de onda tiverem a mesma paridade (pois nesse caso o integrando é uma função ímpar devido ao fator \vec{r}). Transições de dipolo elétrico só são possíveis, portanto, entre estados de paridades *diferentes*.

Há também regras de seleção relativas ao momento angular dos estados inicial e final, que podem ser obtidas diretamente (embora não da forma mais geral, v. nesse sentido o tratamento da seção seguinte) explicitando a dependência angular dos vários ingredientes de que dependem as amplitudes de transição $\vec{\epsilon} \cdot \vec{d}_{nlm, n'l'm'}$. Para isso convém tratar separadamente a emissão de fótons com polarização circular e com polarização linear, respectivamente.

O caso da emissão de um fóton circularmente polarizado pode ser tratado escolhendo o sistema de eixos representado na figura 8.10, no qual $\hat{z} = \hat{k}$, isto é, o eixo z aponta na direção de emissão do fóton. Usando os vetores de polarização circular

$$\vec{\epsilon}_{\pm} = \frac{1}{\sqrt{2}}(\hat{x} \pm i\hat{y})$$

as amplitudes de transição são proporcionais a

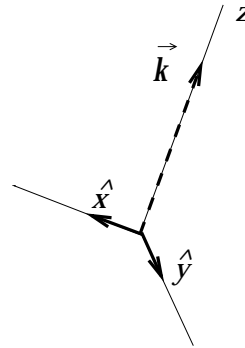


Figura 8.10: Sistema de referência para a emissão de um fóton circularmente polarizado.

$$\begin{aligned}
\vec{\epsilon}_{\pm} \cdot \vec{J}_{nlm, n'l'm'}^{E1} &= \int d^3r \psi_{n'l'm'}^*(\vec{r}) \vec{\epsilon}_{\pm} \cdot \vec{r} \psi_{nlm}(\vec{r}) = \\
&= \int d^3r \psi_{n'l'm'}^*(\vec{r}) \frac{\hat{x} \pm i\hat{y}}{\sqrt{2}} \psi_{nlm}(\vec{r}) = \\
&= \mp \sqrt{\frac{4\pi}{3}} \int d^3r \psi_{n'l'm'}^*(\vec{r}) r Y_{1,\pm 1}(\hat{r}) \psi_{nlm}(\vec{r})
\end{aligned}$$

onde $Y_{1,\pm 1}(\hat{r})$ é uma harmônica esférica. Como a parte angular das funções de onda são também harmônicas esféricas, resulta que as amplitudes de transição são proporcionais às integrais angulares

$$\int d\hat{r} Y_{l'm'}^*(\hat{r}) Y_{1,\pm 1}(\hat{r}) Y_{lm}(\hat{r}) = \sqrt{\frac{3}{4\pi}} \sqrt{\frac{2l+1}{2l'+1}} C_{0\ 0\ 0}^{l\ 1\ l'} C_{m\ \pm 1\ m'}^l \quad (8.31)$$

onde os dois últimos fatores são coeficientes de Clebsch-Gordan. A integral envolvendo três harmônicas esféricas pode ser calculada sem dificuldade fazendo uso da identidade

$$Y_{l_1 m_1}(\hat{r}) Y_{l_2 m_2}(\hat{r}) \equiv \sum_{LM} \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2L+1)}} C_{0\ 0\ 0}^{l_1\ l_2\ L} C_{m_1\ m_2\ M}^{l_1\ l_2\ L} Y_{LM}(\hat{r})$$

e das propriedades de ortonormalidade das harmônicas esféricas. As regras de seleção de momento angular para esse tipo de transição estão contidas nos coeficientes de Clebsch-Gordan que aparecem em (8.31). De fato, as condições para o não anulamento desses coeficientes são, respectivamente

$$\text{a) } C_{0\ 0\ 0}^{l\ 1\ l'} \neq 0 \quad \Longleftrightarrow \quad \begin{cases} |l-1| \leq l' \leq l+1 \\ l+1+l' \text{ par,} \end{cases}$$

$$\text{b) } C_{m\ \pm 1\ m'}^l \neq 0 \quad \Longleftrightarrow \quad \begin{cases} |l-1| \leq l' \leq l+1 \\ m' = m \pm 1. \end{cases}$$

As condições a) na realidade excluem a possibilidade de que se tenha $l = l'$, pois nesse caso $l+1+l'$ é ímpar. As transições de dipolo elétrico só podem ocorrer, portanto, para

$$|l - l'| = 1.$$

Essa regra de seleção inclui a regra de seleção de paridade, dado que a paridade de ψ_{nlm} é $(-1)^l$. Além disso, as condições b) exigem ainda que

$$m' = m \pm 1$$

onde o duplo sinal se refere aos dois estados possíveis de polarização circular, e o eixo z com relação ao qual são definidos os números quânticos m e m' coincide com a direção de emissão do fóton circularmente polarizado.

Para tratar a emissão de fótons polarizados linearmente o sistema de coordenadas mais conveniente é o da figura 8.11, em que agora o eixo z coincide com a direção do vetor de polarização $\vec{\epsilon}$ e a direção de emissão do fóton é identificada com o eixo x . Neste caso

$$\vec{\epsilon} \cdot \vec{r} = z = \sqrt{\frac{4\pi}{3}} Y_{10}(\hat{r})$$

e a integral angular que aparece na amplitude de transição é

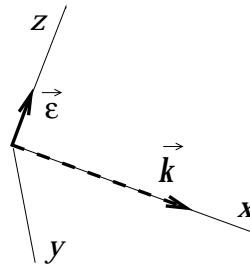
$$\int d\hat{r} Y_{l'm'}^*(\hat{r}) Y_{10}(\hat{r}) Y_{lm}(\hat{r}) = \sqrt{\frac{3}{4\pi}} \sqrt{\frac{2l_1}{2l' + 1}} C_{0 \ 0 \ 0}^{l \ 1 \ l'} C_{m \ 0 \ m'}^{l \ 1 \ l'}$$

que será diferente de zero somente quando estiverem satisfeitas as condições

$$|l - l'| = 1 \quad \text{e} \quad m' = m.$$

Em resumo, transições de dipolo elétrico só podem ocorrer entre estados cujos momentos angulares difiram por uma unidade. O número quântico magnético m , por outro lado, poderá variar de uma unidade, implicando nesse caso na emissão de um fóton circularmente polarizado, ou se manter constante, caso em que o fóton emitido tem polarização linear. A aplicação dessas regras de seleção às transições entre alguns dos níveis de menor energia do espectro do átomo de hidrogênio (sem inclusão de efeitos de spin) é ilustrada na figura 8.12. O diagrama do lado

Figura 8.11: Sistema de referência para a emissão de um fóton linearmente polarizado.



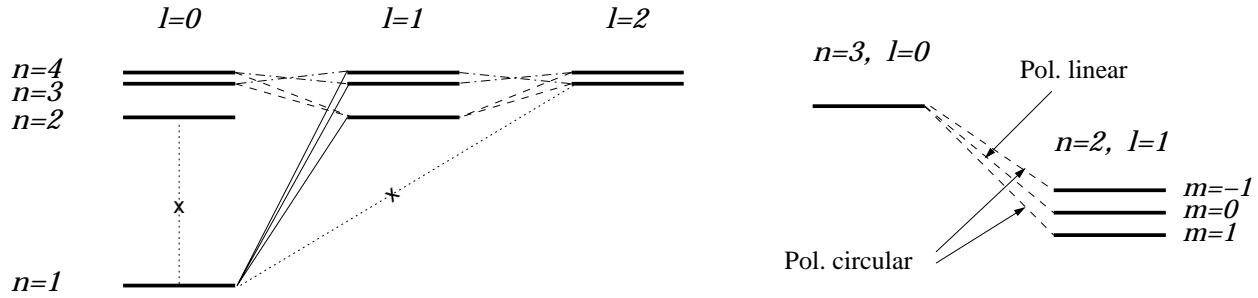


Figura 8.12: Transições de dipolo elétrico correspondentes às linhas de maior frequência das séries de Balmer (linhas tracejadas), Paschen (linhas traço-ponto) e Lyman (linhas cheias). As linhas pontilhadas e cruzadas são exemplos de transições proibidas na aproximação de dipolo elétrico. O lado direito da figura explicita as três componentes magnéticas da linha de Balmer $3s \rightarrow 2p$, $m = -1, 0, 1$ e indica a polarização da radiação emitida em cada caso. Note que o estado $3s$ tem necessariamente $m = 0$.

direito mostra as três componentes de uma transição da série de Balmer associadas aos sub-estados magnéticos do estado final com $l = 1$, indicando os respectivos estados de polarização.

Um problema estreitamente ligado com o da emissão de radiação, como tratado acima, é evidentemente o problema de *absorção* de radiação. As amplitudes de transição envolvidas nele diferem das amplitudes relacionadas com o processo de emissão apenas pela troca dos estados inicial e final. As intensidades (i.e., as probabilidades de transição) associadas ao processo de absorção diferem, no entanto, das intensidades de emissão por efeitos ligados a diferenças dos fatores de espaço de fases apropriados a cada um dos dois casos.

8.2.4 Análise multipolar das transições radiativas.

No tratamento das regras de seleção envolvendo o momento angular para as transições de dipolo elétrico foi possível ver explicitamente que a dependência angular associada à polarização do fóton emitido desempenha um papel essencial. Isso pode ser entendido facilmente à vista das expressões (8.10) e (8.11), que mostram que as propriedades de polarização estão relacionadas diretamente com a parte intrínseca (spin) do momento angular do campo de radiação. Desse modo, a quantidade que deve ser analisada em termos de componentes com momento angular e paridade bem definidos para generalizar o tratamento anterior não é apenas a onda plana associada ao vetor de onda do fóton emitido, mas a quantidade *vetorial* que aparece nas amplitudes de transição

$$\vec{\epsilon} e^{i\vec{k}\cdot\vec{r}}.$$

A questão a ser examinada em seguida é portanto a expansão dessa quantidade vetorial em uma série cujos termos tenham propriedades bem definidas de momento angular e paridade. No caso de transições radiativas entre estados ligados com momento angular e paridade definidos, apenas um número limitado de termos dessa expansão dará contribuições não nulas, sendo as demais excluídas por regras de seleção que generalizam as que foram obtidas na seção anterior.

O ponto de partida para a obtenção de uma expansão desse tipo é completar a base bi-dimensional de vetores de polarização de modo a ter um conjunto completo de *três* vetores, e a forma natural e adequada de fazer isso é introduzir, além dos dois vetores de polarização linear $\vec{\epsilon}_1$ e $\vec{\epsilon}_2$ associados ao vetor de onda \vec{k} , um terceiro vetor unitário na direção do vetor de onda. Na realidade, como visto na obtenção da equação (8.11), para tratar de propriedades de momento angular é mais conveniente trabalhar com vetores de polarização *circular*, o que leva a considerar os três vetores

$$\vec{\epsilon}_{\pm 1} = \mp \frac{1}{\sqrt{2}}(\vec{\epsilon}_1 \pm i\vec{\epsilon}_2) \quad \text{e} \quad \vec{\epsilon}_0 = \hat{k} = \frac{\vec{k}}{k}. \quad (8.32)$$

De fato, usando um sistema de coordenadas tal que $\vec{\epsilon}_1 = \hat{x}$, $\vec{\epsilon}_2 = \hat{y}$ e $\vec{\epsilon}_0 = \hat{z}$, é fácil ver que, se \hat{u} é um vetor unitário qualquer, vale a relação

$$\vec{\epsilon}_\mu \cdot \hat{u} = \sqrt{\frac{4\pi}{3}} Y_{1\mu}(\hat{u}), \quad \mu = -1, 0, 1.$$

de modo que as componentes μ de um vetor qualquer são proporcionais a harmônicas esféricas com $l = 1$ e se comportam portanto como autofunções do momento angular.

Como no sistema de coordenadas adotado o eixo z tem a direção de \vec{k} , as amplitudes de para transições radiativas podem ser escritas em termos de

$$\vec{\epsilon}_\mu e^{ikr \cos \theta} = \sum_l i^l \sqrt{4\pi(2l+1)} j_l(kr) \vec{\epsilon}_\mu Y_{l0}(\hat{r}) \quad (8.33)$$

com $\mu = \pm 1$, onde foi usada a expansão da onda plana em ondas parciais com momento angular bem definido. Usando agora o fato de que os $\vec{\epsilon}_\mu$ se comportam como as $Y_{1\mu}$, os produtos desses vetores por harmônicas esféricas que aparecem nessa expansão podem ser tratados da mesma forma que funções de onda envolvendo dois momentos angulares independentes. Em particular, é possível mudar de representação de forma a tornar diagonal o operador correspondente ao quadrado da *soma* desses dois momentos angulares, além de sua componente z . Os objetos acoplados, correspondente a esta nova representação são dados da forma usual, como

$$\vec{\mathcal{Y}}_{l1JM}(\hat{r}) \equiv \sum_{m\mu} C_{m\mu m}^{l1J} \vec{\epsilon}_\mu Y_{lm}(\hat{r})$$

onde os coeficientes com três momentos angulares são coeficientes de Clebsch-Gordan. Esses objetos são chamados *harmônicas esféricas vetoriais*, e se comportam como autofunções do momento angular *total* com números quânticos J, M , resultantes do acoplamento do momento angular l, m com o “spin” representado pelos $\vec{\epsilon}_\mu$. A ortogonalidade da transformação representada pelos coeficientes de Clebsch-Gordan permite inverter trivialmente a última relação, dando

$$\vec{\epsilon}_\mu Y_{lm}(\hat{r}) = \sum_J C_{m\ \mu\ M}^{l\ 1\ J} \vec{\mathcal{Y}}_{l1JM}(\hat{r})$$

de modo que a expansão (8.33) fica escrita como

$$\vec{\epsilon}_\mu e^{ikr \cos \theta} = \sum_l \sum_J i^l \sqrt{4\pi(2l+1)} j_l(kr) C_{0\ \mu\ \mu}^{l\ 1\ J} \vec{\mathcal{Y}}_{l1J\mu}(\hat{r}).$$

Para o cálculo das amplitudes de transição, os valores relevantes de μ são ± 1 apenas, o que *exclui* o termo $J = 0$ da soma sobre J . Por outro lado, para cada valor de J a soma sobre l tem apenas três termos, isto é, $l = J$ e $l = J \pm 1$. Introduzindo expressões explícitas para os coeficientes de Clebsch-Gordan correspondentes a esses casos resulta

$$\begin{aligned} \vec{\epsilon}_{\mu=\pm 1} e^{ikr \cos \theta} &= \sum_{J \geq 1} i^J \sqrt{\frac{4\pi(2J+1)}{2}} \times \\ &\times \left\{ -\mu j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) - i \left[\sqrt{\frac{J+1}{2J+1}} j_{J-1}(kr) \vec{\mathcal{Y}}_{J-11J\mu}(\hat{r}) - \sqrt{\frac{J}{2J+1}} j_{J+1}(kr) \vec{\mathcal{Y}}_{J+11J\mu}(\hat{r}) \right] \right\} \end{aligned}$$

que pode ainda ser transformada fazendo uso da identidade

$$\begin{aligned} \nabla \times j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) &\equiv \\ &\equiv i \left[\left(\frac{d}{dr} - \frac{J}{r} \right) j_J(kr) \sqrt{\frac{J}{2J+1}} \vec{\mathcal{Y}}_{J+11J\mu}(\hat{r}) + \left(\frac{d}{dr} + \frac{J+1}{r} \right) j_J(kr) \sqrt{\frac{J+1}{2J+1}} \vec{\mathcal{Y}}_{J-11J\mu}(\hat{r}) \right] \end{aligned} \quad (8.34)$$

bem como das relações (cf. equação (5.16))

$$\left(\frac{d}{dr} - \frac{J}{r} \right) j_J(kr) = -k j_{J+1}(kr) \quad \text{e} \quad \left(\frac{d}{dr} + \frac{J+1}{r} \right) j_J(kr) = k j_{J-1}(kr)$$

levando finalmente ao resultado

$$\vec{\epsilon}_{\mu=\pm 1} e^{ikr \cos \theta} = \sum_{J \geq 1} i^J \sqrt{2\pi(2J+1)} \left[-\mu j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) - \frac{1}{k} \nabla \times (j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r})) \right]. \quad (8.35)$$

A decomposição multipolar de amplitude de transição resulta diretamente da substituição dessa expansão em (8.28). De fato, introduzindo a notação abreviada \vec{J}_{IF} , $I \equiv nlm$, $F \equiv n'l'm'$ para a densidade de corrente de transição, e usando a propriedade de conjugação complexa das harmônicas esféricas vetoriais

$$\vec{\mathcal{Y}}_{J1J\mu=\pm 1}^*(\hat{r}) = -\vec{\mathcal{Y}}_{J1J-\mu}(\hat{r})$$

resulta

$$\vec{\epsilon}_{\mu=\pm 1} \cdot \vec{J}_{IF} = \sum_{J \geq 1} (-i)^J \sqrt{2\pi(2J+1)} \left(T_{IF}^{EJ,-\mu}(k) + \mu T_{IF}^{MJ,-\mu}(k) \right) \quad (8.36)$$

onde foram introduzidas as amplitudes correspondentes aos 2^J -polos elétricos

$$T_{IF}^{EJ,\mu}(k) \equiv \frac{1}{k} \int d^3r \left[\nabla \times j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) \right] \cdot \left[\psi_F^*(\vec{r}) \frac{\vec{p}}{m_e c} \psi_I(\vec{r}) \right] \quad (8.37)$$

e aos 2^J -polos magnéticos

$$T_{IF}^{MJ,\mu}(k) \equiv \int d^3r j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) \cdot \left[\psi_F^*(\vec{r}) \frac{\vec{p}}{m_e c} \psi_I(\vec{r}) \right]. \quad (8.38)$$

O que distingue os multipolos elétricos dos magnéticos, para um dado valor de J , é a paridade. De fato, é possível mostrar⁴ que os operadores

$$\left[\nabla \times j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) \right] \cdot \frac{\vec{p}}{m_e c} \quad \text{e} \quad j_J(kr) \vec{\mathcal{Y}}_{J1J\mu}(\hat{r}) \cdot \frac{\vec{p}}{m_e c}$$

têm, respectivamente, paridades $(-1)^J$ e $(-1)^{J+1}$. Uma forma freqüentemente usada para exprimir isso é dizer que os multipolos elétricos têm paridade “natural”, enquanto os multipolos magnéticos têm paridade “não natural”. Além disso, os multipolos de ordem J se comportam como autofunções do momento angular com autovalores J, μ . Essas propriedades permitem obter facilmente regras de seleção para transições de multipolaridade arbitrária, usando as técnicas de momento angular que serão desenvolvidas na seção 9.3.3.

⁴V. e.g. A. R. Edmonds, citado na lista de referências gerais.

Para fazer contacto com o tratamento da seção anterior, é interessante examinar o comportamento dos diferentes multipolos da expansão geral (8.36) no limite de grandes comprimentos de onda. O fator que envolve funções de Bessel e as harmônicas esféricas vetoriais do termo de dipolo elétrico, com o uso da identidade (8.34) pode ser escrito como

$$\frac{-i}{k} \nabla \times j_1(kr) \vec{\mathcal{Y}}_{111-\mu}(\hat{r}) = -j_2(kr) \sqrt{\frac{J}{2J+1}} \vec{\mathcal{Y}}_{111-\mu}(\hat{r}) + j_0(kr) \sqrt{\frac{J+1}{2J+1}} \vec{\mathcal{Y}}_{011-\mu}(\hat{r}).$$

No limite de grandes comprimentos de onda, $kr \rightarrow 0$ e o termo envolvendo $j_2(kr)$ vai a zero como $(kr)^2$, enquanto $j_0(kr) = \sin kr / kr \rightarrow 1$. A harmônica esférica vetorial que acompanha $j_0(kr)$ é, por outro lado,

$$\vec{\mathcal{Y}}_{011-\mu}(\hat{r}) = C_{0-\mu-1}^0 Y_{00}(\hat{r}) \vec{\epsilon}_{-\mu} = \frac{\vec{\epsilon}_{-\mu}}{\sqrt{4\pi}}$$

pois o coeficiente de Clebsch-Gordan que aí aparece é igual a 1. De forma análoga, resulta que todos os demais multipolos (elétricos e magnéticos) se anulam quando $kr \rightarrow 0$, pois todos eles contêm funções de Bessel $j_n(kr)$ com $n \geq 1$, e essas funções são proporcionais a $(kr)^n$ nesse limite. A parte remanescente do termo de dipolo elétrico reproduz o resultado obtido antes diretamente da aproximação $e^{-i\vec{k} \cdot \vec{r}} \simeq 1$. Isso mostra que a amplitude calculada na seção anterior corresponde, de fato, ao limite para grandes comprimentos de onda do termo de dipolo elétrico completo. Em situações mais gerais, além de considerar as possíveis contribuições de outros multipolos, a contribuição do termo de dipolo elétrico deve ser calculada com a expressão completa, dada por (8.37) com $J = 1$.

Problemas.

8.1 - a) Mostre que a transformação

$$b_{\vec{k}+} = \frac{a_{\vec{k}1} - ia_{\vec{k}2}}{\sqrt{2}}, \quad b_{\vec{k}-} = \frac{a_{\vec{k}1} + ia_{\vec{k}2}}{\sqrt{2}}$$

preserva as relações de comutação $[a_\lambda, a_{\lambda'}] = 0$, $[a_\lambda, a_{\lambda'}^\dagger] = \delta_{\lambda\lambda'}$, isto é, as mesmas relações de comutação valem para os b_λ , $b_{\lambda'}$, com $\lambda \equiv \{\vec{k}, \alpha\}$, $\alpha = 1, 2$ ou \pm .

b) Se $|0\rangle$ é o vácuo, considere os estados de um fóton

$$|\vec{k}, \vec{\epsilon}_{1,2}\rangle = a_{\vec{k}, 1,2}^\dagger |0\rangle, \quad |\vec{k}, \vec{\epsilon}_\pm\rangle = b_{\vec{k}, \pm}^\dagger |0\rangle$$

e calcule as “funções de onda” correspondentes a eles, definidas como $\langle 0 | \vec{A}(\vec{r}, t) | \text{estado} \rangle$, onde

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}, \alpha} \sqrt{\frac{2\pi\hbar c}{kV}} \vec{\epsilon}_{\vec{k}, \alpha} \left[a_{\vec{k}, \alpha} e^{i(\vec{k} \cdot \vec{r} - \omega t)} + a_{\vec{k}, \alpha}^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega t)} \right].$$

Mostre que os dois tipos de estado correspondem respectivamente a fótons polarizados linearmente e circularmente. Identifique o fóton circularmente polarizado à direita.

8.2 - a) Calcule, a partir da expressão (8.19), a seção de choque Thompson para um feixe incidente de fótons *não polarizados* mas para uma polarização dada $\vec{\epsilon}'$ do fóton emergente. **Sugestão:** utilize os vetores básicos de polarização $\vec{\epsilon}_i$, $\vec{\epsilon}_i'$ como definidos na figura 8.3, escrevendo a polarização final como $\vec{\epsilon}' = \vec{\epsilon}_1' \cos \varphi' + \vec{\epsilon}_2' \sin \varphi'$.

b) Verifique que essa seção de choque depende, em geral, da direção do vetor de polarização $\vec{\epsilon}'$, de modo que os fótons emergentes são parcialmente polarizados. Definindo o grau de polarização $p(\theta)$, para um valor dado do ângulo de espalhamento θ como

$$p(\theta) = \frac{d\sigma_{\max} - d\sigma_{\min}}{d\sigma_{\max}}$$

onde $d\sigma_{\max}/d\Omega$ e $d\sigma_{\min}/d\Omega$ são respectivamente o valor máximo e o valor mínimo de seção de choque diferencial como função de φ' , com θ fixo, verifique que $p(\theta = \pi/2) = 1$ (polarização completa) e $p(\theta = 0) = 0$ (fótons emergentes não polarizados).

c) Procure justificar a definição dada do grau de polarização. **Sugestão:** a probabilidade de observar um fóton com polarização $\vec{\epsilon}$ em um feixe de fótons de momento $\hbar\vec{k}$ parcialmente polarizado, definido pelos pesos p_1 e $p_2 = 1 - p_1$ para os estados de polarização $\vec{\epsilon}_1$ e $\vec{\epsilon}_2$, é $\sum_{i=1}^2 p_i (\vec{\epsilon}_i \cdot \vec{\epsilon})^2$.

d) Em que direção estão polarizados os fótons espalhados de $\theta = \pi/2$? Você saberia como entender essa polarização, bem como a polarização parcial dos fótons no caso geral, em termos clássicos?

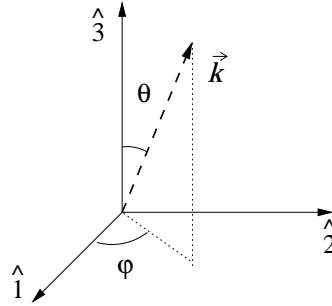
8.3 - Espalhamento Rayleigh. Resolva o problema do espalhamento elástico de fótons não polarizados de frequência muito baixa por um oscilador harmônico em seu estado fundamental. Use

$$H = \sum_{\lambda} \hbar\omega_{\lambda} a_{\lambda}^\dagger a_{\lambda} + \frac{(\vec{p} - \frac{e}{c} \vec{A})^2}{2M} + \frac{1}{2} M\omega_0^2 r^2$$

e faça as seguintes aproximações:

a) Use apenas um estado intermediário, isto é um primeiro estado excitado do oscilador.

Figura 8.13: Sistema de referência para a emissão do fóton no problema 8.4. O eixo de quantização do momento angular é \hat{z} .



b) Para calcular os elementos de matriz use o fato de que $kb \ll 1$, onde k é o número de onda do fóton e b é o parâmetro do oscilador, $b = \sqrt{\hbar/M\omega_0}$. Obtenha a seção de choque diferencial e total. Note que não há conservação de momento.

8.4 - Linha de Balmer no Hidrogênio. Um átomo de Hidrogênio sofre uma transição de dipolo elétrico do nível $3s$ para o nível $2p$ com $\Delta m = 0$, sendo m o “número quântico magnético” (autovalor de L_3 em unidades de \hbar , v. figura 8.12). Escolha um eixo de quantização \hat{z} fixo para o momento angular orbital \vec{L} e defina a direção do fóton emitido através dos ângulos θ e φ indicados na figura 8.13.

a) Ache a distribuição angular do fóton emitido, no caso em que a sua polarização *não* é observada.

b) Calcule a vida média (em *segundos*) do estado $3s$ relativamente a esse tipo de transição.

Sugestão: nas condições propostas a quantidade $\vec{\epsilon} \cdot \vec{r}$ que aparece na amplitude de transição pode ser especializada para $\vec{\epsilon} \cdot \hat{z} r \cos \theta$ devido ao fato de que $\Delta m = 0$ para a transição considerada.

Referências gerais.

Um tratamento do campo de radiação semelhante ao dado aqui pode ser encontrado em

★ **J. J. Sakurai**, *Advanced Quantum Mechanics*, Addison-Wesley 1967, Capítulo 2.

Outros textos “clássicos” que usam esse mesmo tipo de tratamento são

★ **P. A. M. Dirac**, *Quantum Mechanics*, 4th Edition, Oxford 1958, Capítulo X.

★ **W. Heitler**, *The Quantum Theory of Radiation*, Oxford 1936, Capítulo II.

Um trabalho de revisão sobre a interação de sistemas não relativísticos com a radiação é

★ **P. W. Miloni**, *Semiclassical and Quantum-Electrodynamical approaches in Nonrelativistic Radiation Theory*, Phys. Lett. **C25** (1976), pags. 1-81.

As técnicas de expansão em multipolos são discutidas em muitos textos sobre o momento angular na mecânica quântica. Uma referência frequentemente usada para esses assuntos é

★ A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton U. Press 1957.

Um tratamento simples das harmônicas esféricas vetoriais é dado na referência antiga mas também “clássica” de física nuclear

★ **J. Blatt e V. Weisskopf**, *Theoretical Nuclear Physics*, Wiley 1954 (re-publicada por Dover em 1991), Apêndice B.

Capítulo 9

Transformações, simetrias e leis de conservação.

As leis gerais de conservação de energia, momento e momento angular para um sistema físico descrito em termos da mecânica clássica estão associadas a propriedades de invariância da dinâmica do sistema sob a ação geral das transformações espacio-temporais que são, respectivamente, translações temporais, translações espaciais e rotações espaciais. O objetivo deste capítulo é examinar os resultados correspondentes para a mecânica quântica e estendê-los para outros objetos ou propriedades quânticas, eventualmente sem análogo clássico, de que um exemplo é o spin.

Uma questão de natureza algo burocrática, mas com implicações práticas importantes, é a de que as transformações espacio-temporais podem ser descritas de dois pontos de vista alternativos, às vezes chamados de *ativo* e *passivo*, respectivamente. Uma transformação espacio-temporal é dita do tipo ativo quando ela é entendida como aplicada ao próprio sistema físico em questão, isto é, o sistema físico é considerado em outro instante de tempo, levado a outra posição ou orientação enquanto descrito em termos de um referencial espacial e uma escala de tempo fixos. Por outro lado, uma transformação é dita do tipo passivo quando entendida como aplicada ao referencial ou escala de tempo em termos dos quais o sistema é descrito. Embora equivalentes, esses dois pontos de vista devem ser cuidadosamente distinguidos no trabalho formal, pois a passagem de um a outro implica em mudanças de fase, geralmente de aparência sutil mas cruciais para a consistência geral do tratamento. Para ter um exemplo simples disso basta considerar uma translação espacial definida por um dado vetor \vec{a} para um sistema constituído de uma única partícula na posição \vec{r} , com relação a uma dada origem O . Se a translação é entendida dentro do ponto de vista ativo, ela leva a partícula à nova posição $\vec{r}' = \vec{r} + \vec{a}$ com relação à mesma origem O , enquanto que, do ponto de vista passivo, a origem é deslocada para $O' = O + \vec{a}$, com relação à qual a posição da partícula passa a ser $\vec{r}' = \vec{r} - \vec{a}$. Isso faz ver

também que a transformação passiva que consiste na translação definida pelo vetor $-\vec{a}$ leva a um vetor de posição cuja expressão é idêntica ao da translação ativa definida por $+\vec{a}$. Qualquer mistura dos dois pontos de vista ao longo de algum desenvolvimento formal pode portanto levar a erros irrecuperáveis de sinal.

No que segue, salvo menção expressa em contrário, será adotado o ponto de vista *passivo*. Uma forma geral de escrever uma transformação A qualquer das transformações espacio-temporais mencionadas (e até mesmo uma combinação delas) é portanto

$$\{q_i, t\} \longrightarrow \{q'_i, t'\} \equiv A\{q_i, t\} \quad i = 1, 2, 3$$

onde $\{q_i, t\}$ representa posições e instantes identificados no referencial inicial e $\{q'_i, t'\}$ representa as mesmas posições e instantes identificados no referencial transformado. As transformações a serem consideradas são ainda independentes do tempo, o que exclui, em particular, transformações para sistemas que se movem com relação ao sistema de partida. Mais formalmente, isso significa que a relação entre os q'_i e os q_i é independente do tempo.

9.1 Grupos de transformações.

Um fato importante acerca das transformações espacio-temporais consideradas é o fato de que elas constituem famílias caracterizadas pelas seguintes propriedades:

1) A aplicação sucessiva de duas transformações A e B quaisquer da família é equivalente a uma outra transformação $D = BA$ também da família, isto é

$$B(A\{q_i, t\}) = \{q''_i, t''\} = D\{q_i, t\} \equiv (BA)\{q_i, t\}.$$

A transformação D é chamada *produto*, ou *composição* da transformação A com a transformação B (nessa ordem). Esse produto, ou lei de composição, deve ser associativo, o que quer dizer que se C é também uma transformação, $(C(BA)) = ((CB)A)$, mas em geral pode ser *não comutativo*, isto é, em geral $BA \neq AB$. Quando $BA = AB$ para quaisquer dois elementos A e B o grupo é dito *comutativo*, ou *abeliano*.

2) Existe uma transformação “idêntica” 1 tal que $1\{q_i, t\} = \{q_i, t\}$.

3) Para cada transformação A da família existe uma outra transformação A^{-1} , também da família, tal que $A^{-1}A = AA^{-1} = 1$, onde 1 é a transformação idêntica.

Famílias de transformações com essas propriedades são chamadas *grupos* de transformações. Um exemplo extremamente simples de um grupo de transformações que tem apenas dois elementos é a família formada pela transformação P de reflexão espacial, definida por

$$P\{q_i, t\} = \{-q_i, t\},$$

onde os q_i são entendidos como componentes cartesianas de um vetor de posição, além da transformação idêntica. De fato, é imediato que $PP = 1$ e $1P = P1 = P$. Portanto $P^{-1} = P$ e as três propriedades estão satisfeitas. Um exemplo ainda mais simples (embora trivial) de grupo é o que é formado pela transformação idêntica *apenas*. Esta pode ser composta consigo mesma dando novamente a transformação idêntica, que é portanto sua própria inversa, satisfazendo sozinha as três propriedades que caracterizam um grupo. Esse grupo de um único elemento é, por outro lado, parte do grupo de dois elementos $\{1, P\}$, e é portanto o que se chama um *sub-grupo* do grupo maior. A transformação idêntica é aliás um sub-grupo (trivial) de *qualquer* grupo de transformações.

9.1.1 Grupos contínuos, geradores.

Há grupos que têm uma infinidade *contínua* de elementos. Um exemplo simples de um tal grupo é o das rotações em torno do eixo 3, definidas por $t' = t$ e

$$\begin{pmatrix} q'_1 \\ q'_2 \\ q'_3 \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \equiv A_\varphi \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}.$$

Para essas transformações é fácil verificar diretamente que

$$A_{\varphi_2} A_{\varphi_1} = A_{\varphi_1 + \varphi_2} = A_{\varphi_1} A_{\varphi_2}.$$

Por outro lado, é imediato que $A_{\varphi=0} = 1$, logo $A_\varphi^{-1} = A_{-\varphi}$, o que verifica as três propriedades que caracterizam os grupos. Da lei de composição do grupo decorre, em particular, que esse é um grupo comutativo, ou abeliano. Como apropriado para o ponto de vista passivo, o parâmetro φ é o ângulo de rotação do sistema de coordenadas em torno do eixo 3, no sentido anti-horário, e pode portanto variar de forma contínua. Para tais grupos é conveniente considerar transformações “infinitesimais”, no sentido que são muito próximas da transformação idêntica 1, pois elas permitem, em condições razoavelmente gerais, reconstruir todas as demais transformações usando a lei de composição que define o grupo. Para este último exemplo, uma transformação infinitesimal é uma rotação de um ângulo infinitesimal $d\omega$, que pode ser escrita como

$$A_{d\omega} = \mathbf{1} + d\omega \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(d\omega^2) \equiv \mathbf{1} + d\omega \lambda_3 + \mathcal{O}(d\omega^2). \quad (9.1)$$

A rotação de um ângulo finito φ pode ser escrita usando a lei de composição do grupo como

$$A_\varphi = \left(A_{\frac{\varphi}{n}}\right)^n = \left(1 + \frac{\varphi}{n}\lambda_3 + \mathcal{O}\left(\frac{\varphi}{n}\right)^2\right)^n.$$

Tomando o limite $n \rightarrow \infty$ é possível ignorar os termos de ordem superior à primeira em φ/n e usar a relação

$$\lim_{n \rightarrow \infty} \left(1 + \frac{\alpha}{n}\right)^n = e^\alpha$$

para obter o resultado

$$A_\varphi = e^{\varphi\lambda_3}.$$

Não é difícil verificar que uma expansão em série de potências dessa exponencial reproduz a expansão da expressão original de A_φ , que pode portanto ser obtida desse modo em termos da matriz λ_3 que caracteriza as transformações infinitesimais do grupo. Por isso λ_3 é chamado *gerador* do grupo de rotações em torno do eixo 3.

Um grupo contínuo geral tem um número de geradores igual ao número de parâmetros contínuos que rotulam seus elementos. As transformações finitas do grupo podem ser expressas em termos de seus geradores através de recursos análogos aos usados nesse exemplo. É importante notar desde logo, porém, que para grupos contínuos não comutativos cuidados especiais são necessários para ter em conta questões de ordenamento no uso da lei de composição.

9.1.2 Representações.

A ação das transformações consideradas até aqui se dá na realidade ao nível das coordenadas espacio-temporais em termos das quais se faz a descrição dos sistemas físicos. O próximo passo consiste então em estudar os efeitos dessas transformações sobre os objetos cinemáticos básicos envolvidos na descrição da dinâmica desses sistemas.

No contexto da mecânica quântica, os estados de um sistema físico podem ser representados por funções de onda $\psi(q)$ definidas sobre o espaço de configurações do sistema. Nessa notação, q designa coletivamente todas as coordenadas de posição necessárias para identificar um ponto desse espaço. No caso de um sistema com N partículas em três dimensões espaciais, por exemplo, esse argumento envolve $3N$ coordenadas. Dada uma transformação

$$A\{q_i\} = \{q'_i\}$$

em que $\{q_i\}$ e $\{q'_i\}$, $i = 1, 2, 3$ designam agora as antigas e novas coordenadas de um ponto no espaço de três dimensões, é possível definir uma transformação correspondente R_A no espaço de funções de onda através de

$$R_A\psi = \psi'$$

onde as novas funções ψ' são definidas como tendo, para os valores q' , transformados por A , das coordenadas, o mesmo valor que as funções ψ assumem nos valores originais q , isto é

$$\psi'(q') = \psi(q).$$

Em geral $\psi'(q) \neq \psi(q)$, isto é, as funções ψ' e ψ assumem valores diferentes para argumentos iguais, e são portanto funções diferentes. No entanto elas assumem, por definição, o mesmo valor para argumentos relacionados através da transformação A . A transformação R_A é portanto um operador que age no espaço de funções de onda definido de forma que se tenha, para qualquer função de onda $\psi(q)$,

$$(R_A\psi)(Aq) = \psi(q) \quad \text{ou, equivalentemente,} \quad (R_A\psi)(q) = \psi(A^{-1}q) \quad (9.2)$$

onde a notação usada no primeiro membro indica o valor assumido pela função transformada $R_A\psi$ no ponto transformado do espaço de configurações Aq , etc.

Considerando agora um *grupo* de transformações espacio-temporais, é possível verificar que o conjunto associado de operadores definidos desse modo no espaço de funções de onda constitui também um grupo, cujos elementos podem ser postos em correspondência com os elementos do grupo de partida. De fato, usando relações como a (9.2) para definir operadores R_A , R_B e R_{BA} associados respectivamente às duas transformações espacio-temporais A e B e à sua composição BA , resulta

$$(R_{BA}\psi)(q) = \psi((BA)^{-1}q) = \psi(A^{-1}B^{-1}q) = (R_A\psi)(B^{-1}q) = (R_B(R_A\psi))(q),$$

mostrando que à lei de composição do grupo para as transformações A e B corresponde a lei de composição análoga para os operadores

$$R_{BA} = R_BR_A.$$

Com procedimentos semelhantes a esse é possível verificar que as demais propriedades que caracterizam o grupo de transformações espacio-temporais também se transferem para os operadores definidos como em (9.2). O grupo assim formado por esses operadores é chamado uma *representação* do grupo de transformações espacio-temporais considerado.

A representação do grupo de transformações $\{A\}$ em termos dos operadores $\{R_A\}$ construída acima não é única. De fato, a relação (9.2) que define o operador R_A pode, por exemplo, ser generalizada para

$$(U\bar{R}_A\psi)(Aq) = U\psi(q)$$

onde U é um operador unitário qualquer no espaço de Hilbert das funções de onda, sem que isso afete as propriedades de grupo dos novos operadores \bar{R}_A . É claro que para $U = \hat{1}$ $R_A = \bar{R}_A$. Para obter a relação geral entre R_A e \bar{R}_A , basta notar que, segundo a definição (9.2),

$$(R_A U\psi)(Aq) = U\psi(q)$$

para qualquer função de onda $\psi(q)$. Comparando essa relação com a definição de \bar{R}_A , resulta que

$$U\bar{R}_A = R_A U \quad \text{ou} \quad \bar{R}_A = U^\dagger R_A U$$

isto é, os operadores \bar{R}_A são obtidos dos R_A através da transformação unitária U , o que implica, em particular, na preservação das propriedades de grupo.

Uma representação $\{R_A\}$ é chamada *redutível* quando existirem sub-espços próprios do espaço de Hilbert em que ela é definida (isto é, sub-espços diferentes do espaço todo) que são invariantes sob a ação de *todos* os operadores que constituem a representação em questão. Os operadores que constituem a representação do grupo podem ser expressos também em termos de matrizes $(R_A)_{ij}$, definidas com relação a uma base dada $\{\phi_k\}$ do espaço de Hilbert apropriado ao sistema considerado. Quando a representação é redutível, existe uma mudança de base $\{\phi_k\} \rightarrow \{\bar{\phi}_\nu\}$ (equivalente a uma transformação unitária dos R_A) tal que os sub-espços invariantes são associados a projetores $\Pi_I \equiv \sum_{\nu \in I} \bar{\phi}_\nu \bar{\phi}_\nu^*$ e as matrizes que representam os R_A são todas tais que $(R_A)_{\nu\mu} \equiv 0$ sempre que ν e μ pertençam a sub-espços diferentes. Isso significa que todas as matrizes que representam os R_A na nova base de decompões em blocos independentes, cada um dos quais corresponde à restrição da ação dos R_A sobre os vetores de um dos sub-espços invariantes. Cada uma dessas restrições define, por sua vez, uma outra representação do grupo. A *dimensão* de uma dessas representações é o número de vetores linearmente independentes do sub-espço em que ela é definida (isto é, é a dimensão do sub-espço invariante considerado). Uma representação é chamada *irredutível* quando não existirem sub-espços invariantes menores que aquele no qual ela é definida. É possível mostrar que representações irredutíveis de grupos comutativos (ou abelianos) são unidimensionais.

A ação dos operadores que constituem a representação do grupo sobre as variáveis dinâmicas g (operadores hermiteanos no espaço de Hilbert) é definida da forma usual como

$$R_A g \psi(q) = R_A g R_A^{-1} R_A \psi(q) \equiv g' R_A \psi(q) = g' \psi'(q)$$

ou seja

$$g' = R_A g R_A^{-1}.$$

Em particular, o hamiltoniano que caracteriza a dinâmica do sistema físico considerado se transforma como

$$H' = R_A H R_A^{-1}. \quad (9.3)$$

O sistema físico é dito *invariante* por um dado grupo de transformações se $H' = H$ para *todos* os operadores que constituem a representação do grupo. A invariância sob um dado grupo de transformações implica que a forma da equação de movimento é preservada sob as transformações do grupo. De fato, a forma transformada da equação de movimento

$$H\psi = i\hbar \frac{\partial \psi}{\partial t}$$

é, em geral,

$$H'\psi' = i\hbar \frac{\partial \psi'}{\partial t}$$

e, quando $H' = H$, a solução transformada ψ' satisfaz a mesma equação que a solução ψ .

A condição de invariância $H' = H$, com H' dado por (9.3), pode ser expressa também através da relação

$$[R_A, H] = 0,$$

válida para todos os operadores R_A que constituem a representação do grupo. Essa relação mostra explicitamente que a invariância de um sistema físico por um dado grupo de transformações leva à existência de constantes do movimento (v. seção 2.3.3) associadas às transformações do grupo. Em particular, no caso de grupos contínuos, as variáveis dinâmicas relevantes podem ser associadas aos geradores do grupo.

Para concretizar esses fatos gerais com um exemplo simples pode ser útil considerar a sua aplicação ao grupo de dois elementos formado pela transformação de inversão espacial P e pela transformação idêntica, já considerado na página 471. Os operadores que representam esse grupo no espaço de Hilbert das funções de onda $\psi(q)$ de um dado sistema podem ser

construídos usando diretamente a relação (9.2) para cada um dos elementos do grupo. No caso da transformação idêntica

$$(R_1\psi)(q) = \psi(1^{-1}q) = \psi(q)$$

para qualquer função de onda $\psi(q)$. Isso significa que R_1 associa qualquer função de onda a si mesma, e portanto é o operador identidade $\hat{1}$ do espaço de Hilbert considerado. Para o operador de inversão P , por outro lado

$$(R_P\psi)(q) = \psi(P^{-1}q) = \psi(Pq) = \psi(-q)$$

o que identifica o operador R_P com o operador de inversão espacial \mathcal{P} introduzido na discussão do efeito Stark (v. página 318). O grupo de transformações $\{1, P\}$ é portanto representado no espaço de Hilbert considerado pelos operadores $\{\hat{1}, \mathcal{P}\}$. Um sistema é invariante com relação a esse grupo de transformações se o hamiltoniano H que caracteriza sua dinâmica comuta com esses dois operadores. Como a comutação com o operador identidade $\hat{1}$ é trivial, a condição de invariança se reduz na realidade, neste caso à condição

$$[H, \mathcal{P}] = 0. \quad (9.4)$$

Como foi observado anteriormente, o operador \mathcal{P} é tanto hermiteano quanto unitário. A condição de invariança é na realidade também a condição de compatibilidade com H , de modo que as autofunções de H podem ser sempre escolhidas de modo a serem também autofunções da constante do movimento \mathcal{P} (estados estacionários de paridade definida). De fato, da comutatividade de H com \mathcal{P} segue em geral que, se $H\psi_E = E\psi_E$,

$$\mathcal{P}H\psi_E = H\mathcal{P}\psi_E = H\psi'_E = E\psi'_E$$

isto é, $\psi'_E = \mathcal{P}\psi_E$ é também uma autofunção de H com o mesmo autovalor E . Com relação ao operador \mathcal{P} há então dois casos possíveis a considerar:

a) a função ψ_E é também autofunção de \mathcal{P} . Como $\mathcal{P}^2 = \mathcal{P}$ os autovalores possíveis são ± 1 e portanto

$$\psi'_E = \mathcal{P}\psi_E = \pm\psi_E;$$

b) a função ψ_E não é autofunção de \mathcal{P} , isto é, $\psi'_E = \mathcal{P}\psi_E$ não é proporcional a ψ_E . Neste caso, usando o fato que $\mathcal{P}^2 = \mathcal{P}$ é imediato verificar que as duas funções ortogonais definidas por

$$\bar{\psi}_E^{(\pm)} \equiv \psi_E \pm \mathcal{P}\psi_E = \psi_E \pm \psi'_E$$

são autofunções não apenas de H mas também de \mathcal{P} , com autovalores ± 1 respectivamente. A não proporcionalidade de ψ'_E a ψ_E , juntamente com a simetria (9.4), implica portanto na degenerescência pelo menos dupla do autovalor E de H .

9.2 Translações espaciais e momento.

Uma translação espacial (passiva) pode ser caracterizada por um deslocamento da origem do referencial de um vetor $\vec{d} \equiv \{d_1, d_2, d_3\}$ de modo que ela corresponde à transformação do vetor de posição $\vec{q} \equiv \{q_1, q_2, q_3\}$ dada por

$$q'_i = T_{\vec{d}} q_i = q_i - d_i, \quad i = 1, 2, 3 \quad \text{ou} \quad \vec{q}' = T_{\vec{d}} \vec{q} = \vec{q} - \vec{d}.$$

É imediato verificar que essas transformações formam um grupo (contínuo) comutativo, com a lei de composição

$$T_{\vec{d}_2} T_{\vec{d}_1} = T_{\vec{d}_1 + \vec{d}_2} = T_{\vec{d}_1} T_{\vec{d}_2}$$

para quaisquer \vec{d}_1 e \vec{d}_2 . A identidade do grupo corresponde a tomar para \vec{d} o vetor nulo, e $T_{\vec{d}}^{-1} = T_{-\vec{d}}$.

A ação desse grupo sobre as funções de onda que representam os estados de um sistema quântico envolve a construção da representação apropriada $\{R_{\vec{d}}$ no espaço de Hilbert correspondente. Para um sistema de N partículas as funções de onda podem ser escritas como $\psi(\vec{r}_1, \dots, \vec{r}_N) \equiv \psi(\vec{r}_\nu)$. Essas funções podem conter também variáveis de spin. Estas não são, no entanto, relevantes neste caso, e por isso não serão exibidas explicitamente. Usando a relação geral (9.2) a representação é definida por

$$\psi'(\vec{q}_\nu) \equiv (R_{\vec{d}} \psi)(\vec{q}_\nu) = \psi(T_{-\vec{d}} \vec{q}_\nu) = \psi(\vec{q}_\nu + \vec{d}).$$

Como este é um grupo contínuo, é vantajoso obter a representação de seus geradores e em seguida exprimir os $R_{\vec{d}}$, para \vec{d} finito, em termos desses geradores. Para isso basta considerar translações infinitesimais, para as quais se pode escrever

$$\psi'(\vec{q}_\nu) = \psi(\vec{q}_\nu + \vec{\epsilon}) = \psi(\vec{q}_\nu) + \sum_{\nu} \vec{\epsilon} \cdot \nabla_{\nu} \psi(\vec{q}_\nu) + \mathcal{O}(\epsilon^2)$$

ou seja

$$\psi'(\vec{q}_\nu) = \left[1 + \sum_{\nu} \vec{\epsilon} \cdot \nabla_{\nu} \right] \psi(\vec{q}_\nu) + \mathcal{O}(\epsilon^2)$$

de modo que

$$R_{\vec{\epsilon}} = 1 + \sum_{\nu} \vec{\epsilon} \cdot \nabla_{\nu} + \mathcal{O}(\epsilon^2) = 1 + \frac{i}{\hbar} \vec{\epsilon} \cdot \vec{P} + \mathcal{O}(\epsilon^2)$$

onde

$$\vec{P} = \sum_{\nu} \frac{\hbar}{i} \nabla_{\nu}$$

é o operador associado ao momento total do sistema. Isso mostra portanto que os três geradores da representação do grupo de translações espaciais são essencialmente as três componentes do momento total do sistema considerado.

A representação das translações *finitas* no espaço de Hilbert das funções de onda podem agora ser obtidas usando a lei de composição do grupo para escrever

$$R_{\vec{d}} = \lim_{n \rightarrow \infty} \left(1 + \frac{i}{\hbar} \frac{\vec{d}}{n} \cdot \vec{P} \right)^n = e^{\frac{i}{\hbar} \vec{d} \cdot \vec{P}}. \quad (9.5)$$

Dessa forma resulta finalmente que

$$\psi'(\vec{q}_{\nu}) = \psi(\vec{q}_{\nu} + \vec{d}) = R_{\vec{d}} \psi(\vec{q}_{\nu}) = e^{\frac{i}{\hbar} \vec{d} \cdot \vec{P}} \psi(\vec{q}_{\nu}). \quad (9.6)$$

A invariância translacional do sistema, por outro lado, é expressa pela condição $[H, R_{\vec{d}}] = 0$ para qualquer \vec{d} . Tendo em vista a expressão (9.5), essa condição é equivalente a

$$[H, \vec{P}] = 0$$

isto é, o momento total do sistema é uma constante do movimento. Isso significa também que H e \vec{P} possuem um conjunto completo de autofunções comuns $\phi_{E, \vec{K}}(\vec{q}_{\nu})$,

$$H \phi_{E, \vec{K}}(\vec{q}_{\nu}) = E \phi_{E, \vec{K}}(\vec{q}_{\nu}) \quad \text{e} \quad \vec{P} \phi_{E, \vec{K}}(\vec{q}_{\nu}) = \hbar \vec{K} \phi_{E, \vec{K}}(\vec{q}_{\nu}).$$

Isso na realidade implica em condições sobre a estrutura geral dessas autofunções comuns, que podem ser explicitadas notando que, em vista da relação (9.6) e do fato de que elas são autofunções do momento total,

$$\phi_{E, \vec{K}}(\vec{q}_{\nu} + \vec{d}) = e^{\frac{i}{\hbar} \vec{d} \cdot \vec{P}} \phi_{E, \vec{K}}(\vec{q}_{\nu}) = e^{i \vec{d} \cdot \vec{K}} \phi_{E, \vec{K}}(\vec{q}_{\nu}).$$

Como o momento total \vec{P} é canonicamente conjugado à posição do centro de massa

$$\vec{X} = \frac{1}{\sum_{\nu} m_{\nu}} \sum_{\nu} m_{\nu} \vec{q}_{\nu},$$

onde m_{ν} são as massas das partículas, segue que $\phi_{E, \vec{K}}(\vec{q}_{\nu})$ deve ter a forma

$$\phi_{E,\vec{K}}(\vec{q}_\nu) = e^{i\vec{K}\cdot\vec{X}} \varphi_{E,\vec{K}}(\xi_\lambda)$$

onde as variáveis ξ_λ dependem apenas das posições relativas das partículas¹ de modo que a função $\varphi_{E,\vec{K}}(\xi_\lambda)$ seja independente do vetor de posição do centro de massa, \vec{X} . Esse fator “intrínseco” é portanto também uma autofunção do momento total \vec{P} com autovalor zero. A invariância translacional não permite excluir, por si só, que ele dependa do momento total $\hbar\vec{K}$ do sistema. Isso de fato ocorre, no entanto, para hamiltonianos da forma $H = P^2/2M + H_{\text{int}}$, onde H_{int} é um hamiltoniano intrínseco que comuta com \vec{P} , que garante a invariância por transformações mais gerais que as do grupo das translações espaciais (transformações de Galileu).

O grupo das translações espaciais é um grupo comutativo cujas representações irreduzíveis são unidimensionais, como afirmado anteriormente com relação a tais grupos. Os operadores $R_{\vec{d}}$, eq. (9.5), que constituem a representação desse grupo no espaço de Hilbert das funções de onda $\psi(\vec{q}_\nu)$ são operadores *unitários*, e a base ortonormal que define as representações irreduzíveis unidimensionais é formada por autofunções do momento total (além de outras variáveis dinâmicas que sejam necessárias para a constituição de um conjunto completo de observáveis compatíveis). De fato, para qualquer vetor de base que seja autofunção do momento total vale $R_{\vec{d}}\phi_{\vec{K}} = e^{i\vec{d}\cdot\vec{K}}\phi_{\vec{K}}$, de modo que os elementos da representação correspondem simplesmente a multiplicação pelos fatores de fase $e^{i\vec{d}\cdot\vec{K}}$, que são na realidade matrizes unitárias de dimensão 1, com elementos de matriz nulos entre vetores distintos da base.

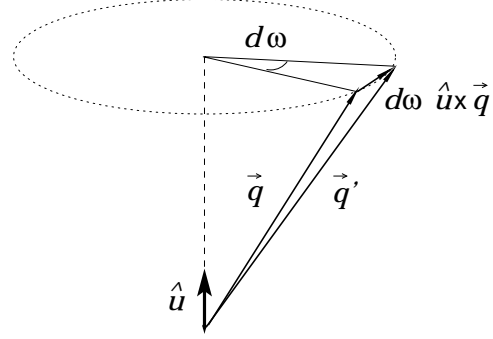
9.3 Rotações, momento angular e spin.

O grupo de rotações em três dimensões é mais complexo que o grupo de translações, discutido na seção anterior, por ser um grupo não comutativo. De fato, é bem conhecido o fato de que rotações espaciais finitas em geral não comutam, isto é, produzem resultados diferentes quando executadas em ordens diferentes. Como o grupo de translações, no entanto, este é um grupo contínuo, cujos elementos podem ser convenientemente expressos em termos de geradores.

Os geradores do grupo de rotações em três dimensões podem ser obtidos considerando inicialmente os efeitos para um vetor de posição \vec{q} de uma rotação de um ângulo infinitesimal

¹Uma escolha possível para as variáveis intrínsecas é $\vec{\xi}_\nu = \vec{q}_\nu - \vec{X}$, isto é as posições das partículas tomadas com relação ao centro de massa. É fácil ver que a transformação $\vec{q}_\nu \rightarrow \vec{q}_\nu + \vec{d}$ não altera os vetores $\vec{\xi}_\nu$, assim definidos. Esses vetores não são, porém, independentes, pois da sua definição segue que $\sum_\nu m_\nu \vec{\xi}_\nu = 0$.

Figura 9.1: Rotação $d\vec{\omega} = \hat{u} d\omega$ que leva o vetor \vec{q} ao vetor \vec{q}' .



$d\omega$ em torno de um eixo definido pelo vetor unitário \hat{u} . Para isso é conveniente descrever a rotação em termos do vetor

$$d\vec{\omega} \equiv \hat{u} d\omega$$

em termos do qual o vetor resultante \vec{q}' pode ser escrito da forma simples (v. figura 9.1)

$$\vec{q}' = \vec{q} + d\vec{\omega} \times \vec{q} + \mathcal{O}(d\omega^2).$$

A partir dessa expressão é fácil obter os efeitos de uma rotação $r_{d\vec{\omega}}$ *do referencial* (o que corresponde à descrição de tipo passivo) sobre as componentes cartesianas $\{q_j\}$, $j = 1, 2, 3$ do vetor \vec{q} . Eles podem ser escritos

$$q'_j = r_{d\vec{\omega}} q_j = q_j - \sum_{kl} \epsilon_{jkl} d\omega_k q_l + \mathcal{O}(d\omega^2) \quad (9.7)$$

onde ϵ_{jkl} é o símbolo antissimétrico, definido como sendo igual a $+1$ ou -1 quando jkl forem respectivamente uma permutação par ou ímpar de 123 e zero caso dois ou mais índices forem iguais, e $d\vec{\omega} \equiv \{d\omega_1, d\omega_2, d\omega_3\}$. Explicitamente

$$\begin{pmatrix} q'_1 \\ q'_2 \\ q'_3 \end{pmatrix} = \begin{pmatrix} 1 & d\omega_3 & -d\omega_2 \\ -d\omega_3 & 1 & d\omega_1 \\ d\omega_2 & -d\omega_1 & 1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} + \mathcal{O}(d\omega^2)$$

que se reduz à relação (9.1) no caso em que $\hat{u} = \hat{3}$ e portanto $d\omega_3 = d\omega$, com $d\omega_1 = d\omega_2 = 0$. Outras especializações do versor \hat{u} permitem identificar os geradores de rotações em torno dos eixos 1 e 2 respectivamente como

$$\lambda_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad \text{e} \quad \lambda_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \quad (9.8)$$

O passo seguinte consiste em obter representações dessas rotações infinitesimais no espaço de Hilbert de funções de onda $\psi(\vec{q}_\nu)$ que descrevem estados de um sistema de partículas *sem spin*. Isso se faz seguindo o esquema padrão resumido na equação (9.2), que neste caso é

$$\psi'(\vec{q}_\nu) = (R_{d\vec{\omega}}\psi)(\vec{q}_\nu) = \psi(r_{d\vec{\omega}}^{-1}\vec{q}_\nu).$$

Como a transformação inversa $r_{d\vec{\omega}}^{-1}$ é simplesmente a rotação infinitesimal $r_{-d\vec{\omega}}$, essa relação pode ser escrita explicitamente em termos das componentes cartesianas $q_{\nu j}$, $j = 1, 2, 3$ dos vetores de posição \vec{q}_ν como

$$\begin{aligned} (R_{d\vec{\omega}}\psi)(\vec{q}_\nu) &= \psi\left(q_{\nu j} + \sum_{kl} \epsilon_{jkl} d\omega_k q_{\nu l} + \mathcal{O}(d\omega^2)\right) \\ &= \left(1 + \sum_{\nu j} \sum_{kl} \epsilon_{jkl} d\omega_k q_{\nu l} \frac{\partial}{\partial q_{\nu j}}\right) \psi(\vec{q}_\nu) + \mathcal{O}(d\omega^2) \\ &= \left(1 + d\vec{\omega} \cdot \sum_{\nu} \vec{q}_\nu \times \nabla_{\vec{q}_\nu}\right) \psi(\vec{q}_\nu) + \mathcal{O}(d\omega^2). \end{aligned}$$

Na última passagem acima foi usada a expressão geral para o produto vetorial em termos do símbolo antissimétrico ϵ_{ijk}

$$\sum_{kl} \epsilon_{jkl} a_k b_l = (\vec{a} \times \vec{b})_j$$

bem como o fato de que o símbolo antissimétrico não se altera por permutações cíclicas dos seus índices. O operador diferencial $\nabla_{\vec{q}_\nu}$ pode ser expresso em termos dos momentos \vec{p}_ν , de modo que o operador $R_{d\vec{\omega}}$ que representa a rotação infinitesimal pode ser finalmente identificado como sendo

$$R_{d\vec{\omega}} = \hat{1} + \frac{i}{\hbar} d\vec{\omega} \cdot \sum_{\nu} (\vec{q}_\nu \times \vec{p}_\nu) = \hat{1} + \frac{i}{\hbar} d\vec{\omega} \cdot \vec{L}$$

onde

$$\vec{L} = \sum_{\nu} (\vec{q}_\nu \times \vec{p}_\nu)$$

é o momento angular total do sistema de partículas sem spin.

O gerador da representação das rotações espaciais no espaço de Hilbert considerado é portanto essencialmente o operador \vec{L} que corresponde ao momento angular orbital total do sistema.

Esse é um operador vetorial com três componentes em termos das quais é possível escrever o operador que representa rotações em torno de qualquer eixo. O gerador de uma rotação em torno do eixo \hat{u} é a componente $L_u = \vec{L} \cdot \hat{u}$, e rotações *finitas* $\vec{\omega} = \hat{u}\omega$ em torno do eixo \hat{u} podem ser escritas em forma exponencial usando

$$R_{\vec{\omega}} = \lim_{n \rightarrow \infty} \left(\hat{1} + \frac{i}{\hbar} \frac{\vec{\omega}}{n} \cdot \vec{L} \right)^n = e^{\frac{i}{\hbar} \vec{\omega} \cdot \vec{L}} = e^{\frac{i}{\hbar} \omega L_u}.$$

O caráter não comutativo do grupo de rotações se manifesta claramente, por exemplo, nas relações de comutação das componentes cartesianas do gerador \vec{L}

$$[L_j, L_k] = i\hbar \sum_l \epsilon_{jkl} L_l,$$

pois elas implicam que rotações *finitas* em torno de eixos diferentes em geral não comutam, isto é, em geral

$$e^{\frac{i}{\hbar} \omega_1 \hat{u} \cdot \vec{L}} e^{\frac{i}{\hbar} \omega_2 \hat{v} \cdot \vec{L}} \neq e^{\frac{i}{\hbar} \omega_2 \hat{v} \cdot \vec{L}} e^{\frac{i}{\hbar} \omega_1 \hat{u} \cdot \vec{L}}$$

se $\hat{u} \neq \hat{v}$.

A invariança rotacional de um sistema de partículas sem spin cuja dinâmica é determinada pelo hamiltoniano H é expressa pela condição $[H, R_{\vec{\omega}}] = 0$ para qualquer $\vec{\omega}$, o que significa qualquer eixo e também qualquer ângulo. Essa condição pode ser expressa em termos do gerador \vec{L} como

$$[H, \vec{L}] = 0.$$

Tendo em conta as relações de comutação das componentes de \vec{L} , isso leva à situação de que uma das componentes de \vec{L} (usualmente tomada como sendo L_3) e o operador $L^2 = L_1^2 + L_2^2 + L_3^2$, que é compatível com qualquer das componentes (e portanto com L_3), podem ser diagonalizados simultaneamente com H . Em outras palavras, os operadores compatíveis L^2 e L_3 são constantes do movimento e existe um conjunto completo de autovetores simultâneos de H , L^2 e L_3 cujos elementos podem ser designados como $|ELM\rangle$. Além disso, como

$$[H, L_{\pm}] = 0 \quad \text{e} \quad L_{\pm} |ELM\rangle \propto |ELM \pm 1\rangle$$

segue que

$$HL_{\pm} |ELM\rangle = L_{\pm} H |ELM\rangle = L_{\pm} E |ELM\rangle = EL_{\pm} |ELM\rangle.$$

Portanto o autovalor de H não depende do valor de M , e os níveis de energia tem pelo menos a degenerescência $2L + 1$ associada ao momento angular.

É possível ainda ter situações com simetria mais restrita como, por exemplo, sistemas cujo hamiltoniano é invariante por rotações em torno de um eixo \hat{u} dado, apenas. Isso significa que o sistema tem simetria axial em torno desse eixo, embora não sendo invariante por rotações mais gerais. Nesse caso o que se tem é

$$[H, R_{\hat{u}\omega}] = 0$$

para qualquer ângulo ω , mas para \hat{u} fixo. Então

$$[H, e^{\frac{i}{\hbar}\omega L_u}] = 0$$

para qualquer ω , e portanto

$$[H, L_u] = 0$$

isto é, a projeção L_u de \vec{L} ao longo do eixo de simetria é uma constante do movimento. O autovalor de H agora depende, em geral, do autovalor de L_u devido à não comutatividade de H com as outras componentes de \vec{L} .

9.3.1 Spin.

No caso de um sistema de partículas sem spin, o momento angular *orbital* total \vec{L} é na realidade o também momento angular *total* do sistema. Como esse é o operador que aparece como gerador da representação do grupo de rotações no espaço de Hilbert correspondente, é natural que se *defina*, o operador associado ao momento angular total \vec{J} de um sistema *em geral*, como o gerador da representação do grupo das rotações espaciais no espaço de Hilbert dos estados desse sistema. Mais formalmente, o momento angular total \vec{J} de um sistema cujos estados são representados pelos vetores de estado $|\psi\rangle$ pode ser *definido* como sendo o operador que aparece na representação de rotações espaciais em termos desses vetores de estado

$$|\psi'\rangle = R_{\vec{\omega}}|\psi\rangle = e^{\frac{i}{\hbar}\vec{\omega}\cdot\vec{J}}|\psi\rangle. \quad (9.9)$$

Essa definição evidentemente não envolve qualquer condição de simetria sobre o sistema. Ela no entanto utiliza propriedades de transformação dos estados para delas extrair a definição do operador a ser associado a uma variável dinâmica.

A natureza explícita do operador assim definido depende, é claro, da natureza dos estados do sistema considerado. No caso de partículas sem spin considerado anteriormente, os operadores

de posição $\{\vec{q}_\nu\}$ constituem um conjunto completo de observáveis compatíveis, os vetores de estado podem ser representados por funções de onda $\psi(\vec{q}_\nu)$ e, como visto, $\vec{J} = \vec{L}$. Se, por outro lado, outros operadores tiverem que ser considerados juntamente com os $\{\vec{q}_\nu\}$ para que se tenha um conjunto completo de observáveis compatíveis, as funções de onda que representam os vetores de estado envolverão variáveis adicionais $\{s_\mu\}$

$$\langle \vec{q}_\nu, s_\mu | \psi \rangle = \psi(\vec{q}_\nu, s_\mu)$$

e o gerador \vec{J} diferirá de \vec{L} sempre que os graus de liberdade adicionais contribuirão para as propriedades de transformação dos estados (ou das funções de onda) do sistema sob rotações espaciais. A relação (9.9) deve ser implementada nesse caso com um gerador da forma

$$\vec{J} = \vec{L} \otimes \hat{1}_S + \hat{1}_L \otimes \vec{S} \quad (9.10)$$

onde $\vec{L} = \sum_\nu (\vec{q}_\nu \times \vec{p}_\nu)$ e o segundo termo, age não trivialmente apenas sobre os graus de liberdade adicionais. Ele corresponde à componente do momento angular total que está associada com esses graus de liberdade, e é portanto o *spin total* do sistema. Como os dois termos desta última expressão agem não trivialmente sobre grupos distintos de graus de liberdade é claro que

$$[\vec{L} \otimes \hat{1}_S, \hat{1}_L \otimes \vec{S}] = 0 \quad \text{ou simplesmente} \quad [\vec{L}, \vec{S}] = 0.$$

Para que a relação (9.9) com um gerador com a forma (9.10) seja de fato uma relação bem construída é preciso que o operador \vec{S} tenha propriedades congruentes com as do operador \vec{L} . Em particular, ele deve ser também um operador vetorial cujas componentes se comportam sob rotações da mesma forma que as componentes correspondentes de \vec{L} . O comportamento das componentes cartesianas de um vetor para uma rotação passiva $d\vec{\omega}$ é dado pela relação (9.7), e essa relação deve portanto valer não só para as componentes de \vec{L} como também para as do operador ainda por determinar \vec{S} . Chamando respectivamente S_j e S'_j as componentes desse vetor no sistema original e no sistema resultante da rotação $d\vec{\omega}$, deve-se ter portanto

$$S'_j = S_j - \sum_{kl} \epsilon_{jkl} d\omega_k S_l + \mathcal{O}(d\omega^2).$$

Por outro lado, se \vec{S} é uma variável dinâmica do sistema para o qual os efeitos da rotação são descritos pelos operadores de rotação $R_{\vec{\omega}}$, equações (9.9) e (9.10), deve valer também a relação de transformação geral para variáveis dinâmicas

$$S'_j = e^{\frac{i}{\hbar} d\vec{\omega} \cdot \vec{S}} S_j e^{-\frac{i}{\hbar} d\vec{\omega} \cdot \vec{S}} = \left(1 + \frac{i}{\hbar} d\vec{\omega} \cdot \vec{S} + \dots\right) S_j \left(1 - \frac{i}{\hbar} d\vec{\omega} \cdot \vec{S} + \dots\right)$$

$$\begin{aligned}
&= S_j + \frac{i}{\hbar} [d\vec{\omega} \cdot \vec{S}, S_j] + \mathcal{O}(d\omega^2) \\
&= S_j + \frac{i}{\hbar} \sum_k d\omega_k [S_k, S_j] + \mathcal{O}(d\omega^2).
\end{aligned}$$

Para que essas duas transformações sejam consistentes é preciso portanto que

$$-\sum_{kl} \epsilon_{jkl} d\omega_k S_l = \frac{i}{\hbar} \sum_k d\omega_k [S_k, S_j]$$

ou seja

$$\sum_k d\omega_k \left([S_k, S_j] + i\hbar \sum_l \epsilon_{jkl} S_l \right) = 0.$$

Essa relação deve valer para qualquer $d\vec{\omega}$, e portanto o parênteses deve ser nulo. Para isso é preciso que se tenha

$$[S_j, S_k] = i\hbar \sum_l \epsilon_{jkl} S_l,$$

isto é, as componentes cartesianas de \vec{S} devem satisfazer as mesmas relações de comutação que as do momento angular orbital \vec{L} . As mesmas relações de comutação são portanto satisfeitas também pelas componentes do momento angular total \vec{J} .

As conseqüências gerais dessas relações de comutação já foram obtidas na seção 5.1.1. Os autovalores de S^2 são da forma $\hbar^2 S(S+1)$, o fator \hbar^2 tendo sido introduzido por razões dimensionais, com S inteiro ou semi-inteiro. Esse operador é compatível com qualquer das componentes S_i (por exemplo S_3) cujos autovalores são $\hbar M_S$ com $-S \leq M_S \leq S$. Valores semi-inteiros do momento angular *orbital* são excluídos pela forma explícita do operador \vec{L} , que é dada em termos dos operadores de momento e posição, mas podem ocorrer no caso de \vec{S} e de \vec{J} . As relações entre os autovalores de \vec{L} , \vec{S} e \vec{J} são determinadas pelas regras relativas à soma de momentos angulares, discutidas na seção 6.2.1.

A condição geral para a invariância rotacional do hamiltoniano, obtida de forma inteiramente análoga à que foi utilizada para o caso de sistemas sem spin, é dada por

$$[H, \vec{J}] = 0$$

o que, em vista das relações de comutação dos operadores envolvidos, implica na compatibilidade de H com L^2 , S^2 , J^2 e J_3 , bem como na independência do autovalor de H com relação ao

autovalor $\hbar M$ de J_3 . Isso leva a uma degenerescência de ordem pelo menos $2J + 1$ dos níveis de energia com momento angular J . Essa degenerescência é uma consequência característica da invariança rotacional.

O sistema mais simples para o qual existe uma contribuição puramente quântica para o momento angular é constituído por uma partícula de spin $1/2$ (v. seção 6.2). Nesse caso, consistentemente com as condições agora obtidas a partir da análise dos efeitos das rotações espaciais, os operadores de spin são usualmente representados em termos das matrizes de Pauli, $\vec{S} \equiv \hbar \vec{\sigma}/2$ e os estados podem ser descritos por funções de onda com duas componentes

$$\langle \vec{r}, s | \psi \rangle = \psi(\vec{r}, s) \equiv \begin{pmatrix} \psi(\vec{r}, s = +1/2) \\ \psi(\vec{r}, s = -1/2) \end{pmatrix}$$

onde os valores $s = \pm 1/2$ da variável de spin estão usualmente associados aos dois autovalores de S_3 . No espaço de Hilbert dessas funções de onda o operador que representa uma rotação (passiva) de um ângulo ω em torno de uma direção $\hat{u} \equiv \{u_1, u_2, u_3\}$ é, portanto,

$$\begin{aligned} R_{\omega \hat{u}} = e^{i\omega \hat{u} \cdot \vec{J}} &= e^{i\omega \hat{u} \cdot \vec{L}} e^{i\frac{\omega}{2} \hat{u} \cdot \vec{\sigma}} = e^{i\omega \hat{u} \cdot \vec{L}} \left(\hat{1} \cos \frac{\omega}{2} + i \hat{u} \cdot \vec{\sigma} \sin \frac{\omega}{2} \right) = \\ &= e^{i\omega \hat{u} \cdot \vec{L}} \begin{pmatrix} \cos \frac{\omega}{2} + i u_3 \sin \frac{\omega}{2} & i(u_1 - i u_2) \sin \frac{\omega}{2} \\ i(u_1 + i u_2) \sin \frac{\omega}{2} & \cos \frac{\omega}{2} - i u_3 \sin \frac{\omega}{2} \end{pmatrix}. \end{aligned}$$

A exponencial contendo o operador do momento angular orbital \vec{L} age sobre a dependência com \vec{r} da função de onda, enquanto a matriz correspondente à parte de spin do operador de rotação recombina as duas componentes da função de onda de partida. Desse modo,

$$\psi'(\vec{r}, s) = R_{\omega \hat{u}} \psi(\vec{r}, s) = \begin{pmatrix} \cos \frac{\omega}{2} + i u_3 \sin \frac{\omega}{2} & i(u_1 - i u_2) \sin \frac{\omega}{2} \\ i(u_1 + i u_2) \sin \frac{\omega}{2} & \cos \frac{\omega}{2} - i u_3 \sin \frac{\omega}{2} \end{pmatrix} \begin{pmatrix} e^{i\omega \hat{u} \cdot \vec{L}} \psi(\vec{r}, s = +1/2) \\ e^{i\omega \hat{u} \cdot \vec{L}} \psi(\vec{r}, s = -1/2) \end{pmatrix}.$$

Em particular, para uma rotação em torno do eixo 3, $u_3 = 1$ e $u_1 = u_2 = 0$ e essa expressão se reduz a

$$\psi'(\vec{r}, s) = \begin{pmatrix} e^{i\frac{\omega}{2}} e^{i\omega \hat{u} \cdot \vec{L}} \psi(\vec{r}, s = +1/2) \\ e^{-i\frac{\omega}{2}} e^{i\omega \hat{u} \cdot \vec{L}} \psi(\vec{r}, s = -1/2) \end{pmatrix}$$

que, para $\omega = 2\pi$, dá

$$\psi'(\vec{r}, s) = -\psi(\vec{r}, s),$$

que reproduz o resultado obtido anteriormente², segundo o qual uma rotação de um ângulo 2π leva a uma troca de sinal da função de onda da partícula de spin $1/2$.

O caso de spin maior que $1/2$ pode ser tratado de forma semelhante, notando que os estados de uma partícula com spin S são representados por funções de onda com $2S + 1$ componentes associadas (por exemplo) aos diferentes autovalores do operador S_3 . A primeira questão a ser considerada é portanto a de determinar uma representação com a dimensão apropriada para os operadores de spin \vec{S} . Tomando como exemplo o caso $S = 1$, o problema consiste em determinar três matrizes (a serem associadas às componentes do operador \vec{S}), de dimensão $2S + 1 = 3$, que satisfaçam as relações de comutação de momento angular. Esse problema pode ser resolvido considerando as matrizes λ_j , $j = 1, 2, 3$, obtidas no contexto das propriedades de transformação sob rotações das componentes de um vetor no espaço de posições, (9.1) e (9.8). De fato, é imediato verificar que essas três matrizes satisfazem as relações de comutação

$$[\lambda_j, \lambda_k] = -\sum_l \epsilon_{jkl} \lambda_l$$

de modo que uma escolha possível para as componentes de \vec{S} nesse caso é

$$S_j \equiv -i\hbar \lambda_j.$$

Essas matrizes são hermiteanas, e foram obtidas em termos de uma base cartesiana no espaço de posições. Mais explicitamente, as componentes cartesianas de um vetor (por exemplo, o vetor de posição \vec{q}) se transformam sob uma rotação infinitesimal $d\vec{\omega}$ de acordo com a relação

$$q'_j = q_j + \sum_k (d\vec{\omega} \cdot \vec{\lambda})_{jk} q_k = q_j + \frac{i}{\hbar} \sum_k (d\vec{\omega} \cdot \vec{S})_{jk} q_k.$$

Nessa particular representação nenhuma das três matrizes é diagonal. É imediato, no entanto, diagonalizar a matriz hermiteana S_3 , cujos autovalores são $\hbar M_S$, $M_S = -1, 0, 1$ os quais correspondem aos autovetores u_{M_S} dados respectivamente por

$$u_{\pm 1} = \mp \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \\ 0 \end{pmatrix}, \quad u_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

²V. página 291.

que correspondem à base considerada, por exemplo, na equação (8.32). Nessa nova base as matrizes S_j são dadas por

$$S_1 = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_2 = \frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_3 = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

para as quais as relações de comutação podem ser mais uma vez verificadas explicitamente.

Esta forma de obter uma representação para o operador de spin no caso $S = 1$ mostra, por outro lado, que as três componentes de spin de uma função de onda $\psi(\vec{r}, s)$, $s = -1, 0, 1$ que descreve o estado de uma partícula de spin 1 se misturam entre si sob uma rotação espacial da mesma forma que as componentes de um *campo vetorial*, e que quando a variável de spin s está associada aos autovalores de S_3 , as componentes $\psi(\vec{r}, s)$ correspondem às componentes ditas *esféricas* (8.32) do campo vetorial.

9.3.2 Representação de rotações gerais em termos de ângulos de Euler.

Uma rotação finita *qualquer* de um sistema de eixos $\{1, 2, 3\}$ que mantém fixa a origem O pode ser sempre descrita como uma rotação de um ângulo dado ω em torno de um eixo também dado \hat{u} que passa por O . Ela é portanto especificada em termos de três parâmetros que são, nesse caso, os dois ângulos que definem o versor \hat{u} no sistema inicial e mais o ângulo de rotação ω . Uma forma alternativa de caracterizar as rotações que é de uso mais comum consiste parametrizá-la em termos de outros três ângulos $\{\alpha, \beta, \gamma\}$, chamados ângulos de Euler, definidos da seguinte forma (v. figura 9.2). A passagem do sistema inicial $\{1, 2, 3\}$ para o sistema final $\{1''', 2''', 3'''\}$ é feita através das seguintes três operações:

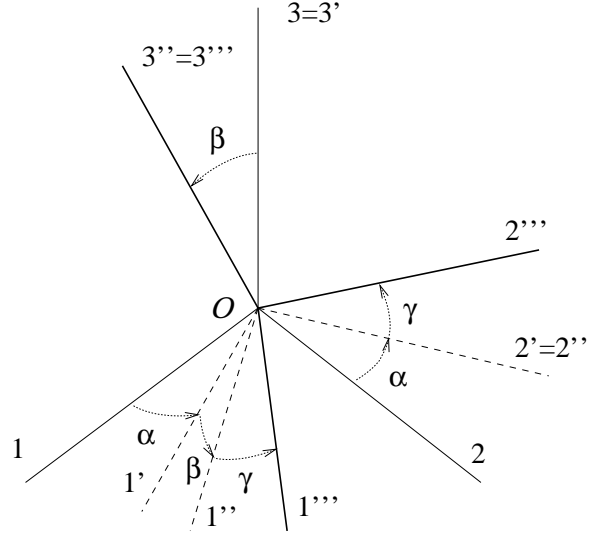
- a) uma rotação de um ângulo α em torno do eixo 3, que leva o sistema inicial ao sistema $3' = 3$, $1'$ e $2'$;
- b) uma rotação de um ângulo β em torno do eixo $2'$ (isto é, em torno da nova posição do eixo 2), que leva o sistema de $\{1', 2', 3'\}$ para $\{1'', 2'' = 2', 3''\}$; e finalmente
- c) uma rotação de um ângulo γ em torno do eixo $3''$, pela qual o sistema $\{1'', 2'', 3''\}$ é levado a $\{1''', 2''', 3''' = 3''\}$.

Desse modo, a rotação $\vec{\omega} \equiv \omega \hat{u}$ é decomposta em três rotações, que podem ser resumidas como

$$\vec{\omega} \longrightarrow \{\alpha \hat{3}, \beta \hat{2}', \gamma \hat{3}''\}.$$

Essas três rotações são aplicadas nessa ordem (da esquerda para a direita) e no sentido anti-horário, em torno dos eixos especificados.

Figura 9.2: Ângulos de Euler $\{\alpha, \beta, \gamma\}$ que levam o sistema de eixos $\{1, 2, 3\}$ a $\{1''', 2''', 3'''\}$. A rotação do ângulo α é feita em torno do eixo 1; a rotação do ângulo β é feita em torno de $2'$ e a rotação do ângulo γ é feita em torno de $3''$.



Por envolver a especificação de três rotações de ângulos apropriados α, β, γ em torno de três eixos distintos, a parametrização de rotações em termos de ângulos de Euler pode parecer mais complicada, e talvez por isso menos conveniente, que a parametrização alternativa em termos de uma única rotação ω em torno de um eixo apropriado \hat{u} . A conveniência da parametrização de Euler é melhor ilustrada com relação ao problema de determinar a rotação que leva o sistema de eixos de partida $\{1, 2, 3\}$ a um sistema final *dado* $\{1''', 2''', 3'''\}$. Para isso é útil notar que a passagem do eixo 3 ao novo eixo $3'''$ é feita através da etapa **b**), que envolve a rotação do ângulo β em torno do eixo $2'$. Esse ângulo β nada mais é, portanto, que o ângulo formado entre os eixos 3 e $3'''$, e o eixo $2'$ para essa rotação deve ser perpendicular ao plano formado por eles. O propósito da etapa **a**), portanto, é levar o eixo 2 à posição $2'$, perpendicular ao plano formado por 3 e $3'''$, e essa é a condição que determina o ângulo α . Uma vez corretamente posicionado o eixo $3'''$, as posições de $1'''$ e $2'''$ podem ser acertadas através da terceira rotação γ . Como pode ser visto, em cada uma das etapas o eixo de rotação é conhecido (eventualmente como resultado da implementação da etapa anterior), e o ângulo de rotação correspondente é determinado por uma condição simples.

Embora levando a um resultado que pode ser visto como mais compacto, a determinação dos elementos da parametrização alternativa para o mesmo problema envolve construções geométricas menos transparentes. De fato, a determinação de \hat{u} e ω pode ser feita do seguinte modo. Em primeiro lugar é preciso observar que, para uma rotação de um ângulo não nulo, pelo menos dois dos novos eixos $\{1''', 2''', 3'''\}$ serão diferentes dos eixos de partida. Sejam então $1''' \neq 1$ e $2''' \neq 2$ (o eixo $3'''$ podendo coincidir *ou não* com o eixo 3). Todos os eixos de rotação capazes de levar 1 a $1'''$ estão contidos no plano, perpendicular ao plano definido por esses dois eixos, que passa pela bissetriz do ângulo

formado por eles. Uma construção análoga vale para os eixos 2 e 2'''. O eixo de rotação \hat{u} é portanto definido pela intersecção dos dois planos assim construídos. Caso também $3''' \neq 3$, a mesma construção aplicada a esse terceiro par de eixos dará um terceiro plano que contém a intersecção dos dois outros planos (caso $3''' = 3$ a intersecção dos dois primeiros planos se dará ao longo de 3, que coincidirá portanto com o eixo de rotação \hat{u}). O ângulo de rotação ω pode então ser determinado a partir da posição relativa das direções \hat{u} , 1 e 1''' (ou \hat{u} , 2 e 2'''), em geral mutuamente oblíquas.

O operador $R_{\vec{\omega}}$ que representa a rotação $\vec{\omega}$ no espaço de Hilbert dos estados de um sistema quântico dado é portanto expresso em termos dos ângulos de Euler como

$$R_{\vec{\omega}} = e^{\frac{i}{\hbar}\vec{\omega}\cdot\vec{J}} = e^{\frac{i}{\hbar}\gamma\hat{3}''\cdot\vec{J}}e^{\frac{i}{\hbar}\beta\hat{2}'\cdot\vec{J}}e^{\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}} = R_{\gamma\hat{3}''}R_{\beta\hat{2}'}R_{\alpha\hat{3}} \equiv R(\alpha, \beta, \gamma). \quad (9.11)$$

Essa expressão pode ser simplificada usando, por um lado, o fato de que o eixo $\hat{2}'$ que é usado na segunda das três rotações é obtido aplicando a rotação $\alpha\hat{3}$ ao eixo $\hat{2}$, de modo que

$$R_{\beta\hat{2}'} = R_{\alpha\hat{3}}R_{\beta\hat{2}}R_{\alpha\hat{3}}^{-1} = e^{\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}}e^{\frac{i}{\hbar}\beta\hat{2}\cdot\vec{J}}e^{-\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}}. \quad (9.12)$$

Por outro lado, a rotação que envolve o ângulo γ pode ser escrita, por motivos análogos, como

$$\begin{aligned} R_{\gamma\hat{3}''} &= R_{\beta\hat{2}'}R_{\gamma\hat{3}}R_{\beta\hat{2}'}^{-1} = (R_{\alpha\hat{3}}R_{\beta\hat{2}}R_{\alpha\hat{3}}^{-1})R_{\gamma\hat{3}}(R_{\alpha\hat{3}}R_{\beta\hat{2}}^{-1}R_{\alpha\hat{3}}^{-1}) = \\ &= R_{\alpha\hat{3}}R_{\beta\hat{2}}R_{\gamma\hat{3}}R_{\beta\hat{2}}^{-1}R_{\alpha\hat{3}}^{-1} = e^{\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}}e^{\frac{i}{\hbar}\beta\hat{2}\cdot\vec{J}}e^{\frac{i}{\hbar}\gamma\hat{3}\cdot\vec{J}}e^{-\frac{i}{\hbar}\beta\hat{2}\cdot\vec{J}}e^{-\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}} \end{aligned} \quad (9.13)$$

onde foi usada a comutatividade de rotações em torno de um mesmo eixo. Levando agora os resultados (9.12) e (9.13) à relação (9.11) resulta, após cancelamentos óbvios,

$$R(\alpha, \beta, \gamma) = e^{\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}}e^{\frac{i}{\hbar}\beta\hat{2}\cdot\vec{J}}e^{\frac{i}{\hbar}\gamma\hat{3}\cdot\vec{J}} = R_{\alpha\hat{3}}R_{\beta\hat{2}}R_{\gamma\hat{3}} \quad (9.14)$$

onde intervêm apenas rotações em torno dos eixos *originais* 2 e 3. A ordem em que aparecem os três ângulos de Euler é no entanto *inversa* à ordem da equação de partida, (9.11).

É bastante comum adotar o ponto de vista *ativo* para escrever os operadores de rotação. Dentro desse ponto de vista, os ângulos de Euler correspondem a três rotações ativas que levam o sistema físico de sua orientação inicial à sua orientação final. Rotações ativas em torno de um eixo diferem das passivas correspondentes por uma mudança de sinal do ângulo de rotação. Desse modo, a expressão que corresponde à (9.14) é

$$R^{(\text{at})}(\alpha, \beta, \gamma) = e^{-\frac{i}{\hbar}\alpha\hat{3}\cdot\vec{J}}e^{-\frac{i}{\hbar}\beta\hat{2}\cdot\vec{J}}e^{-\frac{i}{\hbar}\gamma\hat{3}\cdot\vec{J}} = R_{\alpha\hat{3}}^{(\text{at})}R_{\beta\hat{2}}^{(\text{at})}R_{\gamma\hat{3}}^{(\text{at})}.$$

9.3.3 Matrizes de rotação, tensores esféricos, teorema de Wigner-Eckart.

As matrizes que descrevem os operadores de rotação $R(\alpha, \beta, \gamma)$ em uma base qualquer do espaço de Hilbert dos estados do sistema quântico que esteja sendo considerado constituem também, como já dito, uma representação (em geral redutível) do grupo de rotações. Uma classe particularmente importante de representações matriciais desse tipo é a que se obtém quando a base é formada por autovetores simultâneos de J^2 e de J_3 , além de outros observáveis que sejam necessários para que se tenha um conjunto completo de observáveis compatíveis. Designando os vetores dessa base como $|\lambda JM\rangle$, onde λ denota coletivamente os números quânticos, além de J e M , que sejam necessários para especificar univocamente os vetores de estado,

$$J^2|\lambda JM\rangle = \hbar^2 J(J+1)|\lambda JM\rangle, \quad J_3|\lambda JM\rangle = \hbar M|\lambda JM\rangle.$$

Relações de completeza e ortonormalidade para a base são expressas respectivamente através de

$$\sum_{\lambda JM} |\lambda JM\rangle \langle \lambda JM| = \hat{1} \quad \text{e} \quad \langle \lambda JM | \lambda' J' M' \rangle = \delta_{\lambda\lambda'} \delta_{JJ'} \delta_{MM'}.$$

O efeito do operador de rotação $R(\alpha, \beta, \gamma)$ sobre os vetores dessa base pode ser re-expandido na mesma base dando

$$R(\alpha, \beta, \gamma)|\lambda JM\rangle = \sum_{\lambda' J' M'} |\lambda' J' M'\rangle \langle \lambda' J' M' | R(\alpha, \beta, \gamma) | \lambda JM \rangle$$

que identifica os elementos da matriz de transformação relevante como sendo

$$\langle \lambda' J' M' | R(\alpha, \beta, \gamma) | \lambda JM \rangle = \langle \lambda JM' | R(\alpha, \beta, \gamma) | \lambda JM \rangle \delta_{\lambda\lambda'} \delta_{JJ'}.$$

O fato de que essa matriz é diagonal em J e nos números quânticos λ resulta de que o operador de rotação é uma função de componentes do momento angular total \vec{J} que comutam tanto com J^2 como com os outros observáveis do conjunto completo utilizado para determinação unívoca dos estados da base. Além disso, os elementos de matriz são *independentes de* λ , o que decorre das expressões explícitas para os elementos de matriz das componentes de \vec{J}

$$\langle \lambda JM' | J_3 | \lambda JM \rangle = \hbar M \delta_{MM'}$$

e

$$\langle \lambda JM' | (J_1 \pm iJ_2) | \lambda JM \rangle = \sqrt{J(J+1) - M(M \pm 1)} \delta_{M \pm 1 M'}.$$

Os efeitos do operador de rotação sobre os elementos da base considerada podem portanto ser escritos sob a forma

$$R(\alpha, \beta, \gamma)|\lambda JM\rangle = \sum_{M'} |\lambda JM'\rangle D_{M'M}^J(\alpha, \beta, \gamma) \quad (9.15)$$

onde, para cada um dos valores possíveis de J ,

$$D_{M'M}^J(\alpha, \beta, \gamma) \equiv \langle \lambda JM' | R(\alpha, \beta, \gamma) | \lambda JM \rangle$$

são matrizes de dimensão $2J + 1$, dependentes dos ângulos de Euler. A expressão explícita do operador de rotação em termos das componentes de \vec{J} permite escrever ainda

$$D_{M'M}^J(\alpha, \beta, \gamma) = \langle \lambda JM' | e^{\frac{i}{\hbar} \alpha \hat{J}_3} e^{\frac{i}{\hbar} \beta \hat{J}_2} e^{\frac{i}{\hbar} \gamma \hat{J}_3} | \lambda JM \rangle = e^{i(\alpha M' + \gamma M)} d_{MM'}^J(\alpha, \beta, \gamma)$$

com a definição

$$d_{MM'}^J(\alpha, \beta, \gamma) \equiv \langle \lambda JM' | e^{\frac{i}{\hbar} \beta \hat{J}_2} | \lambda JM \rangle.$$

As matrizes $D_{M'M}^J(\alpha, \beta, \gamma)$, vistas como funções dos ângulos de Euler para um valor dado de J , são *representações irredutíveis* de dimensão $2J + 1$ do grupo das rotações. De fato, aplicando um segundo operador de rotação $R(\alpha', \beta', \gamma')$ sobre a equação (9.15),

$$\begin{aligned} R(\alpha', \beta', \gamma') R(\alpha, \beta, \gamma) |\lambda JM\rangle &= \sum_{M'} R(\alpha', \beta', \gamma') |\lambda JM'\rangle D_{M'M}^J(\alpha, \beta, \gamma) = \\ &= \sum_{M''} |\lambda JM''\rangle D_{M''M'}^J(\alpha', \beta', \gamma') D_{M'M}^J(\alpha, \beta, \gamma). \end{aligned}$$

Usando a lei de composição do grupo, as duas rotações aplicadas sucessivamente equivalem ainda a uma única rotação $R(\bar{\alpha}, \bar{\beta}, \bar{\gamma})$, de modo que, usando mais uma vez a equação (9.15) resulta

$$\sum_{M''} |\lambda JM''\rangle D_{M''M}^J(\bar{\alpha}, \bar{\beta}, \bar{\gamma}) = \sum_{M''} |\lambda JM''\rangle D_{M''M'}^J(\alpha', \beta', \gamma') D_{M'M}^J(\alpha, \beta, \gamma)$$

de onde segue a lei de composição de grupo para as matrizes de rotação

$$D_{M''M}^J(\bar{\alpha}, \bar{\beta}, \bar{\gamma}) = \sum_{M'} D_{M''M'}^J(\alpha', \beta', \gamma') D_{M'M}^J(\alpha, \beta, \gamma).$$

O caráter *irredutível* da representação de dimensão $2J + 1$ dada por $D_{M'M}^J(\alpha, \beta, \gamma)$ significa que não existe nenhum sub-espço do espaço gerado pelos $2J + 1$ vetores $|\lambda JM\rangle$ (com λ e J fixos

e $-J \leq M \leq J$) que seja invariante sob a ação de todas as rotações possíveis. A existência de representações irreduzíveis de dimensão maior que 1 é uma consequência do caráter não comutativo do grupo das rotações. A *dimensão* de uma dada representação irreduzível corresponde à degenerescência característica que resulta da invariância rotacional do hamiltoniano, de ordem $2J + 1$ para um nível de energia com momento angular J .

Além da lei de composição de grupo, outra relação útil entre matrizes de rotação pode ser obtida considerando a transformação dos estados de um sistema constituído de dois subsistemas, aos quais estão associados os operadores de momento angular total $\vec{j}^{(1)}$ e $\vec{j}^{(2)}$ (por exemplo, como geradores das representações $R^{(1)}(\alpha, \beta, \gamma)$ e $R^{(2)}(\alpha, \beta, \gamma)$ do grupo das rotações nos espaços de Hilbert apropriados a cada um dos subsistemas). A base interessante do sistema composto é a base produto constituída dos autovetores simultâneos de $j^{(1)2}$, $j_3^{(1)}$, $j^{(2)2}$ e $j_3^{(2)}$, além de outros observáveis necessários para a identificação unívoca dos vetores $|\lambda j_1 m_1\rangle \otimes |\kappa j_2 m_2\rangle$. Uma base alternativa é aquela que resulta do acoplamento dos momentos angulares dos subsistemas, de modo que o quadrado do momento angular total $\vec{J} = \vec{j}^{(1)} + \vec{j}^{(2)}$ e sua terceira componente são diagonais. A relação entre essas duas bases é dada portanto em termos de coeficientes de Clebsch-Gordan como

$$|\lambda \kappa j_1 j_2 J M\rangle = \sum_{m_1 m_2} C_{m_1 m_2}^{j_1 j_2 J} |\lambda j_1 m_1\rangle \otimes |\kappa j_2 m_2\rangle.$$

Efetuada uma rotação definida pelos ângulos de Euler α, β, γ se obtém

$$R(\alpha, \beta, \gamma) |\lambda \kappa j_1 j_2 J M\rangle = \sum_{m_1 m_2} C_{m_1 m_2}^{j_1 j_2 J} R^{(1)}(\alpha, \beta, \gamma) |\lambda j_1 m_1\rangle \otimes R^{(2)}(\alpha, \beta, \gamma) |\kappa j_2 m_2\rangle$$

que pode ser re-escrita em termos das matrizes de rotação como

$$\begin{aligned} \sum_{M'} |\lambda \kappa j_1 j_2 J M'\rangle D_{M'M}^J(\alpha, \beta, \gamma) = \\ = \sum_{m'_1 m'_2} \sum_{m_1 m_2} C_{m_1 m_2}^{j_1 j_2 J} |\lambda j_1 m'_1\rangle \otimes |\kappa j_2 m'_2\rangle D_{m'_1 m_1}^{j_1}(\alpha, \beta, \gamma) D_{m'_2 m_2}^{j_2}(\alpha, \beta, \gamma). \end{aligned} \quad (9.16)$$

As propriedades de ortogonalidade da transformação descrita pelos coeficientes de Clebsch-Gordan permite escrever os estados da representação produto em termos dos estados acoplados como

$$|\lambda j_1 m'_1\rangle \otimes |\kappa j_2 m'_2\rangle = \sum_{J'} C_{m'_1 m'_2}^{j_1 j_2 J'} R^{(1)} |\lambda \kappa j_1 j_2 J' m'_1 + m'_2\rangle$$

com o que a relação (9.16), projetada sobre o vetor $|\lambda \kappa j_1 j_2 J M'\rangle$, dá

$$D_{M'M}^J(\alpha, \beta, \gamma) = \sum_{m'_1 m'_2} \sum_{m_1 m_2} C_{m'_1 m'_2 M'}^{j_1 j_2 J} C_{m_1 m_2 M}^{j_1 j_2 J} D_{m'_1 m_1}^{j_1}(\alpha, \beta, \gamma) D_{m'_2 m_2}^{j_2}(\alpha, \beta, \gamma) \quad (9.17)$$

que é a relação buscada entre as matrizes de rotação. As três matrizes de rotação que aparecem nessa expressão têm os mesmos argumentos. Os coeficientes de Clebsch-Gordan acoplam os momentos angulares parciais j_1 e j_2 dando o momento angular total J com os números quânticos magnéticos correspondentes respectivamente aos primeiros e segundos índices das matrizes de rotação. Usando novamente as propriedades de ortogonalidade da transformação descrita pelos coeficientes de Clebsch-Gordan essa mesma relação pode ainda ser transformada na relação equivalente

$$\sum_J \sum_{MM'} C_{m'_1 m'_2 M'}^{j_1 j_2 J} C_{m_1 m_2 M}^{j_1 j_2 J} D_{M'M}^J(\alpha, \beta, \gamma) = D_{m'_1 m_1}^{j_1}(\alpha, \beta, \gamma) D_{m'_2 m_2}^{j_2}(\alpha, \beta, \gamma). \quad (9.18)$$

Cabe observar ainda que, projetando a relação (9.16) sobre estados $|\lambda \kappa j_1 j_2 J' M'\rangle$ com $J' \neq J$, a contribuição do lado esquerdo é nula, levando portanto a combinações lineares nulas de produtos de matrizes de rotação do tipo

$$\sum_{m'_1 m'_2} \sum_{m_1 m_2} C_{m'_1 m'_2 M'}^{j_1 j_2 J'} C_{m_1 m_2 M}^{j_1 j_2 J} D_{m'_1 m_1}^{j_1}(\alpha, \beta, \gamma) D_{m'_2 m_2}^{j_2}(\alpha, \beta, \gamma) = 0, \quad J' \neq J.$$

As representações irredutíveis do grupo de rotações $D_{M'M}^J(\alpha, \beta, \gamma)$ permitem definir uma classe importante de objetos através de suas propriedades de transformação sob rotações. Os objetos dessa classe, chamados *tensores esféricos* ou *tensores de Racah* de ordem k , são constituídos de $2k + 1$ componentes T_{kq} , $q = -k, -k + 1, \dots, k - 1, k$ (com k inteiro ou semi-inteiro) que sob uma rotação definida pelos ângulos de Euler α, β, γ se transformam segundo a relação

$$T'_{kq} = \sum_{q'} T_{kq'} D_{q'q}^k.$$

O conjunto dos $2J + 1$ vetores de estado $|\lambda JM\rangle$, para λ e J fixos e $-J \leq M \leq J$ são, nesse sentido, um tensor de ordem J . Os objetos T_{kq} podem ser porém também operadores lineares sobre o espaço de Hilbert, caso em que os objetos transformados são definidos como

$$T'_{kq} \equiv R(\alpha, \beta, \gamma) T_{kq} R^{-1}(\alpha, \beta, \gamma).$$

Exemplos simples de operadores que têm propriedades tensoriais desse tipo são operadores invariantes por rotações, que são tensores esféricos de ordem zero, as componentes do momento angular definidas por (cf. equação (8.32))

$$J_{q=\pm 1} \equiv \mp \frac{1}{\sqrt{2}}(J_1 \pm iJ_2) \quad \text{e} \quad J_{q=0} = J_3$$

onde J_1, J_2 e J_3 são as componentes cartesianas de \vec{J} . As três componentes J_q constituem um tensor esférico de ordem 1, bem como as componentes análogas de qualquer outro operador vetorial.

A relevância desses objetos decorre de uma série de propriedades que eles tem em decorrência da sua relação com as representações irredutíveis do grupo de rotações. Uma dessas propriedades é a de que produtos de tensores esféricos podem ser reduzidos com o uso dos mesmos procedimentos envolvidos no acoplamento de momentos angulares. De fato, o produto dos tensores esféricos $T_{k_1 q_1}$ de ordem k_1 e $S_{k_2 q_2}$ de ordem k_2 se transforma de acordo com a relação

$$\begin{aligned} T'_{k_1 q_1} S'_{k_2 q_2} &= \sum_{q'_1 q'_2} T_{k_1 q'_1} S_{k_2 q'_2} D_{q'_1 q_1}^{k_1}(\alpha, \beta, \gamma) D_{q'_2 q_2}^{k_2}(\alpha, \beta, \gamma) = \\ &= \sum_K \sum_{Q Q'} \sum_{q'_1 q'_2} T_{k_1 q'_1} S_{k_2 q'_2} C_{q'_1 q'_2}^{k_1 k_2 K} C_{q_1 q_2}^{K Q} D_{Q' Q}^K(\alpha, \beta, \gamma) \end{aligned}$$

onde o produto de duas matrizes de rotação foi re-escrito usando a relação (9.18). Usando a ortogonalidade da transformação representada pelos coeficientes de Clebsch-Gordan resulta

$$\sum_{q_1 q_2} T'_{k_1 q_1} S'_{k_2 q_2} C_{q_1 q_2}^{k_1 k_2 K} = \sum_{Q Q'} \sum_{q'_1 q'_2} T_{k_1 q'_1} S_{k_2 q'_2} C_{q'_1 q'_2}^{k_1 k_2 K} D_{Q' Q}^K(\alpha, \beta, \gamma)$$

Como o lado esquerdo dessa relação é a versão transformada de

$$[T_{k_1} S_{k_2}]_{KQ} \equiv \sum_{q_1 q_2} T_{k_1 q_1} S_{k_2 q_2} C_{q_1 q_2}^{k_1 k_2 K} \quad (9.19)$$

resulta que as $2K + 1$ componentes desse objeto constituem por sua vez um tensor esférico de ordem K . Isso mostra que o “acoplamento” de tensores esféricos com coeficientes de Clebsch-Gordan produz novos tensores esféricos cuja ordem corresponde à quantidade que nesses coeficientes está associada ao momento angular total. Em particular, se $K = 0$ (o que exige $k_1 = k_2$ para que o resultado não seja nulo) o produto acoplado, nesse caso chamado *produto escalar* dos dois tensores esféricos, é invariante sob rotações. Um caso particular da definição (9.19) é aquele em que o segundo tensor é o conjunto de sub-estados magnéticos de vetores de estado com momento angular J

$$[T_{k_1} |\lambda J\rangle]_{KQ} = \sum_{q_1 M} T_{k_1 q_1} |\lambda JM\rangle C_{q_1 M}^{k_1 J K}. \quad (9.20)$$

O resultado, nesse caso, é um novo vetor de estado com momento angular K para o qual o autovalor de J_3 é $\hbar Q$.

Uma segunda propriedade importante diz respeito a elementos de matriz de operadores tensoriais esféricos entre autovetores do quadrado do momento angular J^2 e de sua terceira componente J_3 . Usando mais uma vez a ortogonalidade da transformação representada pelos coeficientes de Clebsch-Gordan, agora aplicada à relação (9.20), é possível escrever

$$\langle \lambda' J' M' | T_{kq} | \lambda J M \rangle = \sum_K \langle \lambda' J' M' | [T_k | \lambda J] \rangle_{KQ} C_q^{k J K} = \langle \lambda' J' M' | [T_k | \lambda J] \rangle_{J' M'} C_q^{k J J'}$$

onde foi feito uso da ortogonalidade de autovetores de J^2 e de J_3 associados a autovalores diferentes. O fato importante aqui é que o produto escalar que aparece no último termo é, na realidade, independente de M' . De fato, usando as relações

$$[[T_k J]]_{J' M' \pm 1} = \frac{J_{\pm} [[T_k J]]_{J' M'}}{\sqrt{J'(J' + 1) - M'(M' \pm 1)}} \quad \text{e} \quad |\lambda' J' M' \pm 1\rangle = \frac{J_{\pm} |\lambda' J' M'\rangle}{\sqrt{J'(J' + 1) - M'(M' \pm 1)}},$$

bem como o fato de que $J_{\mp} J_{\pm} = J^2 - J_3(J_3 \pm 1)$, resulta que

$$\langle \lambda' J' M' \pm 1 | [T_k | \lambda J] \rangle_{J' M' \pm 1} = \langle \lambda' J' M' | [T_k | \lambda J] \rangle_{J' M'}.$$

Isso mostra que *toda* a dependência do elemento de matriz considerado com relação aos números quânticos M', q e M está contida no coeficiente de Clebsch-Gordan. Esse resultado é conhecido como *teorema de Wigner-Eckart*, e é usualmente expresso escrevendo o elemento de matriz sob a forma

$$\langle \lambda' J' M' | T_{kq} | \lambda J M \rangle = \frac{1}{\sqrt{2J' + 1}} C_M^{J k J'} \langle \lambda' J' || T_k || \lambda J \rangle \quad (9.21)$$

onde o último fator, chamado *elemento de matriz reduzido*, é independente de M, q e M' . O coeficiente de Clebsch-Gordan que aparece nesta última expressão difere do anterior apenas por uma fase independente desses números quânticos. De fato

$$C_M^{J k J'} = (-1)^{J+k-J'} C_q^{k J J'}.$$

Vale a pena notar explicitamente que a relação (9.21) serve como *definição* do elemento de matriz reduzido. O cálculo desse elemento de matriz é feito na realidade invertendo a relação (9.21), isto é, calculando

$$\langle \lambda' J' || T_k || \lambda J \rangle = \frac{\sqrt{2J' + 1}}{C_M^{J k J'}} \langle \lambda' J' M' | T_{kq} | \lambda J M \rangle$$

e o conteúdo essencial do teorema de Wigner-Eckart é o fato de que o resultado *independe* dos particulares valores de M, q e M' que tenham sido utilizados no cálculo.

Um exemplo importante e típico do uso do teorema de Wigner-Eckart é a obtenção de regras de seleção de momento angular para transições as transições multipolares cujas amplitudes foram obtidas nas equações (8.37) e (8.38). De fato, os operadores que intervêm nessas amplitudes

$$\mathcal{T}_{J\mu}^E \equiv \frac{1}{k} \left[\nabla \times j_J(kr) \vec{\mathcal{Y}}_{J1J\mu} \right] \cdot \frac{\vec{p}}{m_e} \quad \text{e} \quad \mathcal{T}_{J\mu}^M \equiv j_J(kr) \vec{\mathcal{Y}}_{J1J\mu} \cdot \frac{\vec{p}}{m_e}$$

são componentes μ de tensores esféricos de ordem J . Para estados iniciais e finais com momento angular J_I, M_I e J_F, M_F respectivamente, as amplitudes de transição podem ser escritas em termos de elementos de matriz reduzidos como

$$T_{IF}^{E,M J,\mu} = \langle \lambda_F J_F M_F | \mathcal{T}_{J\mu}^{E,M} | \lambda_I J_I M_I \rangle = \frac{1}{\sqrt{2J_F+1}} C_{M_I \mu M_F}^{J_I J J_F} \langle \lambda_F J_F || \mathcal{T}_J^{E,M} || \lambda_I J_I \rangle$$

o que explicita as regras de seleção de momento angular através das condições que devem ser satisfeitas para que o coeficiente de Clebsch-Gordan não se anule, isto é

$$|J_I - J| \leq J_F \leq J_I + J \quad \text{e} \quad M_F = M_I + \mu.$$

Na expansão multipolar que leva às expressões (8.37) e (8.38) deve-se ter $\mu = \pm 1$. Essa restrição se deve à escolha da direção de propagação \hat{k} do fóton emergente como eixo de quantização que define a componente J_3 do momento angular. Através das regras de seleção ela implica que o número quântico magnético do sistema que sofre a transição radiativa muda necessariamente de uma unidade quando definida com relação a esse eixo de quantização. Esse mesmo resultado foi obtido de forma mais direta no limite de grandes comprimentos de onda (v. página 460).

A integral sobre o produto de três harmônicas esféricas (8.31), utilizada para determinar as regras de seleção de dipolo elétrico no limite de grandes comprimentos de onda, fornece um exemplo de cálculo explícito de um elemento de matriz reduzido. Essa integral pode ser vista como o elemento de matriz do tensor esférico $Y_{1\mu}(\hat{r})$ entre as autofunções do momento angular orbital $Y_{lm}(\hat{r})$ e $Y_{l'm'}(\hat{r})$. Desse modo

$$\begin{aligned} \int d\hat{r} Y_{l'm'}^*(\hat{r}) Y_{1\mu}(\hat{r}) Y_{lm}(\hat{r}) &\equiv \langle Y_{l'm'} | Y_{1\mu} | Y_{lm} \rangle = \sqrt{\frac{3}{4\pi}} \sqrt{\frac{2l+1}{2l'+1}} C_{000}^{l1l'} C_{m\mu m'}^{l1l'} = \\ &= \frac{1}{\sqrt{2l'+1}} C_{m\mu m'}^{l1l'} \langle Y_{l'} || Y_1 || Y_l \rangle \end{aligned}$$

o que dá para o elemento de matriz reduzido

$$\langle Y_{l'} || Y_l || Y_l \rangle = \sqrt{\frac{3(2l+1)}{4\pi}} C_{000}^{l1l'}$$

que, como afirma o teorema de Wigner-Eckart, é independente de m, μ e m' .

9.4 Isospin.

As propriedades intrínsecas (como massa e spin) e as propriedades relativas às interações fortes de famílias de partículas como, por exemplo, os nucleons (prótons e neutrons) e os pions (π^+ , π^0 e π^-) revelam semelhanças que sugerem que tais famílias possam ser vantajosamente descritas em termos de “multipletos de carga”, de multiplicidade $2t+1$, análogos aos multipletos constituídos pelos diferentes estados magnéticos de um sistema com spin, com $t = 1/2$ no caso dos nucleons e $t = 1$ no caso dos pions. A adequação dessa descrição se estende ainda a sistemas compostos de várias dessas partículas, cujos estados podem ser classificados em termos de multipletos de carga análogos aos que são constituídos pelos diferentes estados magnéticos do spin total de um sistema composto de vários spins. No contexto da espectroscopia nuclear esse tipo de tratamento foi discutido já em 1936 por Cassen e Condon³, onde se originou a denominação *spin isotópico*, posteriormente substituída por *isospin*.

Nesse esquema, os $2t+1$ constituintes de uma mesma família de partículas são tratados como diferentes estados de carga uma única espécie de partícula dotada de propriedades intrínsecas descritas em termos de variáveis dinâmicas adimensionais t_1, t_2 e t_3 satisfazendo relações de comutação análogas às do momento angular, isto é

$$[t_j, t_k] = i \sum_l \epsilon_{jkl} t_l.$$

A ausência do fator \hbar , presente nas relações de comutação do momento angular, se deve ao caráter adimensional dos t_j . Essas relações de comutação são características das componentes cartesianas de um gerador de rotações \vec{t} (o *isospin* da família de partículas) definido em um espaço abstrato de três dimensões, chamado *espaço de carga*, independente do espaço de posições. Em particular, as componentes desse vetor comutam com o gerador \vec{J} das rotações espaciais.

O caso dos nucleons corresponde a um dubleto de isospin $2t+1 = 2$ ou seja, $t = 1/2$, e uma representação possível dos operadores t_j pode ser escrita em termos das matrizes de Pauli, como

$$t_j = \frac{1}{2} \tau_j, \quad j = 1, 2, 3$$

³B. Cassen e E. U. Condon, *Phys. Rev.* **50**, 846 (1936).

com

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{e} \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Nessa representação o operador t_3 (que corresponde à terceira componente do vetor \vec{t} no espaço de carga) é diagonal, e é usual associar os estados de próton e de neutron do nucleon respectivamente aos seus autovetores com autovalor $+1/2$ e $-1/2$, isto é

$$\text{próton} \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \left| \frac{1}{2} \frac{1}{2} \right\rangle \quad \text{e} \quad \text{neutron} \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \left| \frac{1}{2} -\frac{1}{2} \right\rangle.$$

Com isso a carga do nucleon é a variável dinâmica associada ao operador

$$Q_N \equiv \frac{e}{2} (1 + \tau_3) = e \left(\frac{1}{2} + t_3 \right)$$

enquanto a ação dos operadores $t_{\pm} \equiv t_1 \pm it_2$ sobre os autovetores de t_3 é

$$t_- \left| \frac{1}{2} \pm \frac{1}{2} \right\rangle = \begin{cases} \left| \frac{1}{2} -\frac{1}{2} \right\rangle \\ 0 \end{cases} \quad \text{e} \quad t_+ \left| \frac{1}{2} \pm \frac{1}{2} \right\rangle = \begin{cases} 0 \\ \left| \frac{1}{2} \frac{1}{2} \right\rangle \end{cases}$$

isto é, t_- transforma um próton em um neutron e aniquila um estado de neutron, enquanto t_+ transforma um neutron em um próton e aniquila um estado de próton. Os operadores t_{\pm} são por isso chamados *operadores de troca de carga*. O isospin total \vec{T} de um sistema de vários nucleons é definido como

$$\vec{T} = \sum_k \vec{t}^{(k)}$$

onde $\vec{t}^{(k)}$ é o isospin do k -ésimo nucleon. É claro que todas as regras de adição de momentos angulares se aplicam também à adição de isospins. Em particular, o quadrado do isospin total T^2 comuta com qualquer de suas componentes T_j bem como com os quadrados dos isospins constituintes $t^{(k)2}$

$$[T^2, T_j] = [T^2, t^{(k)2}] = 0$$

mas não comuta com as componentes dos isospins constituintes, $[T^2, t_3^{(k)}] \neq 0$. No caso de apenas dois nucleons existem portanto duas representações possíveis em que os operadores diagonais são respectivamente $(t^{(1)2}, t_3^{(1)}, t^{(2)2}, t_3^{(2)})$ e $(t^{(1)2}, t^{(2)2}, T^2, T_3)$. Essas duas representações são relacionadas uma à outra através de coeficientes de Clebsch-Gordan

$$|\frac{1}{2}\frac{1}{2}TT_3\rangle = \sum_{t_3^{(1)}, t_3^{(2)}} C_{t_3^{(1)} t_3^{(2)} T_3}^{\frac{1}{2} \frac{1}{2} T} |\frac{1}{2}t_3^{(1)}\frac{1}{2}t_3^{(2)}\rangle \quad (9.22)$$

e a sua inversa

$$|\frac{1}{2}t_3^{(1)}\frac{1}{2}t_3^{(2)}\rangle = \sum_T C_{t_3^{(1)} t_3^{(2)} T}^{\frac{1}{2} \frac{1}{2} T} |\frac{1}{2}\frac{1}{2}TT_3\rangle.$$

A carga elétrica total desses estados pode ser expressa em termos da terceira componente T_3 do isospin total como

$$Q_{NN} = e(1 + T_3).$$

No caso de um tripleto de isospin, $2t + 1 = 3$ ou $t = 1$, como é o caso dos pions, uma representação explícita pode ser escrita tomando por empréstimo as matrizes que representam os estados de spin 1, obtidas no final da seção 9.3.1. Usando a representação em que a matriz t_3 é diagonal

$$t_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_2 = \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

os três estados de carga π_+ , π_0 e π_- podem ser associados aos autovetores de t_3

$$\pi_+ \leftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \equiv |11\rangle, \quad \pi_0 \leftrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \equiv |10\rangle, \quad \text{e} \quad \pi_- \leftrightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \equiv |1-1\rangle$$

com o operador de carga dado por

$$Q_\pi \equiv et_3$$

e operadores de troca de carga $t_\pm = t_1 \pm it_2$. Agora é possível considerar também, por exemplo, os estados de isospin de um sistema pion-nucleon. O operador que corresponde ao isospin total desse sistema é

$$\vec{T} = \vec{t}^{(N)} + \vec{t}^{(\pi)}$$

onde $\vec{t}^{(N)}$ e $\vec{t}^{(\pi)}$ são respectivamente os isospins do nucleon e do pion. O operador \vec{T} pode ser tratado com as regras de composição usuais de momentos angulares. Os valores possíveis do

isospin total são portanto $T = 1/2$ e $T = 3/2$, que correspondem respectivamente a um dubleto de carga e a um quadrupletto de carga. Os autovetores simultâneos de $t^{(N)2}$, $t^{(\pi)2}$, T^2 e T_3 são dados por

$$|\frac{1}{2} 1 T T_3\rangle = \sum_{t_3^{(N)}, t_3^{(\pi)}} C_{t_3^{(N)} t_3^{(\pi)} T_3}^{\frac{1}{2} 1 T} |\frac{1}{2} t_3^{(N)} 1 t_3^{(\pi)}\rangle \quad (9.23)$$

e têm uma carga elétrica total dada por

$$Q_{\pi N} = e \left(\frac{1}{2} + t_3^{(N)} + t_3^{(\pi)} \right) = e \left(\frac{1}{2} + T_3 \right).$$

O dubleto $T = 1/2$ consiste portanto de dois estados cujas cargas totais são zero e e , enquanto o quadrupletto $T = 3/2$ contém quatro estados cujas cargas totais são $-e, 0, e$ e $2e$.

É importante notar que nos estados envolvendo várias partículas que têm isospin total bem definido a carga de cada uma das partículas constituintes não é, em geral, bem definida. Esse fato é uma consequência da não comutatividade do isospin total \vec{T} com as terceiras componentes dos isospins constituintes, que são as variáveis dinâmicas que determinam a carga de cada uma das partículas. Os estados com isospin total bem definido serão, em geral, expressos como combinações lineares de estados em que as partículas constituintes aparecem em estados de carga diferentes. Como a carga *total* é determinada pela variável dinâmica com valor bem definido T_3 , no entanto, a soma das cargas das duas partículas em cada uma das componentes será a mesma.

Isso pode ser visto explicitamente tanto no caso do sistema de dois nucleons como no caso do sistema pion-nucleon. Para o primeiro desses sistemas o estado com isospin total igual a zero é dado por

$$|\frac{1}{2} \frac{1}{2} T = 0 T_3 = 0\rangle = \frac{1}{\sqrt{2}} (|\text{próton}_1 \text{ neutron}_2\rangle - |\text{neutron}_1 \text{ próton}_2\rangle)$$

obtido simplesmente substituindo os valores numéricos dos coeficientes de Clebsch-Gordan apropriados em (9.22). como indicado pela notação, o nucleon que está no estado de carga de próton (neutron) na primeira das duas componentes está no estado de neutron (próton) na segunda. O mesmo ocorre com o estado com $T = 1$ e $T_3 = 0$, o qual é dado por

$$|\frac{1}{2} \frac{1}{2} T = 1 T_3 = 0\rangle = \frac{1}{\sqrt{2}} (|\text{próton}_1 \text{ neutron}_2\rangle + |\text{neutron}_1 \text{ próton}_2\rangle).$$

Este último estado é um dos estados de um tripleto de isospin que inclui também os estados com $T = 1$ e $T_3 = \pm 1$

$$|\frac{1}{2} \frac{1}{2} T = 1 T_3 = 1\rangle = |\text{próton}_1 \text{ próton}_2\rangle \quad \text{e} \quad |\frac{1}{2} \frac{1}{2} T = 1 T_3 = -1\rangle = |\text{neutron}_1 \text{ neutron}_2\rangle.$$

O fato de que a carga dos nucleons constituintes nesses dois estados está também bem definida resulta de que eles são os únicos estados de dois nucleons com carga elétrica total $2e$ e 0 respectivamente. Desse modo a soma em (9.22) tem nesses casos um único termo, no qual o coeficiente de Clebsch-Gordan é igual a 1.

Para o sistema pion-nucleon, o estado com $T = 1/2$ e carga total nula (isto é, $T_3 = -1/2$) pode ser escrito explicitamente, a partir de (9.23), como

$$|\frac{1}{2} 1 T = \frac{1}{2} T_3 = -\frac{1}{2}\rangle = \sqrt{\frac{3}{2}} |\pi_- \text{ próton}\rangle - \frac{1}{\sqrt{3}} |\pi_0 \text{ neutron}\rangle.$$

A combinação linear ortogonal a essa, das mesmas duas componentes, corresponde ao estado com $T = 3/2$ e $T_3 = -1/2$. Por outro lado, o estado de carga máxima $2e$ tem $T_3 = 3/2$, o que implica que $T = 3/2$. Esse estado corresponde ao estado de único em que a soma das terceiras componentes $t_3^{(N)}$ e $T_3^{(\pi)}$ é máxima. Portanto

$$|\frac{1}{2} 1 T = \frac{3}{2} T_3 = \frac{3}{2}\rangle = |\pi_+ \text{ próton}\rangle.$$

De forma análoga ao que ocorre no caso do grupo das rotações espaciais, o operador \vec{T} correspondente ao isospin total é o gerador de rotações infinitesimais *no espaço de cargas*. A invariança da dinâmica sob esse tipo de rotação, usualmente chamada simplesmente *invariança de carga*, é expressa através da comutatividade do hamiltoniano com esse gerador, isto é

$$[H, \vec{T}] = 0.$$

Tendo em conta as relações de comutação das componentes do isospin total, isso significa que os operadores T^2 e T_3 são compatíveis com H e também comutam entre si, o que leva à existência de um base de autovetores simultâneos desse três operadores, $\{|ETq\rangle\}$, $-T \leq q \leq T$. Consequência característica da invariança de carga é o fato de que os autovalores de H não dependem do autovalor q de T_3 , de modo que o espectro de energia consiste de *multipletos de carga* com $2T + 1$ estados degenerados para isospin total T .

Interações electromagnéticas violam a invariança de carga devido à sua dependência explícita com a carga elétrica das partículas que constituem o sistema. Por exemplo, a interação coulombiana em um sistema de muitos nucleons pode ser escrita em termos dos operadores de isospin como

$$\begin{aligned}
V_C &= \frac{1}{2} \sum_{jk} \frac{e^2}{|\vec{r}_j - \vec{r}_k|} \left(\frac{1}{2} + t_3^{(j)} \right) \left(\frac{1}{2} + t_3^{(k)} \right) = \\
&= \frac{1}{8} \sum_{jk} \frac{e^2}{|\vec{r}_j - \vec{r}_k|} + \frac{1}{4} \sum_{jk} \frac{e^2}{|\vec{r}_j - \vec{r}_k|} \left(t_3^{(j)} + t_3^{(k)} \right) + \frac{1}{2} \sum_{jk} \frac{e^2}{|\vec{r}_j - \vec{r}_k|} t_3^{(j)} t_3^{(k)}.
\end{aligned}$$

O primeiro termo desta última expressão é independente do isospin e é portanto invariante por uma rotação no espaço de carga, mas o mesmo não acontece com os outros dois termos. Como os operadores vetoriais (no espaço de carga) $\vec{t}^{(j)}$ correspondem a tensores esféricos de ordem 1 (nesse espaço), o segundo têm o caráter de componente de um tensor esférico de ordem 1, enquanto o terceiro pode ser escrito como a soma de um escalar com uma componente de tensor esférico de ordem 2 no espaço de cargas. De fato

$$t_3^{(j)} t_3^{(k)} \equiv t_{q_j=0}^{(j)} t_{q_k=0}^{(k)} = \sum_K C_{0 \ 0 \ 0}^{1 \ 1 \ K} \left[t^{(j)} t^{(k)} \right]_{K0} = -\frac{1}{\sqrt{3}} \left[t^{(j)} t^{(k)} \right]_{00} + \sqrt{\frac{2}{3}} \left[t^{(j)} t^{(k)} \right]_{20}$$

onde foi usado o valor explícito dos coeficientes de Clebsch-Gordan (que é nulo no caso $K = 1$) e os colchetes são as componentes de tensores esféricos acoplados de ordem K

$$\left[t^{(j)} t^{(k)} \right]_{K0} \equiv \sum_{q_j q_k} C_{q_j \ q_k \ 0}^{1 \ 1 \ K} t_{q_j}^{(j)} t_{q_k}^{(k)}.$$

Usando o teorema de Wigner-Eckart resulta portanto que a interação coulombiana V_C em geral tem elementos de matriz não nulos entre estados com isospins que diferem de modo que $\Delta T = 1$ (para os termos tensoriais de ordem 1) ou $\Delta T = 2$ (para os termos tensoriais de ordem 2). É claro, no entanto, que

$$[V_C, t_3^{(j)}] = [V_C, T_3] = 0$$

pois apenas as terceiras componentes dos operadores de isospin aparecem em V_C . O anulamento do primeiro desses comutadores exprime a conservação da carga de cada um dos nucleons envolvidos na interação coulombiana, enquanto o anulamento do segundo exprime a conservação da carga total do sistema de nucleons.

Esses fatos mostram que o isospin total *não* pode ser um bom número quântico (isto é, o operador T^2 não é uma constante do movimento) para sistemas de muitos nucleons, como são os núcleos atômicos, devido à presença de interações coulombianas entre os prótons. Na medida em que a dinâmica desses sistemas seja dominada por interações fortes, invariantes por rotações no espaço de carga, o isospin total é no entanto uma quantidade *aproximadamente* conservada. Na

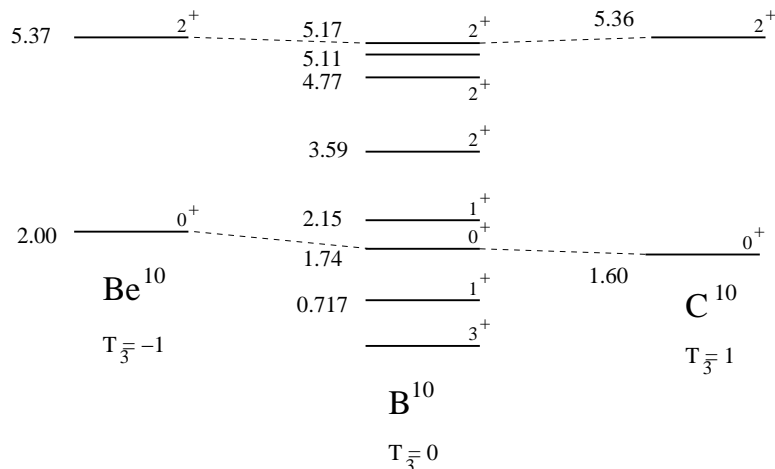


Figura 9.3: Diagrama isobárico para três núcleos com $A = Z + N = 10$. As diferenças de energia coulombiana foram devidamente subtraídas. À esquerda de cada um dos níveis está indicada a energia de excitação em MeV (com relação ao estado fundamental do B^{10}), e à direita o momento angular e a paridade do nível (quando conhecidos). As linhas tracejadas ligam os níveis pertencentes a tripletos de isospin. Os níveis restantes, no B^{10} , correspondem a estados com $T = 0$.

realidade, subtraindo diferenças de energia associadas às interações coulombianas, os multipletos de isospin se tornam claramente visíveis nos espectros de núcleos com números de massa iguais mas com números atômicos diferentes. Um exemplo disso pode ser visto na figura 9.3, que mostra os níveis de menor energia dos núcleos Be^{10} ($N = 6$, $Z = 4$ e portanto $T_3 = -1$), B^{10} ($N = Z = 5$) e C^{10} ($N = 4$, $Z = 6$, $T_3 = 1$) após subtração das diferenças de energia coulombiana. São claramente visíveis vários níveis do B^{10} que não têm contrapartida nos outros dois núcleos, e que correspondem portanto a singletos de isospin, bem como dois tripletos que correspondem a $T = 1$. O caráter aproximado da classificação em termos do isospin total é revelado pela degenerescência apenas aproximada dos níveis correspondentes aos tripletos.

Outro exemplo que mostra a utilidade do isospin em processos dominados por interações fortes provém da análise de algumas propriedades do espalhamento pion-nucleon. A interação pion-nucleon inclui certamente efeitos eletromagnéticos (em particular, efeitos coulombianos no caso de espalhamento π_{\pm} -próton), que violam a conservação do isospin total. No entanto, na medida em que os efeitos devidos às interações fortes sejam dominantes, a suposição da conservação do isospin total pode servir como uma primeira aproximação útil para simplificar a descrição da dinâmica do processo de espalhamento. Como consequência dessa suposição, a

amplitude de espalhamento deve ser diagonal no isospin total T e na sua terceira componente T_3 . Desse modo, o espalhamento pion-nucleon deve poder ser bem descrito em termos de amplitudes de espalhamento f_{TT_3} , com $T = 3/2$ e $T = 1/2$ e com os valores de T_3 compatíveis com cada um desses valores do isospin total. Além disso, se a interação responsável pelo espalhamento é invariante por uma rotação no espaço de carga, o operador de transição \mathcal{T} deve ser um *escalar* nesse espaço, ou seja, um tensor esférico de ordem zero. Nessas condições o teorema de Wigner-Eckart permite escrever as amplitudes de espalhamento como

$$f_{TT_3} = \langle TT_3 | \mathcal{T} | TT_3 \rangle = C_{T_3 0 T_3}^{T 0 T} \frac{\langle T || \mathcal{T} || T \rangle}{\sqrt{2T+1}}.$$

O coeficiente de Clebsch-Gordan que aparece nessa expressão é independente de T_3 e igual a 1, de modo que as amplitudes de espalhamento são na realidade independentes da carga total do sistema, determinada por T_3 . O espalhamento é portanto determinado completamente em termos de duas amplitudes apenas, $f_{\frac{3}{2}}$ e $f_{\frac{1}{2}}$.

Essas amplitudes de espalhamento correspondem a estados com isospin total bem definido. Como visto anteriormente, nesses estados as cargas das partículas não é em geral bem definida. Invertendo a relação (9.23) é possível porém exprimis estados envolvendo partículas com cargas bem definidas em termos de estados com isospin total bem definido. Em particular

$$|\pi_- \text{ próton} \rangle = \frac{1}{\sqrt{3}} \left| \frac{3}{2} \quad -\frac{1}{2} \right\rangle + \sqrt{\frac{2}{3}} \left| \frac{1}{2} \quad -\frac{1}{2} \right\rangle$$

donde segue que o processo acessível à experiência

$$\pi_- + \text{próton} \longrightarrow \pi_- + \text{próton}$$

é descrito pela amplitude

$$\begin{aligned} f_- &= \langle \pi_- \text{ próton} | \mathcal{T} | \pi_- \text{ próton} \rangle = \\ &= \frac{1}{3} \left\langle \frac{3}{2} \quad -\frac{1}{2} \right| \mathcal{T} \left| \frac{3}{2} \quad -\frac{1}{2} \right\rangle + \frac{2}{3} \left\langle \frac{1}{2} \quad -\frac{1}{2} \right| \mathcal{T} \left| \frac{1}{2} \quad -\frac{1}{2} \right\rangle = \frac{1}{3} f_{\frac{3}{2}} + \frac{2}{3} f_{\frac{1}{2}}. \end{aligned}$$

pois a conservação do isospin total implica no anulamento dos termos cruzados envolvendo estados iniciais e finais com isospins diferentes. Analogamente, o processo

$$\pi_+ + \text{próton} \longrightarrow \pi_+ + \text{próton}$$

é descrito pela amplitude

$$f_+ = f_{\frac{3}{2}}$$

e o processo com troca de carga

$$\pi_- + \text{próton} \longrightarrow \pi_0 + \text{neutron}$$

pela amplitude

$$f_t = \frac{\sqrt{2}}{3} (f_{\frac{3}{2}} - f_{\frac{1}{2}}) .$$

para obter este último resultado é necessário usar a expressão do estado final em termos de estados com isospin total bem definido

$$|\pi_0 \text{ neutron}\rangle = \sqrt{\frac{2}{3}} \left| \frac{3}{2} - \frac{1}{2} \right\rangle - \frac{1}{\sqrt{3}} \left| \frac{1}{2} - \frac{1}{2} \right\rangle .$$

As seções de choque correspondentes a cada uma das três amplitudes de espalhamento acessíveis à experiência são então

$$\begin{aligned} \sigma_+ &= |f_+|^2 = \left| f_{\frac{3}{2}} \right|^2 \\ \sigma_- &= |f_-|^2 = \frac{1}{9} \left[\left| f_{\frac{3}{2}} \right|^2 + 4 \left| f_{\frac{1}{2}} \right|^2 + 4 \operatorname{Re} (f_{\frac{3}{2}}^* f_{\frac{1}{2}}) \right] \\ \sigma_t &= |f_t|^2 = \frac{1}{9} \left[2 \left| f_{\frac{3}{2}} \right|^2 + 2 \left| f_{\frac{1}{2}} \right|^2 - 4 \operatorname{Re} (f_{\frac{3}{2}}^* f_{\frac{1}{2}}) \right] . \end{aligned}$$

Dessas relações resulta que

$$\sigma_t + \sigma_- = \frac{1}{9} \left(3\sigma_+ + 6 \left| f_{\frac{1}{2}} \right|^2 \right)$$

que pode ser resolvida para o termo envolvendo a amplitude $f_{\frac{1}{2}}$ dando

$$\left| f_{\frac{1}{2}} \right|^2 = \frac{3(\sigma_- + \sigma_t) - \sigma_+}{2}$$

que permite obter essa quantidade diretamente do valor de seções de choque elásticas e com troca de carga medidas, para pions incidindo sobre nucleons. Experimentalmente, para energias

dos pions entre 100 MeV e 300 MeV no sistema de laboratório, tem-se que $\sigma_+ \approx 3(\sigma_- + \sigma_t)$. Isso significa que a seção de choque total de espalhamento de π_+ por prótons é, nessas energias, essencialmente igual a três vezes a seção de choque total de espalhamento (incluindo nesse caso o canal de troca de carga) de π_- por prótons, indicando que $\left|f_{\frac{1}{2}}\right|^2 \approx 0$.

A consequência desse resultado é que as três seções de choque, σ_+ , σ_- e σ_t são essencialmente determinadas por uma única amplitude $f_{\frac{3}{2}}$, e devem portanto manter entre si as relações

$$\sigma_+ : \sigma_- : \sigma_t = 9 : 1 : 2$$

se, como suposto, esses processos de espalhamento são dominados nessas energias por interações invariantes por rotações no espaço da carga. Os valores experimentais para essas relação são, à energia de 120 MeV

$$\sigma_+ : \sigma_- : \sigma_t = 93 : 11 : 22$$

e, à energia de 144 MeV

$$\sigma_+ : \sigma_- : \sigma_t = 158 : 17 : 31.$$

As seções de choque σ_+ e $\sigma_- + \sigma_t$ mostram uma ressonância muito pronunciada em torno da energia de 200 MeV no laboratório. Essa ressonância (chamada “ressonância delta”) foi descoberta no sistema π_+ +próton por Anderson, Fermi e colaboradores em 1953⁴. Em termos de isospin, o canal π_+ +próton corresponde a $T_3 = 3/2$, e portanto a um caso puro de isospin $T = 3/2$. A análise precedente mostra no entanto que a mesma ressonância, como observada no canal π_- +próton, é também consistente com o valor do isospin total $T = 3/2$. A invariança (aproximada) por rotações no espaço de carga implica que a ressonância delta seja de fato um quadrupletto de isospin com $-3/2 \leq T_3 \leq 3/2$, o que inclui os canais π_0 +próton, π_0 +neutron e π_- + *neutron*, dos quais os dois últimos são menos acessíveis em termos experimentais.

9.5 Paridade e paridade intrínseca.

A representação do grupo constituído pela operação de inversão espacial e pelo seu quadrado (que coincide com a identidade) no espaço de Hilbert de funções de onda já foi tratada na seção 9.1.2 como um exemplo simples de implementação do procedimento geral a ser utilizado. Tendo em vista o tratamento subsequente dos grupos de translações e de rotações espaciais, cabe, em primeiro lugar, rever os resultados obtidos anteriormente para a operação de inversão de modo a sublinhar os elementos comuns presentes nos dois contextos.

⁴H. L. Anderson, E. Fermi, R. Martin e D. E. Nagle, *Phys. Rev.* **91**, 155 (1953).

A representação dada pelos operadores do espaço de Hilbert R_1 e R_P , definidos por

$$R_1\psi(q) = \psi(q) \quad \text{e} \quad R_P\psi(q) = \psi(-q)$$

permite identificar dois conjuntos distintos de funções, que são o conjunto das funções *pares* $\psi_+(q)$ e o das funções *ímpares* $\psi_-(q)$, caracterizados pelos efeitos de R_P através das relações

$$R_P\psi_{\pm}(q) = \pm\psi_{\pm}(q).$$

Esses dois conjuntos podem ser associados, por sua vez, a *duas* outras representações do mesmo grupo, definidas respectivamente por

$$R_1^{(+)} = \hat{1}, \quad R_P^{(+)} = \hat{1} \quad \text{e} \quad R_1^{(-)} = \hat{1}, \quad R_P^{(-)} = -\hat{1}.$$

Na primeira, tanto a operação de inversão espacial quanto a identidade são representadas pelo operador identidade do espaço de Hilbert, enquanto na segunda a operação de inversão é representada pelo operador $-\hat{1}$, enquanto a identidade é representada também pelo operador identidade do espaço de Hilbert. Embora a primeira dessas representações seja a representação trivial, é claro que todas as operações do grupo são mapeadas corretamente nos dois casos. Essas duas representações são unidimensionais, e portanto irredutíveis. As funções de onda pares se transformam segundo a representação irredutível $(R_P^{(+)}, R_1^{(+)})$ enquanto as funções ímpares se transformam segundo a representação $(R_P^{(-)}, R_1^{(-)})$. Da mesma forma que uma função de onda com momento angular J pode ser caracterizada pela propriedade que ela se transforma sob rotações segundo a representação irredutível $D_{M'M}^J(\alpha, \beta, \gamma)$ do grupo de rotações, funções de onda com paridade positiva e negativa podem ser caracterizadas como aquelas que se transformam sob inversão espacial segundo as representações $(+)$ e $(-)$, respectivamente.

Para sistemas invariantes por inversão espacial e que sejam constituídos de um conjunto fixo de partículas, a classificação das autofunções do hamiltoniano em termos de sua transformação por uma dessas duas representações irredutíveis é suficiente para dar conta de todos os fatos experimentais. O mesmo acontece no caso de um sistema envolvendo um conjunto fixo de partículas carregadas interagindo com o campo de radiação, em que pode ocorrer a emissão ou absorção de um ou mais fótons. Existem, no entanto, situações em que o conjunto de partículas não é fixo e que exigem a ampliação desse esquema simples.

Um processo que ilustra situações desse tipo é a desintegração do deuteron d (estado ligado de um próton e um neutron) pela absorção de um pion

$$\pi_- + d \longrightarrow \text{neutron} + \text{neutron}.$$

De fato, esse processo de captura do π_- é observada a partir de estados em que o seu momento angular orbital é zero (estado fundamental do deutério piônico, que é um átomo de deutério

no qual o elétron é substituído por um pion negativo). Além disso, o deuteron é sabidamente um objeto com momento angular total $I = 1$ e paridade positiva, de modo que o estado inicial para o processo de captura tem segundo a classificação simples acima, momento angular total $J = 1$ e paridade positiva.

A conservação do momento angular total implica portanto que o estado final de dois neutrons deve ser também um estado com momento angular total $J = 1$. As situações possíveis para que essa condição esteja satisfeitas são listadas no quadro abaixo, que inclui também a paridade de cada um dos estados finais considerados.

Spin total	l	Simetria	Paridade
0	1	$(-1) \times (-1)^l = +1$	$(-1)^l = -1$
1	0, 2	$(+1) \times (-1)^l = +1$	$(-1)^l = +1$
1	1	$(+1) \times (-1)^l = -1$	$(-1)^l = -1$

A primeira coluna do quadro mostra o valor do spin total dos dois neutrons. Os estados de spin são antissimétricos pela troca dos dois neutrons para $S = 0$ e simétricos para $S = 1$. A segunda coluna mostra valores possíveis do momento angular orbital relativo dos dois neutrons, compatíveis com o valor suposto para o spin total e com o valor exigido $J = 1$ para o momento angular total. Estados com valores pares e ímpares do momento angular orbital são respectivamente simétricos e antissimétricos pela troca dos dois neutrons. A terceira coluna mostra a simetria ou antissimetria global do estado final, tendo em conta a simetria ou antissimetria tanto da função de onda de spin quanto orbital, e a última coluna identifica a paridade do estado considerado, que é dada em termos do valor do momento angular orbital. Como o estado final envolve dois férmions idênticos, ele deve ser *antissimétrico* pela troca dos dois neutrons, o que exclui as duas primeiras linhas do quadro. A única combinação de spin total e momento angular orbital compatível com a estatística de Fermi-Dirac é portanto a que aparece na última linha, mas ela corresponde a um estado de paridade *negativa*. Em vista das propriedades do estado inicial, isso implicaria numa mudança de paridade na transição em que o pion é absorvido. Como uma enorme variedade de fatos experimentais indica que as interações fortes, responsáveis pelo processo de absorção do pion, são interações que conservam a paridade, uma solução para essa questão pode ser buscada nas propriedades intrínsecas do pion, presente no estado inicial mas não no estado final dessa transição. De fato é possível preservar a paridade nesse processo atribuindo uma *paridade intrínseca* negativa ao pion, o que torna também *negativa* a paridade do estado inicial.

Essa idéia pode ser formalizada através de uma extensão simples da representação (R_P, R_1) do grupo de inversão espacial considerada acima e também na seção 9.1.2. Ali o operador R_P

foi definido pela relação

$$(R_P \psi)(q) = \psi(P^{-1}q) \quad (9.24)$$

válida para qualquer função $\psi(q)$ do espaço de Hilbert. A definição do operador que representa a operação de inversão pode no entanto ser modificada substituindo a relação (9.24) por

$$R_P^{(I)} \psi(q) = I \psi(P^{-1}q) \quad (9.25)$$

onde $I \equiv \pm \hat{1}$. Essa relação na realidade define *dois* operadores $R_P^{(I)}$, correspondentes a cada uma das alternativas adotadas para I . Comparação com a definição (9.24) mostra imediatamente que esses novos operadores se relacionam com R_P através de

$$R_P^{(I)} = I R_P, \quad I = \pm \hat{1}.$$

Mantendo a definição $R_1^{(I)} = R_1 = \hat{1}$, é imediato verificar que $(R_P^{(I)}, R_1^{(I)})$ são representações do grupo de inversão espacial para ambas as escolhas de I . O fator I pode ser associado à *paridade intrínseca* do sistema. O fato de que o π_- no processo considerado se comporta como tendo paridade intrínseca negativa leva à atribuição $I = (-1)^{n_\pi} \hat{1}$ para um sistema contendo n_π pions. Isso não apenas dá conta da conservação da paridade no processo de desintegração de deuteron por captura de pions, mas implica em relações de paridade em outros processos envolvendo pions que podem ser experimentalmente testadas.

Finalmente, a definição (9.25) leva agora às *quatro* representações irredutíveis

$$R_1^{(I+)} = \hat{1}, \quad R_P^{(I+)} = I \quad \text{e} \quad R_1^{(I-)} = \hat{1}, \quad R_P^{(I-)} = -I, \quad I = \pm \hat{1}$$

que, respectivamente, caracterizam as funções de onda de paridade positiva e negativa de sistemas com paridade intrínseca positiva ($I = +\hat{1}$) e negativa ($I = -\hat{1}$). Um esquema de classificação envolvendo a paridade das funções de onda e a paridade intrínseca do sistema envolve portanto quatro tipos de estados. Há, de fato dois tipos de estado de paridade positiva, estados associados a funções de onda pares de sistemas com paridade intrínseca positiva e estados associados a funções de onda ímpares de sistemas com paridade intrínseca negativa, bem como dois tipos de estado de paridade negativa.

9.6 Invariança por translações temporais e a lei dinâmica de evolução.

Na mecânica quântica de partículas as variáveis espaciais e o tempo desempenham papéis claramente diferentes, pois enquanto as variáveis espaciais estão diretamente associadas a variáveis

dinâmicas quantizadas (operadores de posição), o tempo aparece como um parâmetro externo associado à descrição da evolução dinâmica do sistema considerado⁵. Em cada tempo t o estado do sistema é descrito por um vetor de estado $|\varphi(t)\rangle$ que pode ser representado, na base dos autovetores (impróprios) $\{|q\rangle\}$ dos operadores de posição relevantes para o sistema considerado, por funções de onda dependentes do tempo $\varphi(q, t) \equiv \langle q|\varphi(t)\rangle$. A dependência do vetor de estado (ou da função de onda) com o tempo é dada pelo operador de evolução unitário $U(t, t_0)$ que é a solução da equação diferencial

$$i\hbar \frac{dU(t, t_0)}{dt} = H(t)U(t, t_0)$$

com a condição inicial $U(t_0, t_0) = 1$, onde $H(t)$ é o hamiltoniano (que em geral pode depender explicitamente do tempo) que descreve a dinâmica do sistema ou, equivalentemente, da equação integral

$$U(t, t_0) = \hat{1} - \frac{i}{\hbar} \int_{t_0}^t dt' H(t') U(t', t_0). \quad (9.26)$$

A unitariedade dos operadores $U(t, t_0)$ garante a existência dos seus respectivos inversos como $U^{-1}(t, t_0) = U^\dagger(t, t_0)$. Além disso existe uma lei de composição *para intervalos de tempo consecutivos*, $U(t_1, t)U(t, t_0) = U(t_1, t_0)$. Para ver isso basta escrever a equação (9.26) para este último operador separando a integral em duas partes

$$U(t_1, t_0) = 1 - \frac{i}{\hbar} \int_{t_0}^t dt' H(t') U(t', t_0) - \frac{i}{\hbar} \int_t^{t_1} dt' H(t') U(t', t_0)$$

e notar que os dois primeiros termos correspondem a $U(t, t_0)$. Por outro lado

$$U(t_1, t) = 1 - \frac{i}{\hbar} \int_t^{t_1} dt' H(t') U(t', t)$$

donde segue que

$$U(t_1, t)U(t, t_0) = U(t, t_0) - \frac{i}{\hbar} \int_t^{t_1} dt' H(t') U(t', t) U(t, t_0).$$

⁵É claro que esta assimetria no tratamento de variáveis espaciais e do tempo é extremamente problemática para uma formulação relativística. Ela é fortemente atenuada, porém, no contexto de uma teoria de campos quantizada, em que o processo de quantização é aplicado aos valores assumidos pelos campos, dando às variáveis espaciais um papel de rótulos externos semelhante ao que é desempenhado pelo tempo. Esse fato determina, em última análise, a preferência por teorias de campo quantizadas, no contexto relativístico, em detrimento de uma mecânica quantizada de partículas.

Portanto os operadores $U(t_1, t)U(t, t_0)$ e $U(t_1, t_0)$ satisfazem a mesma equação integral (ou, equivalentemente, a mesma equação diferencial com a mesma condição inicial), e portanto são iguais. A restrição a intervalos de tempo consecutivos da lei de composição a torna evidentemente insuficiente para definir uma estrutura de grupo sobre o conjunto de operadores de evolução.

A relação entre o vetor de estado inicial $|\varphi_0\rangle$ do sistema, quando esse estado é preparado no instante t_0 , e o vetor de estado no tempo t é dada em termos do operador de evolução como

$$|\varphi(t)\rangle = U(t, t_0)|\varphi_0\rangle.$$

Para uma escolha genérica do estado inicial $|\varphi_0\rangle$, o vetor de estado $|\varphi(t)\rangle$ depende em geral não apenas do tempo transcorrido desde a preparação do estado inicial, $\Delta t \equiv t - t_0$, mas também do valor tomado para o tempo inicial t_0 . De fato, devido à dependência explícita do hamiltoniano com o tempo, os operadores de evolução $U(t_0 + \Delta t, t_0)$ dependem tanto de Δt como de t_0 . Uma questão relevante é portanto determinar condições sobre a dinâmica do sistema que garantam que a evolução temporal de um estado inicial genérico dependa unicamente do tempo transcorrido desde a preparação, e não do particular instante em que a preparação ocorre. Essa propriedade é na realidade a análoga temporal da propriedade de invariância translacional, que exige que as propriedades do sistema dependam apenas das posições relativas de seus constituintes. Ela é expressa pela condição

$$U(t_0 + \Delta t, t_0)|\varphi_0\rangle = U(\Delta t, 0)|\varphi_0\rangle$$

para quaisquer valores de t_0 e de Δt , e para qualquer estado inicial $|\varphi_0\rangle$. Isso significa que os operadores de evolução devem ser iguais para quaisquer t_0 e Δt . Tomando um intervalo de tempo $\Delta t = \delta t$ infinitesimal

$$\begin{aligned} U(t_0 + \delta t, t_0) &= 1 + \delta t \left. \frac{dU(t, t_0)}{dt} \right|_{t=t_0} + \mathcal{O}(\delta t^2) = 1 - \frac{i}{\hbar} \delta t H(t_0) + \mathcal{O}(\delta t^2) = \\ &= 1 - \frac{i}{\hbar} \delta t H(0) + \mathcal{O}(\delta t^2), \end{aligned}$$

que deve valer para qualquer t_0 , o que implica que o hamiltoniano não deve depender explicitamente do tempo, sendo portanto uma constante do movimento (v. eq. (2.24)). Neste caso a equação (9.26) admite a solução formal

$$U(t, t_0) = e^{-\frac{i}{\hbar} H (t-t_0)}$$

na qual a dependência apenas da diferença do tempo inicial e do tempo final é explícita.

9.7 Inversão temporal.

A evolução temporal de um sistema físico pode ser descrita *classicamente* dando a configuração (isto é, a localização espacial) de seus constituintes j em função do tempo, $q_k(t)$. Neste contexto, a operação de inversão temporal T pode ser caracterizada simplesmente como consistindo na inversão do sentido do tempo nessa descrição, isto é

$$q_k(t) \xrightarrow{T} q'_k(t) \equiv q_k(t_0 - t) \quad (9.27)$$

onde t_0 é um instante qualquer que seja escolhido para definir a origem dos tempos para o movimento temporalmente invertido. A evolução alternativa $q'_k(t)$, resultante dessa operação, pode ser ou não, por sua vez, uma solução possível das equações de movimento do sistema. Quando a versão temporalmente invertida de *qualquer* evolução consistente com as equações de movimento também for consistente com essas equações (isto é, quando ela for também uma das possíveis evoluções temporais do sistema), este é dito *invariante* sob inversão temporal.

Por outro lado, as equações de movimento clássicas são tipicamente equações de segunda ordem no tempo (ou, em uma formulação hamiltoniana, pares de equações acopladas de primeira ordem), e por isso exigem mais que a especificação da configuração do sistema num dado instante (por exemplo, $q_k(t_0)$) para a determinação unívoca de uma dada evolução temporal. Uma especificação do estado do sistema que é suficiente para esse fim é aquela que inclui também os valores das velocidades nesse mesmo instante, $\dot{q}_k(t_0)$, onde o ponto indica diferenciação com relação ao argumento. Como a operação de inversão temporal inverte também as velocidades presentes no sistema, a questão da invariança sob inversão temporal deve ser examinada comparando evoluções determinadas por condições subsidiárias envolvendo a mesma configuração mas velocidades opostas no instante considerado (por exemplo, $q'_k(0) = q_k(t_0)$ e $\dot{q}'_k(0) = -\dot{q}_k(t_0)$). A especificação da evolução temporal através das funções do tempo $q_k(t)$ permite implementar essa comparação sem maiores problemas, através do cálculo das derivadas temporais pertinentes.

A evolução temporal de um sistema *quântico*, no entanto, é dada em termos da dependência temporal de um vetor de estado $|\varphi(t)\rangle$, o qual deve satisfazer uma equação de movimento que é *de primeira ordem* no tempo. Desse modo, a especificação da condição inicial $|\varphi(t=0)\rangle = |\varphi_0\rangle$ determina completamente a sua evolução temporal subsequente. De fato, o vetor de estado *em um instante dado* contém não apenas informação acerca da configuração do sistema, através da distribuição de probabilidade $|\langle q|\varphi(t)\rangle|^2$, onde q designa coletivamente os autovalores dos operadores de posição pertinentes, mas contém também informação acerca da densidade de corrente de probabilidade, através das quantidades (cf. eq. (3.15))

$$j_k(q, t) = \text{Re} \left(\langle \varphi(t) | q \rangle \langle q | \frac{p_k}{m_k} | \varphi(t) \rangle \right)$$

onde os j_k são componentes da densidade de corrente de probabilidade, os p_k são operadores de momento e os m_k as massas associadas aos p_k . Uma definição razoável do que se deva entender por inversão temporal no contexto da mecânica quântica deve incluir a troca de sinal de todas as densidades de corrente de probabilidade sem alterar as densidades de probabilidade, que correspondem às configurações espaciais do sistema. Isso exige evidentemente modificações mais profundas da sequência temporal de estados $|\varphi(t)\rangle$ que a simples inversão do sentido do tempo ao longo da sequência.

Para chegar a uma formulação satisfatória da operação de inversão temporal na mecânica quântica é conveniente considerar inicialmente o caso de mais simples de um sistema com análogo clássico, cujas variáveis dinâmicas podem ser escritas em termos de operadores de posição q_k e seus momentos canonicamente conjugados p_k , os quais satisfazem as relações canônicas de comutação

$$[q_j, p_k] = i\hbar \delta_{jk}.$$

A dinâmica desse sistema será tomada como sendo determinada por um hamiltoniano que não depende explicitamente do tempo, de modo a assegurar a invariância por translações temporais, como discutido na seção anterior. Nessas condições, a definição (9.27) dada para o caso clássico deve ser substituída, no caso quântico, por

$$|\varphi(t)\rangle \xrightarrow{T} |\varphi'(t)\rangle \equiv K|\varphi(t_0 - t)\rangle \quad (9.28)$$

onde K é um operador a determinar cuja função será a de inverter o sinal das densidades de corrente de probabilidade sem no entanto alterar as densidades de probabilidade, qualquer que seja o estado considerado do sistema.

A condição de que a ação de K não altere as densidades de probabilidade para qualquer estado $|\varphi(t)\rangle$, isto é

$$|\langle q|K|\varphi(t)\rangle|^2 = |\langle q|\varphi(t)\rangle|^2, \quad q \text{ e } |\varphi(t)\rangle \text{ quaisquer}$$

implica não apenas que K preserva a norma (dada pela integral da densidade de probabilidade) mas também que as duas amplitudes de probabilidade que intervêm nessa relação só podem diferir por um fator de fase, eventualmente dependente de q , isto é

$$\langle q|K|\varphi(t)\rangle = e^{i\xi_q} \langle q|\varphi(t)\rangle.$$

Essa condição pode ser satisfeita tomando para K um operador unitário que é diagonal na base dos autovetores (impróprios) dos operadores de posição, isto é

$$K^\dagger = K^{-1}, \quad K|q\rangle = e^{-i\xi_q}|q\rangle,$$

o que implica que ele *comuta* com os operadores de posição. No entanto, essa escolha para o operador K é *inconsistente* com a condição de que ele deve também inverter o sinal das densidades de corrente. De fato, as densidades de corrente para o estado transformado por K são, nessas condições,

$$\text{Re} \left(\langle \varphi(t) | K^\dagger | q \rangle \langle q | K K^\dagger \frac{p_k}{m_k} K | \varphi(t) \rangle \right) = \text{Re} \left(\langle \varphi(t) | q \rangle \langle q | \frac{K^\dagger p_k K}{m_k} | \varphi(t) \rangle \right)$$

o que exige, para troca de sinal, que $K^\dagger p_k K = -p_k$. Essa condição é porém incompatível com as relações de comutação canônicas, pois ela exige que

$$[q_j, K^\dagger p_k K] = -i\hbar \hat{1} \delta_{jk}$$

enquanto a aplicação direta da transformação K sobre essas relações mantém o sinal do segundo membro.

Essa inconsistência pode ser evitada se o operador K for escolhido não como sendo um operador linear, mas um operador *anti-linear*. Operadores anti-lineares diferem dos operadores lineares quanto à propriedade

$$K(z_1 |\varphi_1\rangle + z_2 |\varphi_2\rangle) = z_1^* K |\varphi_1\rangle + z_2^* K |\varphi_2\rangle$$

que difere da propriedade correspondente dos operadores lineares pela conjugação complexa dos coeficientes da combinação linear de estados considerada. Essa diferença exige, por consistência, que o *adjunto* K^\dagger do operador anti-linear K seja definido pela relação

$$\langle \varphi_1 | K | \varphi_2 \rangle = \langle \varphi_2 | K^\dagger | \varphi_1 \rangle \quad K \text{ anti-linear, quaisquer } |\varphi_1\rangle, |\varphi_2\rangle \quad (9.29)$$

que difere da definição usual do adjunto g^\dagger de um operador linear g

$$\langle \varphi_1 | g | \varphi_2 \rangle = \langle \varphi_2 | g^\dagger | \varphi_1 \rangle^* \quad g \text{ linear, quaisquer } |\varphi_1\rangle, |\varphi_2\rangle \quad (9.30)$$

por uma conjugação complexa adicional. Um operador anti-linear que tem a propriedade

$$K K^\dagger = K^\dagger K = \hat{1}$$

é chamado *antiunitário*. É imediato verificar que operadores antiunitários preservam a norma, mas introduzem uma conjugação complexa adicional no produto escalar de dois vetores diferentes. De fato, supondo K anti-unitário,

$$\langle \psi | K | \varphi_1 \rangle = \langle \varphi_1 | K^\dagger | \psi \rangle$$

e, tomando $|\psi\rangle = K|\varphi_2\rangle$, essa quantidade se torna o produto escalar desse vetor por $K|\varphi_1\rangle$ e é dada por

$$\langle\varphi_1|K^\dagger K|\varphi_2\rangle = \langle\varphi_2|\varphi_1\rangle^*.$$

A norma de $K|\varphi_1\rangle$ é porém obtida fazendo $|\varphi_2\rangle = |\varphi_1\rangle$, o que reproduz a norma deste último vetor.

A consistência da definição (9.29) com a anti-linearidade pode ser verificada aplicando-a ao elemento de matriz

$$\langle\psi|K(z_1|\varphi_1\rangle + z_2|\varphi_2\rangle) = z_1^*\langle\psi|K|\varphi_1\rangle + z_2^*\langle\psi|K|\varphi_2\rangle$$

onde a conjugação complexa de z_1 e z_2 resulta da anti-linearidade de K . Usando agora a definição modificada (9.29) para K^\dagger , os dois elementos de matriz do lado direito dessa relação podem ser escritos como

$$\langle\psi|K|\varphi_1\rangle = \langle\varphi_1|K^\dagger|\psi\rangle \quad \text{e} \quad \langle\psi|K|\varphi_2\rangle = \langle\varphi_2|K^\dagger|\psi\rangle. \quad (9.31)$$

Usando essas expressões, o elemento de matriz inicial fica dado em termos de K^\dagger como

$$\begin{aligned} \langle\psi|K(z_1|\varphi_1\rangle + z_2|\varphi_2\rangle) &= z_1^*\langle\varphi_1|K^\dagger|\psi\rangle + z_2^*\langle\varphi_2|K^\dagger|\psi\rangle = \\ &= \langle z_1\varphi_1 + z_2\varphi_2 | K^\dagger |\psi\rangle \end{aligned}$$

onde $\langle z_1\varphi_1 + z_2\varphi_2 |$ é o bra associado ao ket $z_1|\varphi_1\rangle + z_2|\varphi_2\rangle$ e foi feito uso da anti-linearidade do produto escalar com relação ao seu primeiro argumento. Esta última relação é consistente com a definição modificada (9.29) do operador adjunto K^\dagger . Se esse operador tivesse sido definido utilizando a relação (9.30), haveria uma conjugação complexa adicional nas relações (9.31), o que levaria a

$$\langle\psi|K(z_1|\varphi_1\rangle + z_2|\varphi_2\rangle) = \langle z_1^*\varphi_1 + z_2^*\varphi_2 | K^\dagger |\psi\rangle^*.$$

Esse resultado é inconsistente com a definição (9.30) devido à conjugação complexa das constantes z_1 e z_2 .

A operação de inversão temporal pode então ser caracterizada, no contexto da mecânica quântica de sistemas com análogo clássico, pela relação (9.28), na qual o operador K é um operador *anti-unitário* com as propriedades

$$K^\dagger q_j K = q_j \quad \text{e} \quad K^\dagger p_k K = -p_k$$

que, em vista da antilinearidade de K , é agora consistente com as relações de comutação canônicas. A primeira dessas relações garante a compatibilidade de K com os operadores de posição e através disso garante também a preservação das densidades de probabilidade. A troca de sinal dos momentos, por outro lado, garante a troca de sinal das densidades de corrente. Desse modo, cada um dos estados da seqüência temporalmente invertida $|\varphi'(t)\rangle$ tem, por sua vez, densidades de probabilidade iguais mas distribuições de coerente invertidas em relação aos estados da seqüência original.

A determinação explícita da transformação anti-unitária K depende da particular representação adotada para os operadores e vetores de estado. No caso da representação de Schrödinger, em que os operadores de posição são representados por operadores de multiplicação (reais) e os operadores de momento são representados por operadores diferenciais da forma

$$p_k \longrightarrow \frac{\hbar}{i} \frac{\partial}{\partial q_k}$$

a escolha mais simples e usual é tomar K como sendo o operador de conjugação complexa $K_0 = K_0^\dagger$. Nesse caso, de fato

$$K_0 q_j K_0 = q_j \quad \text{e} \quad K_0 \frac{\hbar}{i} \frac{\partial}{\partial q_k} K_0 = -\frac{\hbar}{i} \frac{\partial}{\partial q_k}.$$

9.7.1 Sistemas com spin.

A transformação antiunitária do operador correspondente ao momento angular orbital, no caso de sistemas com análogo clássico, é dada por

$$K^\dagger \vec{L} K = K^\dagger \sum_j (\vec{r}_j \times \vec{p}_j) K = \sum_j \vec{r}_j \times (-\vec{p}_j) = -\vec{L}$$

consistentemente com a transformação do momento angular clássico sob inversão temporal. No caso em que o sistema considerado envolver partículas com spin, é preciso, por consistência, que o comportamento sob inversão temporal dos operadores associados ao spin seja igual ao do momento angular orbital, isto é

$$K^\dagger \vec{S} K = \sum_j K^\dagger \vec{s}_j K = \sum_j (-\vec{s}_j).$$

Essa relação deve de fato ser vista como uma condição adicional a ser satisfeita, no caso, pelo operador anti-unitário K . Ela garante também propriedades de transformação adequadas para o momento angular total $\vec{J} = \vec{L} + \vec{S}$. A determinação explícita de K depende também neste caso

da representação adotada tanto para as variáveis dinâmicas com análogo clássico como para os operadores de spin. No caso de ser usada a representação de Schrödinger, e para partículas de spin 1/2 com operadores de spin escritos em termos das matrizes de Pauli $\vec{\sigma}^{(j)} \equiv \{\sigma_1^{(j)}, \sigma_2^{(j)}, \sigma_3^{(j)}\}$ como

$$\vec{S} = \sum_j \vec{s}_j = \sum_j \frac{\hbar}{2} \vec{\sigma}^{(j)}$$

esse operador pode ser realizado como $K = K_0 T$, onde K_0 é o operador de conjugação complexa e T é um operador *unitário* agindo sobre os spins. Desse modo o que se exige é que

$$T^\dagger K_0 \sum_j \vec{\sigma}^{(j)} K_0 T = \sum_j \left(-\vec{\sigma}^{(j)} \right).$$

Como as matrizes de Pauli σ_1 e σ_3 são reais e a matriz σ_2 é puramente imaginária,

$$K_0 \sigma_1 K_0 = \sigma_1, \quad K_0 \sigma_2 K_0 = -\sigma_2 \quad \text{e} \quad K_0 \sigma_3 K_0 = \sigma_3$$

de modo que a ação do operador T deve ser tal que

$$T^\dagger \sigma_1 T = -\sigma_1, \quad T^\dagger \sigma_2 T = \sigma_2 \quad \text{e} \quad T^\dagger \sigma_3 T = -\sigma_3$$

o que corresponde a uma rotação de um ângulo π em torno do eixo 2, isto é

$$T = e^{\frac{i}{\hbar} \pi \hat{2} \cdot \vec{S}} = e^{i \frac{\pi}{2} \sum_j \sigma_2^{(j)}} = \prod_j e^{i \frac{\pi}{2} \sigma_2^{(j)}} = \prod_j i \sigma_2^{(j)}. \quad (9.32)$$

Casos em que há partículas com outros valores do spin, ou quando se utilize representações alternativas para as variáveis de spin, podem ser tratados de forma semelhante. Também semelhante é o caso de partículas envolvendo outros graus de liberdade internos além do spin. No caso de partículas com isospin $t = 1/2$, com operadores de isospin representado em termos de matrizes de Pauli agindo no espaço de carga, $\vec{t}_j = \vec{\tau}^{(j)}$, a suposição de que a inversão temporal não afeta esse espaço se traduz em termos das condições sobre K

$$K^\dagger \vec{\tau}^{(j)} K = \vec{\tau}^{(j)}.$$

realizando esse operador novamente como $K = K_0 \tilde{T}$ as condições a serem exigidas de \tilde{T} são agora

$$\tilde{T}^\dagger \tau_1 \tilde{T} = \tau_1, \quad \tilde{T}^\dagger \tau_2 \tilde{T} = -\tau_2 \quad \text{e} \quad \tilde{T}^\dagger \tau_3 \tilde{T} = \tau_3$$

que podem ser realizadas em termos de uma inversão no espaço de carga seguida de uma rotação de π em torno do eixo 2 desse mesmo espaço.

9.7.2 Invariança sob inversão temporal.

A questão da *invariança* da dinâmica sob a operação de inversão temporal definida como na equação (9.28), com a realização apropriada do operador anti-unitário K , depende da natureza do hamiltoniano em termos do qual é escrita a equação de movimento. A forma geral dessa equação é (no caso de um Hamiltoniano que não depende explicitamente do tempo)

$$H|\varphi(t)\rangle = i\hbar \frac{d|\varphi(t)\rangle}{dt}$$

donde segue que a equação de movimento satisfeita pela solução temporalmente invertida é

$$KHK^\dagger|\varphi'(t)\rangle = i\hbar \frac{d|\varphi'(t)\rangle}{dt}.$$

A forma do segundo membro dessa equação é igual à da equação de partida devido a uma *dupla* troca de sinal, devida respectivamente à transformação da variável temporal $t \rightarrow t_0 - t$ e à antilinearidade da transformação anti-unitária K . O sistema considerado será dito invariante sob a operação de inversão temporal sempre que essas duas equações de movimento coincidirem, isto é

$$KHK^\dagger = H \quad \text{ou} \quad [K, H] = 0.$$

Nesse caso, portanto, tanto a solução $|\varphi(t)\rangle$ como a solução temporalmente invertida $|\varphi'(t)\rangle = K|\varphi(t_0 - t)\rangle$ são soluções *da mesma* equação de movimento, satisfazendo condições iniciais relacionadas entre si por

$$|\varphi'(0)\rangle = K|\varphi(t_0)\rangle.$$

Uma forma particularmente sugestiva de exprimir esse fato é considerar a quantidade

$$P_{a \rightarrow b}(t) \equiv \left| \langle \varphi_b | e^{-\frac{i}{\hbar} H t} | \varphi_a \rangle \right|^2$$

que representa a probabilidade de que o sistema cuja dinâmica é descrita pelo hamiltoniano (independente do tempo) H , quando preparado no estado $|\varphi_a\rangle$ no instante zero seja encontrado no estado $|\varphi_b\rangle$ no instante t . Se a dinâmica determinada por H for invariante sob inversão temporal, o elemento de matriz que aparece nessa expressão pode ser re-escrito sucessivamente como

$$\begin{aligned} \langle \varphi_b | e^{-\frac{i}{\hbar} H t} | \varphi_a \rangle &= \langle \varphi_b | K^\dagger e^{\frac{i}{\hbar} H t} K | \varphi_a \rangle = \langle \varphi_b | K^\dagger e^{\frac{i}{\hbar} H t} | \varphi'_a \rangle = \\ &= \langle \varphi'_a | e^{-\frac{i}{\hbar} H t} K | \varphi_b \rangle = \langle \varphi'_a | e^{-\frac{i}{\hbar} H t} | \varphi'_b \rangle \end{aligned}$$

onde foi usada a relação que define o adjunto de um operador anti-linear e foram introduzidas as definições $|\varphi'_{a,b}\rangle \equiv K|\varphi_{a,b}\rangle$. Estes estados correspondem a versões temporalmente invertidas (no sentido que densidades de corrente e spins são invertidos) dos estados inicial e final envolvidos em $P_{a\rightarrow b}(t)$, de modo que a invariância sob inversão temporal garante a igualdade das probabilidades

$$P_{a\rightarrow b}(t) = P_{b'\rightarrow a'}(t) \quad (9.33)$$

que correspondem respectivamente à probabilidade de obter o estado b um tempo t após a preparação do estado a e à probabilidade de obter o inverso temporal do estado $|\varphi_a\rangle$ um tempo t após a preparação do estado temporalmente invertido de $|\varphi_b\rangle$. Deve-se notar que os papéis dos estados inicial e final são invertidos nessas duas probabilidades. A relação (9.33) é conhecida como *princípio de micro-reversibilidade*, e é na realidade equivalente à invariância sob inversão temporal. O princípio de micro-reversibilidade se aplica *aos quadrados dos módulos dos elementos de matriz* e não a probabilidades de transição nas quais intervenham fatores de espaço de fases, como por exemplo seções de choque em problemas de espalhamento. De fato, a troca de papéis dos estados inicial e final implica em mudanças de espaço de fases que em geral tornarão estas probabilidades de transição desiguais.

A invariância sob inversão temporal equivale à compatibilidade do operador K com o hamiltoniano, e portanto tem conseqüências também para o espectro de energia do sistema. No caso de um sistema com *spin total inteiro* o operador anti-unitário K satisfaz à relação $K^2 = \hat{1}$. Essa propriedade é imediata para sistemas com análogo clássico tratados na representação em que K pode ser realizado como o operador de conjugação complexa K_0 , e pode ser verificada facilmente para sistemas com um número *par* de partículas de spin $1/2$, tratadas em termos do operador $K = K_0T$, com T definido por (9.32). Nesse caso os possíveis autovalores de K são $\kappa = \pm 1$, e existe um conjunto completo de autovetores comuns de H e K

$$H|\phi_{E\kappa}\rangle = E|\phi_{E\kappa}\rangle, \quad K|\phi_{E\kappa}\rangle = \kappa|\phi_{E\kappa}\rangle, \quad \kappa = \pm 1.$$

Os autovetores de K com autovalores $\kappa = +1$ e $\kappa = -1$ são ditos respectivamente *temporalmente pares* e *temporalmente ímpares*, ou “reais” e “imaginários”. Estas últimas denominações são motivadas pelo fato de se poder ver o operador anti-unitário K_0T como uma “conjugação complexa generalizada”. No entanto, se $|\phi_{E\kappa=-1}\rangle$ é um vetor temporalmente ímpar, é claro que $i|\phi_{E\kappa=-1}\rangle$ é temporalmente par, o que significa que *todos* os autovetores de um hamiltoniano invariante sob inversão temporal de um sistema com spin total inteiro podem ser escolhidos de forma a serem vetores de estado temporalmente pares.

As questões relativas à inversão temporal em um sistema *sem spin* são bem ilustradas pela evolução temporal de um pacote de ondas gaussiano determinada pelo hamiltoniano livre $H = p^2/2M$, que foi extensamente discutida na seção 3.3.3, e particularmente pela

obtenção de condições iniciais cuja evolução temporal leva a um *estreitamento* do pacote (v. pag. 145 e seguintes). Na representação de Schrödinger, o operador anti-unitário apropriado pode ser tomado como o operador de conjugação complexa K_0 , e a realidade de H implica portanto na invariança sob inversão temporal. A base de ondas planas, autofunções simultâneas do momento \vec{p} e do hamiltoniano, envolve funções de onda complexas, mas o espectro do hamiltoniano é fortemente degenerado e permite a construção de uma base alternativa de autofunções de H *reais* (ou temporalmente pares)

$$\phi_k^{(+)}(\vec{r}) = \sqrt{\frac{2}{L^3}} \cos \vec{k} \cdot \vec{r} \quad \text{e} \quad \phi_k^{(-)}(\vec{r}) = \sqrt{\frac{2}{L^3}} \sin \vec{k} \cdot \vec{r}.$$

A condição inicial gaussiana com momento médio $\hbar \vec{k}_0$ e posição média \vec{r}_0

$$\varphi(\vec{r}, t = 0) = \left[\frac{1}{b\sqrt{\pi}} \right]^{\frac{3}{2}} e^{i\vec{k}_0 \cdot \vec{r}} e^{-\frac{(\vec{r} - \vec{r}_0)^2}{2b^2}} \quad (9.34)$$

é também uma função complexa, da qual a inversa temporal $K_0 \varphi(\vec{r}, t = 0) = \varphi^*(\vec{r}, t = 0)$ difere apenas pela inversão do momento, $\vec{k}_0 \rightarrow -\vec{k}_0$. É fácil ver que essa transformação inverte a densidade de corrente do pacote original. A evolução temporal da condição inicial invertida temporalmente difere da evolução temporal da condição inicial original apenas por uma inversão na direção do deslocamento do pacote. Em particular, nos dois casos há um alargamento idêntico da distribuição espacial de probabilidade.

Quando uma nova condição inicial é definida pela inversão temporal do pacote de ondas evoluído até o tempo $t_0 = T$ a partir da condição inicial (9.34), isto é

$$\varphi'(\vec{r}, t = 0) = K_0 \varphi(\vec{r}, T) = \varphi^*(\vec{r}, T),$$

a sua respectiva evolução temporal recapitulará, durante um tempo também igual a T e em seqüência temporal inversa, a evolução temporal do pacote original, consistentemente com o princípio de micro-reversibilidade. O alargamento do pacote original dá lugar portanto a um estreitamento na dinâmica temporalmente invertida. Neste caso a operação de inversão temporal representada por K_0 implica não apenas na mudança de sinal no expoente da onda plana mas também na troca de sinal dos demais termos complexos que aparecem na forma evoluída do pacote gaussiano, dada pela equação (3.31). Isso é necessário, de fato, para que a densidade de corrente seja efetivamente invertida em todos os pontos.

No caso de um sistema com spin total semi-inteiro, no entanto $K^2 = -1$, como pode ser visto, por exemplo, no caso de um sistema contendo um número ímpar de partículas de spin $1/2$, em que há portanto um número também ímpar de fatores no operador T definido em (9.32). Uma consequência da anti-unitaridade de K é, nesse caso

$$K^\dagger = -K^2 K^\dagger = -K$$

de modo que

$$\langle \varphi | K | \varphi \rangle = \langle \varphi | K^\dagger | \varphi \rangle = -\langle \varphi | K | \varphi \rangle$$

o que implica em

$$\langle \varphi | K | \varphi \rangle = 0$$

para qualquer vetor de estado $|\varphi\rangle$. Isso significa que o vetor $K|\varphi\rangle$, cuja norma é igual à norma de $|\varphi\rangle$, é *ortogonal* a esse vetor. Então, dado um autovetor qualquer de H com autovalor E , isto é

$$H|\phi_E\rangle = E|\phi_E\rangle$$

a invariança sob inversão temporal implica em que se tenha também

$$KH|\phi_E\rangle = HK|\phi_E\rangle = EK|\phi_E\rangle,$$

isto é, o vetor $K|\phi_E\rangle$ é também autovetor de H com o mesmo autovalor E . Como esse vetor é ortogonal a $|\phi_E\rangle$, segue que os autovalores de H são *pelo menos* duplamente degenerados. Além disso, se existe um terceiro autovetor $|\phi'_E\rangle$ de H associado ao mesmo autovalor, ortogonal tanto a $|\phi_E\rangle$ como a $K|\phi_E\rangle$, então é fácil verificar que $K|\phi'_E\rangle$ é também autovetor de H , associado ao mesmo autovalor E , e ortogonal a cada um dos três vetores considerados. Isso mostra portanto que a degenerescência dos autovalores do hamiltoniano é, nas condições consideradas, sempre *par*. A degenerescência par, decorrente da invariança sob inversão temporal de sistemas com spin total ímpar é chamada *degenerescência de Kramers*.

Problemas.

9.1 - Considere as transformações de coordenadas, correspondentes a rotações do referencial de ângulos φ em torno do eixo z , descritas por

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \equiv \Omega_z(\varphi) \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

a) Verifique a lei de composição de grupo $\Omega_z(\varphi_2)\Omega_z(\varphi_1) = \Omega_z(\varphi_1 + \varphi_2)$ para essas transformações.

b) Mostre que o grupo dessas transformações é comutativo (abeliano).

c) O gerador ω_z desse grupo é definido através da relação

$$\Omega_z(d\varphi) = 1 + d\varphi \omega_z + \mathcal{O}(d\varphi^2).$$

Verifique que

$$\omega_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e mostre, expandindo a exponencial, usando a forma especial do gerador e re-somando a expansão, que

$$\Omega_z(\varphi) = e^{\varphi \omega_z}.$$

Sugestão: $\omega_z^3 = -\omega_z$.

d) Considere a representação desse grupo de transformações no espaço de Hilbert de funções de onda $\psi(\vec{r})$, com $\vec{r} \equiv \{x, y, z\}$, obtida pondo

$$[R_{\Omega_z}(\varphi)\psi](\Omega_z(\varphi)\vec{r}) = \psi(\vec{r}).$$

Mostre que o gerador dessa representação é dado por $-iL_z/\hbar$, isto é

$$R_{\Omega_z}(d\varphi) = 1 - \frac{i}{\hbar} d\varphi L_z$$

onde $L_z \equiv \frac{\hbar}{i}(\vec{r} \times \nabla)_z$ é o operador associado à componente z do momento angular orbital. Mostre também que o operador correspondente a uma rotação finita é $R_{\Omega_z}(\varphi) = e^{-\frac{i}{\hbar}\varphi L_z}$.

9.2 - Uma partícula está em um estado cuja parte angular $\phi_{10}(\theta, \varphi)$ é um autovetor de \vec{L}^2 e de L_z (sendo $\vec{L} \equiv \vec{r} \times \vec{p}$ o momento angular orbital da partícula) com autovalores $\hbar^2 l(l+1)$ e $\hbar m$, respectivamente, com $l = 1$ e $m = 0$. Considere uma rotação do sistema de coordenadas de um ângulo α em torno do eixo \hat{y} , com o que $\{\theta, \varphi\} \rightarrow \{\theta', \varphi'\}$, onde θ e φ são os ângulos esféricos de \vec{r} no sistema original, e θ' e φ' seus correspondentes no novo sistema.

a) Exprima $\phi'_{10}(\theta, \varphi) = e^{i\alpha L_y} \phi_{10}(\theta, \varphi)$ em termos de $\phi_{1m}(\theta, \varphi)$, $m = -1, 0, 1$. **Sugestão:** expanda a exponencial e decompõe L_y em termos de L_+ e L_- .

b) Calcule $\phi'_{10}(\theta', \varphi')$.

9.3 - A função de onda de spin de uma partícula é $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Qual é a nova função de onda u' quando o sistema de coordenadas é rodado de um ângulo α em torno do eixo \hat{y} ?

9.4 - Supondo conservação de isospin mostre que a seção de choque para o processo

$$p + p \longrightarrow d + \pi^+$$

é o dobro da seção de choque para o processo

$$p + n \longrightarrow d + \pi^0.$$

Sugestão: escreva os estados inicial e final em termos de estados de isospin total bem definido. Esses processos podem então ser associados a amplitudes que dependem apenas do isospin total dos estados inicial e final, isto é, que não dependem da carga dos sistemas inicial e final. Isto resulta do teorema de Wigner-Eckart, quando se supõe que o operador de transição é um escalar no espaço de isospin (isto é, um operador isoescalar), pois nesse caso

$$\langle I^{(f)} I_3^{(f)} | T | I^{(i)} I_3^{(i)} \rangle = C_{I_3^{(f)} 0 I_3^{(i)}}^{I^{(f)} 0 I^{(i)}} \langle I^{(f)} || T || I^{(i)} \rangle$$

onde o coeficiente de Clebsh-Gordan é simplesmente $\delta_{I^{(f)} I^{(i)}} \delta_{I_3^{(f)} I_3^{(i)}}$.

9.5 - O operador anti-unitário K associado à inversão temporal de um sistema constituído de partículas sem spin deve satisfazer as relações

$$K^\dagger q_j K = q_j \quad \text{e} \quad K^\dagger p_k K = -p_k.$$

Mostre que, na representação de momentos, em que os p_k são realizados como operadores de multiplicação, e os q_j são realizados pelos operadores diferenciais

$$q_j \longrightarrow i\hbar \frac{\partial}{\partial p_j}$$

de modo que $[q_j, p_k] = i\hbar \hat{1} \delta_{jk}$ e os vetores de estado são representados por funções dos momentos, $\tilde{\psi}(p_k)$, K pode ser realizado como $K = K_0 \mathcal{P}$, onde K_0 é o operador de conjugação complexa e \mathcal{P} é o operador de inversão espacial, definido por $\mathcal{P}\psi(q_j) = \psi(-q_j)$.

9.6 - Considere, sob o ponto de vista da operação de inversão temporal, os seguintes sistemas e hamiltonianos:

a) uma partícula de massa M e carga elétrica e sujeita a um campo magnético externo, constante (isto é, independente do tempo) e uniforme \vec{B} . Para este sistema

$$H_a = \frac{1}{2M} \left(\vec{p} - \frac{e}{c} \vec{A}(\vec{r}) \right)^2 \quad \text{com} \quad \vec{A}(\vec{r}) = \frac{1}{2} \vec{B} \times \vec{r};$$

b) um spin $1/2$, ao qual está associado um momento magnético $\vec{\mu} = g\mu_0\vec{s}$, onde μ_0 é o magneton apropriado, g é o fator giromagnético e $\vec{s} = \hbar\vec{\sigma}/2$, sendo $\vec{\sigma}$ as matrizes de Pauli, imerso em um campo magnético externo constante e uniforme \vec{B} . Para este sistema

$$H_b = -\vec{\mu} \cdot \vec{B};$$

e, finalmente,

c) o ion He^+ (que consiste de um núcleo de spin zero e carga $+2e$ com um único elétron externo), incluindo um termo de interação spin-órbita, na aproximação em que a massa do núcleo é considerada infinita. Para este sistema

$$H_c = \frac{p^2}{2m_e} - \frac{2e^2}{r} + V_{LS}(r)\vec{s} \cdot \vec{L} \quad \text{com} \quad \vec{L} = \vec{r} \times \vec{p}.$$

Mostre que, nos dois primeiros casos, os hamiltonianos não são invariantes sob inversão temporal (supondo que esta operação não modifique o campo magnético externo), mas que H_c é invariante sob inversão temporal. Tendo em vista que o termo de interação spin-órbita está relacionado com a interação do momento magnético associado ao spin do elétron com o campo magnético resultante de seu movimento orbital, discuta a diferença entre os casos **b)** e **c)** com relação à inversão temporal.

9.7 - Considere os dois processos

$$\gamma_{\vec{k}\vec{\epsilon}} + {}^1\text{H} \longrightarrow \text{p} + e_{\vec{k}'\sigma}$$

e

$$e_{-\vec{k}'-\sigma} + \text{p} \longrightarrow {}^1\text{H} + \gamma_{-\vec{k}-\vec{\epsilon}}$$

(isto é, respectivamente a foto-dissociação do hidrogênio e a captura radiativa de um electron por um próton) que são relacionados por inversão temporal. Supondo a invariança do operador de transição por inversão temporal, relacione as seções de choque para os dois processos.

Sugestão: use a invariança para relacionar os dois elementos de matriz de transição, sem tentar calcula-los explicitamente.

Referências gerais

A literatura específica dedicada à questão de transformações e simetrias na mecânica quântica é muito extensa. Referências clássicas (possivelmente difíceis) são

★ **H. Weyl**, *The theory of groups and Quantum Mechanics*, tradução para o inglês publicada por Dover, sem data.

★ **E. P. Wigner**, *Group Theory and its applications to the Quantum Mechanics of atomic spectra*, Academic Press, 1959 (tradução americana).

Textos específicos menos clássicos são, por exemplo

★ **M. Hamermesh**, *Group Theory and its application to physical problems*, Addison-Wesley, 1962,

★ **L. M. Falicov**, *Group Theory and its Physical Applications*, U. of Chicago Press, 1966.

Uma apresentação simples da teoria de grupos contínuos, inclusive de sua aplicação à classificação das partículas elementares, em pleno florescimento à época de sua publicação, pode ser encontrada em

★ **H. J. Lipkin**, *Lie Groups for Pedestrians* (Second Edition), North Holland, 1968.

Da mesma época é o artigo de revisão acompanhado de reproduções de trabalhos originais clássicos e/ou importantes para as áreas de física nuclear e física das partículas elementares

★ **F. J. Dyson**, *Symmetry Groups in Nuclear and Particle Physics*, Benjamin, 1966.

Além disso a maioria dos textos gerais dedica capítulos específicos a essas questões. Exemplos disso são

★ **A. Messiah**, *op. cit.*, Capítulos XIII (sobre rotações e momento angular, v. também o Apêndice C), XIV (que contém uma discussão do isospin) e XV (que trata de outras simetrias, notadamente da simetria sob inversão temporal).

★ **L. Landau e E. M. Lifshitz**, *op. cit.*, especialmente Capítulos IV (momento angular e rotações), VIII (spin) e XII (que inclui um tratamento de grupos finitos relevantes para a teoria quântica dos sólidos).

Existe também um grande número de textos especificamente dedicados à questão do grupo de rotações e momento angular, dos quais os mais freqüentemente citados são

★ **R. A. Edmonds**, *op. cit.*

- ★ **M. E. Rose**, *Elementary Theory of Angular Momentum*, Wiley, 1957.
- ★ **D. M. Brink e G. R. Satchler**, *Angular Momentum*, Oxford U. Press, 1962.
- ★ **J. Schwinger**, *Quantum Theory of Angular Momentum*, L. Biedenharn e H. Van Dam, Eds., Academic Press, 1965.
- ★ **U. Fano e G. Racah**, *Irreducible Tensorial Sets*, Academic Press, 1959.

Capítulo 10

Espalhamento por um potencial de curto alcance.

O espectro contínuo do Hamiltoniano de uma partícula sujeita à ação de um potencial central de curto alcance $V(r)$ foi relacionado, no capítulo 5, com os processos de espalhamento da partícula por esse potencial. Como foi também mencionado ali, o problema de dois corpos que interagem através de um potencial que depende apenas da sua posição relativa pode ser reduzido a esse caso, através de uma mudança de variáveis para coordenadas relativas e coordenadas de centro de massa. O caráter de curto alcance do potencial foi definido através da condição

$$\lim_{r \rightarrow \infty} rV(r) = 0$$

que garante que, na região assintótica $r \rightarrow \infty$, as autofunções do espectro contínuo do Hamiltoniano que são autofunções simultâneas do momento angular da partícula (“ondas parciais”) diferem das soluções livres correspondentes apenas por defasagens $\delta_l(E)$. Secções de choque de espalhamento, em particular, puderam ser expressas em termos dessas defasagens. As autofunções do espectro contínuo de particular interesse nesse contexto são as que satisfazem as *condições de contorno de espalhamento*

$$\phi_k^\pm(\vec{r}) \longrightarrow e^{i\vec{k} \cdot \vec{r}} + f_k^{(\pm)}(\hat{r}) \frac{e^{\pm ikr}}{r}, \quad k = \sqrt{\frac{2ME}{\hbar^2}} \quad (10.1)$$

para $r \rightarrow \infty$, onde \hat{r} indica as coordenadas angulares de \vec{r} . A função $f_k^{(+)}(\hat{r})$ é chamada *amplitude de espalhamento*, e se relaciona com a seção de choque diferencial simplesmente como

$$\frac{d\sigma}{d\Omega} = \left| f_k^{(+)}(\hat{r}) \right|^2.$$

Ela foi expressa em termos das defasagens das diferentes ondas parciais como

$$f_k^{(+)}(\hat{r}) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos \theta) = \left(f_k^{-}(\hat{r}) \right)^*.$$

O objetivo deste capítulo é estender esse tratamento de duas formas de certo modo complementares. Por um lado, a relação do espectro contínuo com os processos de espalhamento será examinada de uma forma fisicamente mais precisa, com base em um tratamento dependente do tempo, no qual se estuda a evolução de um pacote de ondas sob a ação do potencial. Nesse contexto é possível examinar, em particular, as condições físicas em que um tratamento como o dado anteriormente é válido. Por outro lado, o tratamento estacionário será desenvolvido formalmente através da introdução e manipulação de equações integrais equivalentes à equação de Schrödinger e que incorporam também as condições de contorno de espalhamento. Isso permite obter de forma simples e transparente uma série de propriedades gerais das funções de onda e de observáveis relevantes nos processos de colisão. As duas extensões serão tratadas em parte concomitantemente, devido ao fato de que a reformulação da equação de Schrödinger em forma integral é bastante útil para o desenvolvimento do tratamento dependente do tempo.

10.1 A equação de Lippmann-Schwinger.

A equação de autovalores relevante para o espalhamento de partículas de massa M por um potencial de curto alcance $V(\vec{r})$ (o qual, a menos de menção explícita em contrário, não precisará ser tomado aqui como sendo um potencial central) pode ser escrita como

$$\left(E - \frac{p^2}{2M} \right) |\phi_E\rangle = V |\phi_E\rangle \quad (10.2)$$

ou, utilizando a representação de Schrödinger em que \vec{r} é realizado como um operador de multiplicação

$$(k^2 + \nabla^2) \phi_{\vec{k}}(\vec{r}) = \frac{2MV(\vec{r})}{\hbar^2} \phi_{\vec{k}}(\vec{r}) \quad (10.3)$$

onde $k^2 \equiv 2ME/\hbar^2$ e a rotulação da função de onda $\phi_{\vec{k}}(\vec{r}) \equiv \langle \vec{r} | \phi_{\vec{k}} \rangle$ foi adaptada de forma a indicar mais explicitamente que ela deverá eventualmente satisfazer condições de contorno de espalhamento do tipo (10.1). Uma forma extremamente útil de tratar o conteúdo da equação diferencial (10.3) é transforma-la em uma equação integral fazendo uso das funções de Green

$$G_k^{\pm}(\vec{r}, \vec{r}') = \frac{e^{\pm ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}$$

que satisfazem a equação

$$(k^2 + \nabla^2) G_k^\pm(\vec{r}, \vec{r}') = -4\pi\delta(\vec{r} - \vec{r}')$$

e diferem entre si por suas propriedades assintóticas, para valores grandes de $|\vec{r}|$ e com $|\vec{r}'|$ finito, como discutido a seguir. Utilizando essas funções, de fato é possível escrever

$$\begin{aligned} \phi_k^\pm(\vec{r}) &= e^{i\vec{k}\cdot\vec{r}} - \frac{M}{2\pi\hbar^2} \int d^3r' G_k^\pm(\vec{r}, \vec{r}') V(\vec{r}') \phi_k^\pm(\vec{r}') \\ &= e^{i\vec{k}\cdot\vec{r}} - \frac{M}{2\pi\hbar^2} \int d^3r' \frac{e^{\pm ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} V(\vec{r}') \phi_k^\pm(\vec{r}') \end{aligned} \quad (10.4)$$

Essa é uma equação integral inhomogênea (note a presença do função a determinar $\phi_k^\pm(\vec{r})$ como um dos fatores do integrando), conhecida como *equação de Lippmann-Schwinger*¹, que desempenha um papel central na teoria formal do espalhamento. A sua solução claramente satisfaz à equação diferencial (10.3) *bem como às condições de contorno de espalhamento*, (10.1). De fato, o termo inhomogêneo de (10.4) satisfaz

$$(k^2 + \nabla^2) e^{i\vec{k}\cdot\vec{r}} = 0$$

enquanto o termo integrado reproduz o lado direito da equação (10.3). No que se refere às condições de contorno, notando que o módulo da variável de integração \vec{r}' é limitado pelo alcance do potencial $V(\vec{r}')$, é possível usar na região assintótica $|\vec{r}| \rightarrow \infty$ a expansão

$$|\vec{r} - \vec{r}'| = \sqrt{r^2 + r'^2 - 2\vec{r} \cdot \vec{r}'} \rightarrow r \left(1 - \frac{\vec{r} \cdot \vec{r}'}{r^2} + \mathcal{O}\left(\frac{r'}{r}\right)^2 \right) \simeq r - \hat{r} \cdot \vec{r}',$$

com $\hat{r} \equiv \vec{r}/r$, que permite escrever, nas mesmas condições,

$$\frac{e^{\pm ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} \rightarrow \frac{e^{\pm ikr}}{r} e^{\mp ik\hat{r}\cdot\vec{r}'}.$$

ou ainda, levando esse resultado à equação (10.4),

$$\phi_k^\pm(\vec{r}) \rightarrow e^{i\vec{k}\cdot\vec{r}} - \frac{e^{\pm ikr}}{r} \frac{M}{2\pi\hbar^2} \int d^3r' e^{\mp ik\hat{r}\cdot\vec{r}'} V(\vec{r}') \phi_k^\pm(\vec{r}').$$

¹B. A. Lippmann e J. Schwinger, *Phys. Rev.* **79**, 469-480 (1950).

Isso não apenas mostra que as funções $\phi_k^\pm(\vec{r})$ satisfazem às condições de contorno de espalhamento mas ainda permite identificar as amplitudes de espalhamento $f_k^{(\pm)}(\hat{r})$ como sendo dadas por

$$f_k^{(\pm)}(\hat{r}) = -\frac{M}{2\pi\hbar^2} \int d^3r' e^{\mp i\vec{k}\cdot\vec{r}'} V(\vec{r}') \phi_k^\pm(\vec{r}') \equiv -\frac{M}{2\pi\hbar^2} \langle \pm \underline{k} \hat{r} | V | \phi_k^\pm \rangle. \quad (10.5)$$

Nesta última forma foi usada a notação $|\underline{k} \hat{r}\rangle$ para o autovetor do momento cuja função de onda é

$$\langle \vec{r}' | \underline{k} \hat{r} \rangle \equiv e^{i\vec{k}\cdot\vec{r}'},$$

cujas normalizações diferem por um fator da usual, $\langle \vec{r}' | \vec{k} \rangle \equiv e^{i\vec{k}\cdot\vec{r}'} / (2\pi)^{3/2}$, que corresponde a $\langle \vec{k}' | \vec{k} \rangle = \delta(\vec{k} - \vec{k}')$. É claro que a equação (10.5) tem a utilidade apenas formal (embora considerável) de exprimir as amplitudes de espalhamento diretamente em termos de elementos de matriz do potencial envolvendo os autoestados *exatos* $|\phi_k^\pm\rangle$. Como procedimento para o cálculo efetivo das amplitudes de espalhamento ela é no entanto de utilidade restrita, pois o problema da obtenção desses autoestados é equivalente, seja em termos de dificuldade, seja em termos dos procedimentos envolvidos, ao problema da obtenção das próprias amplitudes de espalhamento.

10.1.1 A expansão de Born.

O caráter *inhomogêneo* da equação de Lippmann-Schwinger permite obter expansões para as funções de onda $\phi_k^\pm(\vec{r})$ simplesmente iterando recursivamente a equação integral. Introduzindo, para simplificar a notação, a definição

$$U(\vec{r}) \equiv -\frac{M}{2\pi\hbar^2} V(\vec{r})$$

o resultado desse processo pode ser expresso como

$$\begin{aligned} \phi_k^\pm(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} &+ \int d^3r' G_k^\pm(\vec{r}, \vec{r}') U(\vec{r}') e^{i\vec{k}\cdot\vec{r}'} \\ &+ \int d^3r' G_k^\pm(\vec{r}, \vec{r}') U(\vec{r}') \int d^3r'' G_k^\pm(\vec{r}', \vec{r}'') U(\vec{r}'') e^{i\vec{k}\cdot\vec{r}''} + \dots \end{aligned} \quad (10.6)$$

ou ainda na forma mais compacta e um pouco mais abstrata

$$|\phi_k^\pm\rangle = |\underline{k}\rangle + G_k^\pm U |\underline{k}\rangle + G_k^\pm U G_k^\pm U |\underline{k}\rangle + \dots$$

Essa série é conhecida como *série de Born* para a função de onda $\phi_k^\pm(\vec{r})$ (ou para o vetor de estado $|\phi_k^\pm\rangle$). A substituição dessa série na expressão (10.5) para as amplitudes de espalhamento $f_k^{(\pm)}(\hat{r})$ dá a *expansão de Born* para essas amplitudes

$$\begin{aligned} f_k^{(\pm)}(\hat{r}) &= \int d^3r' e^{\mp i k \hat{r} \cdot \vec{r}'} U(\vec{r}') e^{i \vec{k} \cdot \vec{r}'} \\ &+ \int d^3r' e^{\mp i k \hat{r} \cdot \vec{r}'} U(\vec{r}') \int d^3r'' G_k^\pm(\vec{r}', \vec{r}'') U(\vec{r}'') e^{i \vec{k} \cdot \vec{r}''} + \dots \end{aligned}$$

que pode ser escrita também como

$$f_k^{(\pm)}(\hat{r}) = \langle \pm k \hat{r} | U | \vec{k} \rangle + \langle \pm k \hat{r} | U G_k^\pm U | \vec{k} \rangle + \langle \pm k \hat{r} | U G_k^\pm U G_k^\pm U | \vec{k} \rangle + \dots \quad (10.7)$$

Ela tem a forma de uma expansão perturbativa das amplitudes de espalhamento *exatas* em potências do potencial $U(\vec{r})$ (ou, equivalentemente, $V(\vec{r})$). É fácil verificar, em particular, que o truncamento dessa expansão no seu primeiro termo (o que constitui a chamada “primeira aproximação de Born” para a amplitude de espalhamento) reproduz o resultado perturbativo de primeira ordem obtido na seção 4.1.4. A expressão correspondente para a seção de choque é de fato

$$\frac{d\sigma}{d\Omega} = |f_k^{(+)}(\hat{r})|^2 \simeq |\langle k \hat{r} | U | \vec{k} \rangle|^2 = \frac{M^2}{4\pi^2 \hbar^4} \left| \int d^3r' e^{-i(k\hat{r} - \vec{k}) \cdot \vec{r}'} V(\vec{r}') \right|^2 = \frac{2\pi M^2}{\hbar^4} |\tilde{V}(k\hat{r} - \vec{k})|^2$$

com a definição utilizada anteriormente para a transformada de Fourier do potencial, $\tilde{V}(\vec{k})$. Vale a pena notar explicitamente que a primeira aproximação de Born para a amplitude de espalhamento

$$f_k^{+(\text{Born})} = -\frac{\sqrt{2\pi}M}{\hbar^2} \tilde{V}(k\hat{r} - \vec{k})$$

é dada essencialmente pela transformada de Fourier do potencial, e depende portanto dos ângulos de \hat{r} apenas através do momento transferido $\vec{q} \equiv k\hat{r} - \vec{k}$. Por outro lado, diferentes valores de $|\vec{k}|$ correspondem a diferentes ângulos de espalhamento para um dado momento transferido, de modo que, em situações nas quais a amplitude de espalhamento é fortemente dominada pelo primeiro termo da expansão de Born, é proveitoso e usual considerar a seção de choque diferencial de espalhamento não como função dos ângulos \hat{r} , mas como função do momento transferido \vec{q} .

A utilidade da expansão de Born para o cálculo de amplitudes de espalhamento e de seções de choque depende evidentemente das propriedades de convergência dessa expansão. A questão

geral da convergência da série de Born foi tratada por Jost e Pais² e posteriormente revisitada por Weinberg³. Condições gerais de convergência podem ser expressas em termos do espectro do núcleo $G_k^\pm V$ da equação de Lippmann-Schwinger (10.4), visto como um operador (não hermiteano) no espaço de Hilbert, mas seu uso é pouco freqüente.

As aplicações mais usuais se restringem quase sempre à primeira aproximação de Born apenas. Uma forma de avaliar a qualidade dessa aproximação consiste em usar a expressão formal (10.5) para a amplitude de espalhamento $f_k^{(+)}(\vec{r})$ e escrever a função de onda com condições de contorno de espalhamento que aí aparece como a soma de uma onda plana e de uma “onda espalhada”, isto é

$$\phi_k^\pm(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} + \psi_k^{\text{esp}}(\vec{r}).$$

Desse modo

$$f_k^{(+)}(\vec{r}) = -\frac{M}{2\pi\hbar^2} \left(\langle \underline{k}\hat{r} | V | \underline{k} \rangle + \langle \underline{k}\hat{r} | V | \psi_k^{\text{esp}} \rangle \right)$$

onde o primeiro termo corresponde à primeira aproximação de Born e o segundo termo deve ser relativamente pequeno para que essa aproximação seja boa. É claro, porém, que a onda espalhada representa a distorção da onda plana produzida pelo potencial $V(\vec{r})$, e que a sua contribuição para o segundo termo provém da região em que esse potencial é diferente de zero. Uma situação típica em que a primeira aproximação de Born *não* dá bons resultados é a de espalhamento ressonante, que implica em distorções consideráveis da função de onda na região sensível, envolvendo todos os termos da expansão de Born.

10.2 Descrição dependente do tempo de um processo de espalhamento por um potencial.

Um ingrediente essencial para uma descrição “realística” da evolução temporal de um processo de espalhamento é a localização das partículas envolvidas nele. De fato, o arranjo típico envolve um estágio inicial, em que as partículas são preparadas a uma distância suficientemente grande do potencial para que os efeitos deste sejam desprezíveis, com um momento inicial experimentalmente bem definido, escolhido de forma que se elas aproximem do potencial espalhador; e ainda, após a evolução temporal durante a qual se dá a interação com o potencial, um estágio final, em que o momento final da partícula espalhada é medido também a uma distância em que

²R. Jost e A Pais, *Phys. Rev.* **82**, 840 (1951). V. também W. Kohn, *Rev. Mod. Phys.* **26**, 292 (1954).

³S. Weinberg, *Phys. Rev.* **130**, 776 (1963) e *ibid.* **131**, 440 (1963), especialmente a última destas duas referências.

os efeitos do potencial tenham se tornado novamente desprezíveis. A teoria quântica impõe, no entanto, limitações intrínsecas para a localização e para a determinação simultânea do momento de partículas, expressas em termos das relações de incerteza posição-momento. Por isso é importante examinar inicialmente as diversas ordens de grandeza envolvidas numa situação experimental típica com o objetivo de, por um lado, avaliar a importância dessas limitações e, por outro lado, identificar os parâmetros relevantes para o resultado da experiência.

10.2.1 Escalas típicas em processos de espalhamento.

Um estado quântico típico de uma partícula com posição e momento experimentalmente definidos pode ser representado por meio de um pacote de ondas do tipo

$$\psi_0(\vec{r} - \vec{r}_0) = \frac{1}{(2\pi)^{3/2}} \int d^3k \tilde{u}(\vec{k} - \vec{k}_0) e^{i\vec{k} \cdot (\vec{r} - \vec{r}_0)} = \frac{e^{i\vec{k}_0 \cdot (\vec{r} - \vec{r}_0)}}{(2\pi)^{3/2}} \int d^3k' \tilde{u}(\vec{k}') e^{i\vec{k}' \cdot (\vec{r} - \vec{r}_0)} \quad (10.8)$$

onde a amplitude $\tilde{u}(\vec{k} - \vec{k}_0)$ é uma função normalizada do momento \vec{k} que difere apreciavelmente de zero apenas numa vizinhança de \vec{k}_0 . Na segunda forma dada à integral foi introduzido como variável de integração o vetor $\vec{k}' \equiv \vec{k} - \vec{k}_0$, que é efetivamente limitado pelo alcance da função $\tilde{u}(\vec{k} - \vec{k}_0) = \tilde{u}(\vec{k}')$. Exemplo de um tal pacote de ondas é o pacote gaussiano tratado na seção 3.3.3, para o qual essa amplitude é dada por

$$\tilde{u}(\vec{k} - \vec{k}_0) = \left[\frac{2b}{\sqrt{\pi}} \right]^{\frac{3}{2}} e^{-2b^2(\vec{k} - \vec{k}_0)^2}, \quad \int d^3k |\tilde{u}(\vec{k} - \vec{k}_0)|^2 = 1.$$

Com essa particular escolha, a equação (10.8) dá

$$\psi_0(\vec{r} - \vec{r}_0) = \left[\frac{1}{\sqrt{\pi}b} \right]^{\frac{3}{2}} e^{i\vec{k}_0 \cdot (\vec{r} - \vec{r}_0)} e^{-\frac{(\vec{r} - \vec{r}_0)^2}{2b^2}} \quad (10.9)$$

que corresponde a um pacote de ondas de incerteza mínima, de largura b em torno do ponto \vec{r}_0 , cujo momento médio é $\hbar\vec{k}_0$. Para a discussão que segue não será necessário, porém, usar uma forma explícita para $\tilde{u}(\vec{k} - \vec{k}_0)$, bastando notar que a localização mencionada dessa amplitude no espaço de momentos implica numa (des)localização complementar de $\psi_0(\vec{r} - \vec{r}_0)$ no espaço de posições em torno de \vec{r}_0 , e que as limitações impostas pelas relações de incerteza são tipicamente muito menores que as decorrentes de limitações experimentais. Em particular, não será necessário supor que o pacote tenha incerteza mínima.

A evolução temporal de um pacote de ondas livre, representado inicialmente pela equação (10.8), é dada por

$$\begin{aligned}
\psi(\vec{r}, t) &= e^{-\frac{i}{\hbar} H_0 t} \psi_0(\vec{r} - \vec{r}_0) = \frac{1}{(2\pi)^{3/2}} \int d^3 k \tilde{u}(\vec{k} - \vec{k}_0) e^{-i \frac{\hbar k^2}{2M} t} e^{i \vec{k} \cdot (\vec{r} - \vec{r}_0)} = \\
&= \frac{1}{(2\pi)^{3/2}} \int d^3 k' \tilde{u}(\vec{k}') e^{-i \frac{\hbar}{2M} (\vec{k}' + \vec{k}_0)^2 t} e^{i (\vec{k}' + \vec{k}_0) \cdot (\vec{r} - \vec{r}_0)}
\end{aligned}$$

onde $H_0 = p^2/2M$ e, na última expressão, a variável de integração foi tomada como sendo $\vec{k}' \equiv \vec{k} - \vec{k}_0$. Os argumentos das exponenciais que aí aparecem podem ser reagrupados como

$$-\frac{\hbar}{2M} (\vec{k}' + \vec{k}_0)^2 t + (\vec{k}' + \vec{k}_0) \cdot (\vec{r} - \vec{r}_0) = \vec{k}' \cdot (\vec{r} - \vec{r}_0 - \frac{\hbar \vec{k}_0}{M} t) - \frac{\hbar k_0^2}{2M} t - \frac{\hbar k'^2}{2M} t + \vec{k}_0 \cdot (\vec{r} - \vec{r}_0)$$

onde o primeiro termo corresponde ao deslocamento do pacote com velocidade uniforme $\hbar \vec{k}_0/M$, o segundo dá origem a um fator de fase global e o terceiro é responsável pela dispersão do pacote. O último termo é, como o que corresponde ao fator de fase global, independente da variável de integração, e dá origem ao fator de onda plana a ser modulada pela amplitude do pacote no espaço de posições, exatamente como no caso do pacote gaussiano de incerteza mínima (v. equação (10.9)). É fácil ver da expressão para o terceiro termo que os efeitos de dispersão serão na realidade desprezíveis quando

$$\frac{\hbar k'^2}{2M} t \ll 1 \quad \text{ou} \quad t \ll \frac{2M}{\hbar k'^2}$$

para todos os valores de k' para os quais a amplitude $\tilde{u}(\vec{k}')$ seja apreciavelmente diferente de zero. Usando a dispersão em momento σ_k (que vale $1/\sqrt{2}b$ no caso do pacote gaussiano) para caracterizar os maiores valores de k' que satisfazem essa condição, a condição para que a dispersão do pacote seja desprezível pode ser escrita como

$$t \ll \tau_k \equiv \frac{2M}{\hbar \sigma_k^2}$$

onde o tempo τ_k pode ser visto como um tempo característico para o processo dispersivo. Vale a pena observar explicitamente que esse tempo não depende do momento médio \vec{k}_0 do pacote (que define a velocidade do deslocamento), mas apenas de sua largura no espaço de momentos.

Num processo típico de colisão, o pacote de ondas incidente é preparado a uma distância L do alvo, que deve ser percorrida antes que a colisão ocorra. O tempo τ_L envolvido nesse deslocamento é então dado por

$$\tau_L \simeq \frac{ML}{\hbar k_0}$$

cuja comparação com o tempo característico de dispersão dá

$$\frac{\tau_k}{\tau_L} \sim \frac{2M}{\hbar\sigma_k^2} \frac{\hbar k_0}{ML} = \frac{2k_0}{\sigma_k^2 L}.$$

Alguns exemplos específicos são úteis para avaliar quantitativamente o conteúdo dessa relação. No caso de um próton preparado com energia cinética de 1 MeV, o momento médio k_0 é dado por

$$k_0^2 = \frac{2m_p E}{\hbar^2} = 0.0478 \times 10^{26} \text{ cm}^{-2}$$

ou seja, $k_0 \simeq 0.22 \times 10^{13} \text{ cm}^{-1}$. Tomando $\sigma_k \sim 100 \text{ cm}^{-1}$, o que corresponde a uma localização espacial na escala de 0.01 cm para um pacote mínimo, e $L \sim 100 \text{ cm}$, o que se obtém é

$$\frac{\tau_k}{\tau_L} \sim 10^8.$$

Para um elétron com energia cinética de 1 eV, por outro lado,

$$k_0^2 = \frac{2m_e E}{\hbar^2} = 0.260 \times 10^{16} \text{ cm}^{-2}$$

ou seja $k_0 \simeq 0.51 \times 10^8 \text{ cm}^{-1}$ que, para os mesmos valores de σ_k e de L dá

$$\frac{\tau_k}{\tau_L} \sim 10^2.$$

Isso mostra que, mesmo na situação desfavorável de projéteis leves e energia cinética pequena, o “tempo de trânsito” τ_L é tipicamente muito menor que o tempo característico de dispersão τ_k , de modo que os efeitos de dispersão podem ser desprezados durante a propagação livre do pacote de ondas.

O ponto seguinte a ser considerado é o de que uma experiência de espalhamento típica envolve um feixe contendo um grande número de partículas incidentes, o qual é dirigido para um alvo que contém, por sua vez, um grande número de centros espalhadores. Parâmetros normalmente utilizados para caracterizar as propriedades de um feixe externo de partículas carregadas produzido por um acelerador incluem a intensidade, usualmente dada em termos de uma corrente elétrica, e a resolução em energia, usualmente expressa através de uma porcentagem da energia média.

Quanto à multiplicidade de pacotes e a dispersão instrumental de energia no feixe, como a localização das partículas implica por sua vez, através da cinemática quântica, numa dispersão não nula no momento e portanto também na energia, é interessante comparar resoluções em

energia típicas dos feixes utilizados com esta dispersão “intrínseca”, que pode ser estimada como

$$\Delta E_k \sim \frac{\hbar^2 \sigma_k^2}{2M} = E \left(\frac{\sigma_k}{k} \right)^2$$

onde E é a energia média. Para os exemplos considerados acima, o último fator é respectivamente da ordem de 10^{-20} e de 10^{-11} , o que mostra que a indefinição instrumental da energia do feixe é muitas ordens de grandeza maior que a dispersão intrínseca associada à localização dos pacotes de onda. Nessas condições, diferentes pacotes presentes no feixe tendem a estar associados a valores da energia que são suficientemente diferentes para serem em princípio distinguíveis, tornando os efeitos de cada um dos pacotes também distinguível dos efeitos dos demais. Os efeitos médios resultantes das colisões produzidas por uma tal coleção de pacotes podem portanto ser obtidos combinando *incoerentemente* (isto é, ao nível das probabilidades, e não das amplitudes quânticas) os efeitos dos pacotes considerados individualmente. Um feixe com essas características é por isso dito *incoerente*, e esta propriedade é a que permite, em última análise, reduzir o problema do espalhamento envolvendo o feixe ao problema do espalhamento de um único pacote de ondas. A multiplicidade de centros espalhadores no alvo, por outro lado, pode ser caracterizada por meio de uma distância média entre eles definida como $d \sim \rho^{-1/3}$, onde ρ é a densidade de centros espalhadores. A propriedade das partículas incidentes que é relevante neste contexto é o comprimento de onda médio de de Broglie

$$\frac{\lambda_0}{2\pi} \equiv \frac{1}{k_0} = \sqrt{\frac{\hbar^2}{2ME}}.$$

Considerando mais uma vez os dois exemplos já utilizados, os valores de λ_0 são respectivamente da ordem de 10^{-13} cm e de 10^{-8} cm. Para um alvo com a densidade típica de um sólido ou líquido normal $\rho \sim 10^{23}$ cm $^{-3}$ e portanto $d \sim 10^{-8}$ cm o que é, em particular, muito menor que as dimensões admitidas para o pacote de ondas. No caso em que $d/\lambda_0 \gg 1$, no entanto, as ondas espalhadas provenientes de centros diferentes chegarão ao detector com diferenças de fase muito grandes de modo que efeitos de interferência tendam em média a se anular. O resultado disso é que a intensidade observada será dada essencialmente pela soma das intensidades devidas a cada um dos centros espalhadores considerado individualmente, o que permite reduzir o problema de espalhamento por muitos centros também a um problema de espalhamento por um único alvo. A situação é no entanto claramente outra no caso do feixe de electrons, e nesse caso de fato há efeitos de interferência suficientemente importantes para que sejam usados para estudar propriedades da distribuição espacial de centros espalhadores no alvo (difração de elétrons).

O problema que será considerado no restante desta seção é o do espalhamento incoerente de um feixe também incoerente. Ele abrange, pelo menos, um grande número de situações de

interesse na física atômica e nuclear pode, como visto, ser reduzido ao problema do espalhamento de um único pacote de ondas por um único centro espalhador.

10.2.2 Tratamento dependente do tempo do espalhamento de um pacote de ondas por um potencial de curto alcance.

A estratégia geral a ser seguida para estudar a evolução temporal de um pacote de ondas incidente sobre um centro espalhador representado por um potencial de curto alcance $V(\vec{r})$ é na realidade a mesma que foi utilizada no Capítulo 3 para estudar de uma forma geral a evolução temporal de sistemas quânticos. O primeiro passo consiste em expandir o estado inicial do pacote em estados estacionários do Hamiltoniano, isto é, determinar os coeficientes c_n em

$$\psi(\vec{r}, t = 0) = \sum_n c_n \phi_n(\vec{r})$$

onde as funções de onda $\phi_n(\vec{r})$ satisfazem

$$\left[\frac{p^2}{2M} + V(\vec{r}) \right] \phi_n(\vec{r}) = E_n \phi_n(\vec{r}) \quad (10.10)$$

o que determina o estado quântico no tempo t simplesmente através de

$$\psi(\vec{r}, t) = \sum_n c_n e^{-\frac{iE_n t}{\hbar}} \phi_n(\vec{r}).$$

No caso do problema de espalhamento há complicações ligadas ao espectro contínuo, o que implica em substituir o índice que aparece nas somas por índices contínuos e as somas por integrais. As autofunções envolvidas na expansão (e conseqüentemente também os coeficientes da expansão) dependem ainda, nesse caso, da escolha de condições de contorno que definam completamente as autofunções a serem utilizadas, de modo que o problema da determinação dos coeficientes da expansão deve ser tratado conjuntamente com o da escolha mais conveniente de condições de contorno para as autofunções do espectro contínuo.

A solução desse problema duplo, de escolha de condições de contorno e determinação de coeficientes de expansão, pode porém ser obtida de forma simples utilizando um resultado que pode ser expresso da seguinte forma. Se $\phi_k^\pm(\vec{r})$ são as soluções da equação de Lippmann-Schwinger respectivamente com onda espalhada emergente e imergente, e se $\tilde{u}(\vec{k} - \vec{k}_0)$ é a transformada de Fourier do pacote de ondas tal qual foi utilizada na equação (10.8), então

a) se $\vec{k}_0 \cdot \vec{r}_0 < 0$, no limite assintótico $|\vec{r}_0| \rightarrow \infty$ vale a relação

$$\psi_0(\vec{r} - \vec{r}_0) \simeq \chi^+(\vec{r}) \equiv \frac{1}{(2\pi)^{3/2}} \int d^3k e^{-i\vec{k} \cdot \vec{r}_0} \phi_k^+(\vec{r}) \tilde{u}(\vec{k} - \vec{k}_0). \quad (10.11)$$

A condição $\vec{k}_0 \cdot \vec{r}_0 < 0$ é satisfeita, em particular, para pacotes de onda cujo momento médio tem sentido oposto ao seu vetor médio de posição, ou seja, para pacotes de onda que *se aproximam* da origem, e se aplica portanto aos pacotes de onda incidentes (ou pacotes “pré-colisão”) do problema de espalhamento; e ainda

b) se $\vec{k}_0 \cdot \vec{r}_0 > 0$, no limite assintótico $|\vec{r}_0| \rightarrow \infty$ vale a relação

$$\psi_0(\vec{r} - \vec{r}_0) \simeq \chi^-(\vec{r}) \equiv \frac{1}{(2\pi)^{3/2}} \int d^3k e^{-i\vec{k} \cdot \vec{r}_0} \phi_k^-(\vec{r}) \tilde{u}(\vec{k} - \vec{k}_0), \quad (10.12)$$

onde se pode notar que condição $\vec{k}_0 \cdot \vec{r}_0 > 0$ se aplica agora a pacotes que *se afastam* da origem (ou pacotes “pós-colisão”). Como os pacotes de onda iniciais no processo de espalhamento são pacotes pré-colisão preparados na região assintótica $|\vec{r}| \rightarrow \infty$, o resultado **a)** resolve o problema de sua expansão em autofunções do Hamiltoniano, representadas nesse caso pelas soluções da equação e Lippmann-Schwinger com ondas *emergentes* $\phi_k^+(\vec{r})$, de modo que a sua evolução temporal aparece sob a forma

$$\psi(\vec{r}, t) = e^{-\frac{i}{\hbar} H t} \psi_0(\vec{r} - \vec{r}_0) \simeq \frac{1}{(2\pi)^{3/2}} \int d^3k e^{-i\vec{k} \cdot \vec{r}_0} e^{-i\frac{\hbar k^2 t}{2M}} \phi_k^+(\vec{r}) \tilde{u}(\vec{k} - \vec{k}_0). \quad (10.13)$$

A aproximação envolvida nessa relação provém daquela existente na expressão (10.11), e será explicitada a seguir através de sua demonstração.

A verificação dos resultados (10.11) e (10.12) pode ser feita calculando as funções $\chi^\pm(\vec{r})$ da seguinte forma. Usando a equação de Lippmann-Schwinger (10.4) para as autofunções $\phi_k^\pm(\vec{r})$ que aparecem no integrando essas funções podem ser reescritas como

$$\begin{aligned} \chi^\pm(\vec{r}) &= \frac{1}{(2\pi)^{3/2}} \int d^3k e^{-i\vec{k} \cdot \vec{r}_0} \left[e^{i\vec{k} \cdot \vec{r}} - \frac{M}{2\pi\hbar^2} \int d^3r' \frac{e^{\pm i k |\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} V(\vec{r}') \phi_k^\pm(\vec{r}') \right] \tilde{u}(\vec{k} - \vec{k}_0) \\ &= \psi_0(\vec{r} - \vec{r}_0) - \\ &\quad - \frac{1}{(2\pi)^{3/2}} \int d^3k \frac{M}{2\pi\hbar^2} \int d^3r' e^{-i\vec{k} \cdot \vec{r}_0} \frac{e^{\pm i k |\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} V(\vec{r}') \phi_k^\pm(\vec{r}') \tilde{u}(\vec{k} - \vec{k}_0). \end{aligned}$$

O termo inhomogêneo da equação de Lippmann-Schwinger simplesmente reconstrói o pacote inicial, enquanto o termo que contém a função de Green envolve integrações em \vec{k} e \vec{r}' . Este último termo pode ser simplificado sempre que a dependência com \vec{k} de $\phi_k^\pm(\vec{r}')$, em pontos \vec{r}' onde $V(\vec{r}')$ seja significativamente diferente de zero, possa ser desprezada

para o intervalo de valores de \vec{k} em que $\tilde{u}(\vec{k} - \vec{k}_0)$ seja, por sua vez, significativamente diferente de zero. Nessas condições é possível utilizar a aproximação

$$\phi_{\vec{k}}^{\pm}(\vec{r}') \simeq \phi_{\vec{k}_0}^{\pm}(\vec{r}')$$

com a qual esse objeto pode ser excluído da integração sobre \vec{k} , dando então

$$\chi^{\pm}(\vec{r}) \simeq \psi_0(\vec{r} - \vec{r}_0) - \frac{M}{2\pi\hbar^2} \int d^3r' V(\vec{r}') \phi_{\vec{k}_0}^{\pm}(\vec{r}') I^{\pm}(\vec{r}, \vec{r}')$$

onde

$$I^{\pm}(\vec{r}, \vec{r}') = \frac{1}{(2\pi)^{3/2}} \int d^3k e^{-i\vec{k} \cdot \vec{r}_0} \frac{e^{\pm i\vec{k} \cdot |\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} \tilde{u}(\vec{k} - \vec{k}_0).$$

Como no tratamento do pacote livre, essas integrais podem ser re-escritas usando $\vec{k}' \equiv \vec{k} - \vec{k}_0$ como variável de integração e usando o curto alcance da função $\tilde{u}(\vec{k}')$ para introduzir a expansão

$$k = |\vec{k}' + \vec{k}_0| \simeq k_0 + \vec{k}' \cdot \hat{k}_0$$

onde $\hat{k}_0 \equiv \vec{k}_0/|\vec{k}_0|$. Isso permite finalmente escreve-las em termos do pacote inicial como

$$\begin{aligned} I^{\pm}(\vec{r}, \vec{r}') &\simeq \frac{1}{|\vec{r} - \vec{r}'|} \frac{e^{i\vec{k}_0 \cdot (\pm \hat{k}_0 |\vec{r} - \vec{r}'| - \vec{r}_0)}}{(2\pi)^{3/2}} \int d^3k' e^{i\vec{k}' \cdot (\pm \hat{k}_0 |\vec{r} - \vec{r}'| - \vec{r}_0)} \tilde{u}(\vec{k}') = \\ &= \frac{1}{|\vec{r} - \vec{r}'|} \psi_0(\pm \hat{k}_0 |\vec{r} - \vec{r}'| - \vec{r}_0). \end{aligned}$$

A função $\psi_0(\vec{r})$ é no entanto, por construção, significativamente diferente de zero apenas numa região de dimensões da ordem de b em torno de $\vec{r} = 0$. Para verificar a relação (10.11) basta então notar que, se $\vec{k}_0 \cdot \vec{r}_0 < 0$ e $\vec{r}_0 \rightarrow \infty$ o argumento de ψ_0 na última expressão será sempre diferente de zero e muito grande *no caso de* $I^+(\vec{r}, \vec{r}')$, devido ao sinal, nesse caso positivo, do termo do argumento que envolve \hat{k}_0 . No caso em que $\vec{k}_0 \cdot \vec{r}_0 > 0$, por outro lado, o argumento de ψ_0 será sempre não nulo e suficientemente grande no caso de $I^-(\vec{r}, \vec{r}')$, o que leva à relação (10.12).

As relações (10.11) e (10.12) podem ser interpretadas dizendo que estados de espalhamento com condições de contorno de ondas *emergentes* são adequados para a expansão de pacotes pré-colisão, ao passo que estados com ondas *imergentes* são adequados para a expansão de pacotes pós-colisão. Essa diferença pode ser entendida notando que, em ambas as situações, o pacote é reconstituído *apenas* pela onda plana da solução de espalhamento estacionária, e que portanto as contribuições das ondas espalhadas se anulam por

interferência destrutiva na integração sobre \vec{k} . Admitindo como intuitivamente razoável que a interferência construtiva de ondas espalhadas emergentes só possa ocorrer *depois* da colisão, e que a de ondas imergentes só possa ocorrer *antes* da colisão, segue então que o uso de condições de contorno “não naturais” (ondas emergentes e imergentes, respectivamente para pacotes pré- e pós-colisão) permite efetivamente ignorar a contribuição de ondas espalhadas no processo de reconstrução do pacote através de suas componentes de Fourier.

O passo seguinte consiste em calcular a expressão (10.13) com t suficientemente grande para que os produtos da colisão atinjam o sistema de detecção, localizado a uma distância do alvo que é suficientemente grande para que a função $\phi_k^+(\vec{r})$ possa ser substituída pela sua forma assintótica

$$\phi_k^+(\vec{r}) \longrightarrow e^{i\vec{k}\cdot\vec{r}} + f_k^{(+)}(\hat{r}) \frac{e^{ikr}}{r}.$$

O termo de (10.13) contendo a onda plana corresponde simplesmente à propagação livre do pacote de ondas inicial. Na aproximação em que o alargamento do pacote durante a propagação livre é desprezado, essa relação pode ser expressa como

$$\psi(\vec{r}, t) \simeq e^{-i\frac{\hbar k_0^2 t}{2M}} \psi_0(\vec{r} - \vec{v}_0 t - \vec{r}_0) + \psi_{\text{esp}}(\vec{r}, t),$$

onde $\vec{v}_0 \equiv \hbar \vec{k}_0 / M$, com o pacote espalhado $\psi_{\text{esp}}(\vec{r}, t)$ dado por

$$\begin{aligned} \psi_{\text{esp}}(\vec{r}, t) &= \frac{1}{(2\pi)^{3/2}} \int d^3k \, e^{-i\vec{k}\cdot\vec{r}_0} e^{-i\frac{\hbar k^2 t}{2M}} f_k^{(+)}(\hat{r}) \frac{e^{ikr}}{r} \tilde{u}(\vec{k} - \vec{k}_0) \simeq \\ &\simeq f_{k_0}^{(+)}(\hat{r}) \frac{1}{(2\pi)^{3/2}} \int d^3k \, e^{-i\vec{k}\cdot\vec{r}_0} e^{-i\frac{\hbar k^2 t}{2M}} \frac{e^{ikr}}{r} \tilde{u}(\vec{k} - \vec{k}_0). \end{aligned}$$

Na última passagem, a substituição da amplitude de espalhamento para momento de módulo k pela quantidade correspondente para o momento médio, de módulo k_0 , é justificável sempre que o intervalo de momentos em que $\tilde{u}(\vec{k} - \vec{k}_0)$ é significativamente diferente de zero for suficientemente pequeno para que a variação da amplitude de espalhamento com k possa aí ser desprezada. Nessas condições, a integral resultante sobre \vec{k} pode também ser expressa em termos do pacote de ondas inicial, substituindo mais uma vez a variável de integração por $\vec{k}' \equiv \vec{k} - \vec{k}_0$ e usando a expansão

$$k \simeq k_0 + \vec{k}' \cdot \hat{k}_0.$$

Desse modo se obtém, após algumas transformações algébricas simples,

$$\begin{aligned}\psi_{\text{esp}}(\vec{r}, t) &\simeq e^{-i\frac{\hbar k_0^2 t}{2M}} \frac{f_k^{(+)}(\hat{r})}{r} \frac{e^{i\vec{k}_0 \cdot (\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)}}{(2\pi)^{3/2}} \int d^3 k' e^{i\vec{k}' \cdot (\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)} \tilde{u}(\vec{k}') = \\ &= e^{-i\frac{\hbar k_0^2 t}{2M}} \frac{f_{k_0}^{(+)}(\hat{r})}{r} \psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0).\end{aligned}$$

O resultado final para o pacote de ondas, num tempo t suficientemente grande para que a região de interferência construtiva esteja já suficientemente afastada do alvo para que as expressões assintóticas das autofunções do hamiltoniano sejam válidas, é dado portanto pela expressão

$$\psi(\vec{r}, t) \simeq e^{-i\frac{\hbar k_0^2 t}{2M}} \left[\psi_0(\vec{r} - \vec{v}_0 t - \vec{r}_0) + \frac{f_{k_0}^{(+)}(\hat{r})}{r} \psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \right]. \quad (10.14)$$

Os dois termos que aí aparecem podem ser facilmente interpretados da seguinte forma (v. fig. 10.1). A menos de uma fase global irrelevante, o primeiro termo representa o pacote incidente propagado livremente sem distorção (devido às aproximações feitas nesse sentido) e portanto deslocado de uma distância $\vec{v}_0 t$ de sua posição inicial. Esse termo corresponde portanto a uma amplitude de probabilidade que difere de zero apenas numa região de dimensões da ordem de b na direção frontal definida por \vec{k}_0 . O segundo termo, por outro lado, representa uma amplitude espalhada cuja dependência com a coordenada *radial* r é dada em termos do perfil do pacote inicial ao longo da direção de incidência. De fato, a amplitude em \vec{r}

$$\psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)$$

é dada (a menos do fator geométrico $1/r$) pela amplitude do pacote *inicial* calculada no ponto $\hat{k}_0 r - \vec{v}_0 t$. A dependência *angular* da amplitude espalhada é dada porém pela amplitude de espalhamento $f_{k_0}^{(+)}(\hat{r})$. Desse modo, os dois termos de (10.14) são *simultaneamente* diferentes de zero apenas numa região de dimensões da ordem b em torno da direção frontal. Para direções diferentes dessa a amplitude correspondente ao pacote transmitido se anula, e a única contribuição para a amplitude total é a que provém de ψ_{esp} . Na direção frontal, porém, a amplitude total aparece como a soma de duas contribuições não nulas, devidas respectivamente ao pacote transmitido e ao pacote espalhado, sendo que esta última tem como fator a amplitude de espalhamento na direção frontal $f_{k_0}^{(+)}(\hat{k}_0)$.

A interferência, na direção frontal, entre a amplitude transmitida e a amplitude espalhada tem o papel importante de garantir a preservação da norma da função de onda ao longo do tempo, o que é uma consequência do caráter hermiteano do Hamiltoniano que gera a evolução temporal. Isso indica ainda que a amplitude de espalhamento na direção frontal deve conter

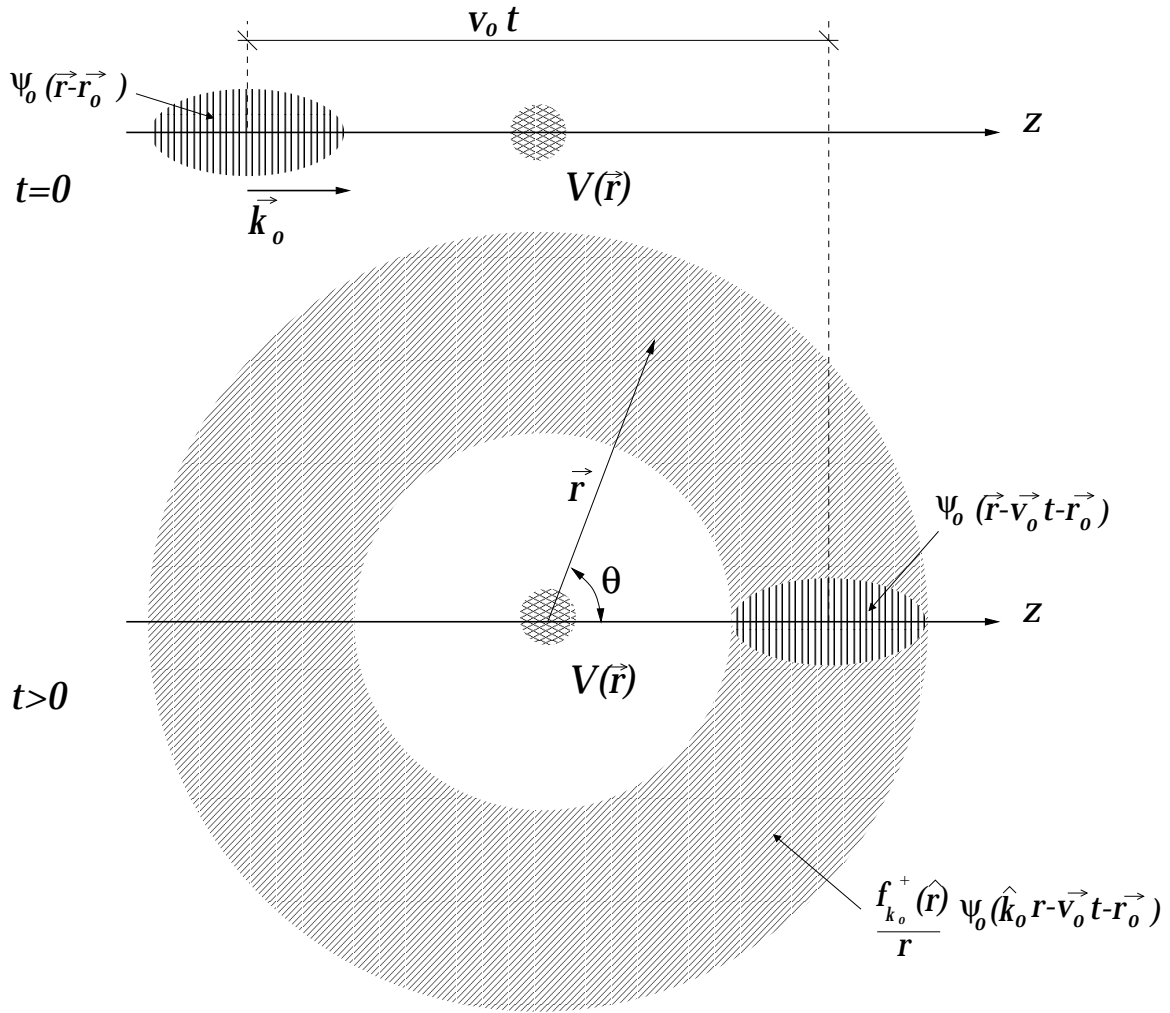


Figura 10.1: Parte superior ($t = 0$): pacote incidente localizado em torno de \vec{r}_0 . Parte inferior ($t > 0$, suficientemente grande): representação esquemática do pacote de ondas espalhado e do pacote de ondas transmitido, com interferência na direção frontal.

informação sobre a probabilidade espalhada *total*, dado que esta é a quantidade envolvida na conservação da norma da função de onda. De fato, como será mostrado na seção seguinte, a condição de conservação de norma, aplicada à expressão (10.14), conduz diretamente ao teorema óptico, que relaciona a seção de choque total à amplitude de espalhamento na direção frontal (v. seção 5.1.5).

Resta, finalmente, relacionar o resultado (10.14) com a seção de choque de espalhamento. Em termos experimentais, o elemento de seção de choque diferencial $d\sigma$ para espalhamento incoerente de um feixe incoerente, correspondente ao elemento de ângulo sólido $d\Omega$ na direção de observação \hat{r} , é definido como

$$d\sigma = \frac{\text{número de pacotes espalhados medidos em } d\Omega}{(\text{número de pacotes incidentes}) \times (\text{número de alvos por unidade de área})}.$$

O segundo fator do denominador corresponde ao número de alvos por unidade de área, numa projeção sobre um plano perpendicular à direção de incidência⁴. Dado o caráter incoerente do feixe e dos processos de espalhamento envolvendo diferentes alvos, $d\sigma$ pode ser expresso em termos de quantidades calculadas para um único pacote e de um único alvo como

$$d\sigma = \frac{\text{probabilidade em } d\Omega \text{ para um pacote}}{\text{probabilidade incidente por unidade de área}} \equiv \frac{N}{D}.$$

Tanto o numerador N quanto o denominador D dessa relação podem ser calculados a partir de (10.14). O numerador pode, de fato, ser calculado como

$$N = \int_{-\infty}^{+\infty} dt \vec{j}(\vec{r}, t) \cdot \hat{r} r^2 d\Omega$$

com a densidade de corrente de probabilidade

$$\vec{j}(\vec{r}, t) = \frac{\hbar}{2Mi} [\psi^*(\vec{r}, t) \nabla \psi(\vec{r}, t) - (\nabla \psi(\vec{r}, t))^* \psi(\vec{r}, t)].$$

Para calcular essa quantidade é conveniente escrever o pacote de ondas inicial explicitamente como

$$\psi_0(\vec{r} - \vec{r}_0) = e^{i\vec{k}_0 \cdot \vec{r}} u(\vec{r} - \vec{r}_0) \quad (10.15)$$

onde o último fator corresponde à distribuição *espacial* da amplitude do pacote, relacionada com $\tilde{u}(\vec{k} - \vec{k}_0)$ por uma transformação de Fourier. Desse modo a função de onda $\psi(\vec{r}, t)$ na região assintótica $r \rightarrow \infty$ é dada por

⁴O hábito de especificar a “espessura do alvo” em unidades de gramas por centímetro quadrado está relacionado com este fator.

$$\psi(\vec{r}, t) \longrightarrow e^{-i\frac{\hbar k_0^2 t}{2M}} \left[e^{i\vec{k}_0 \cdot \vec{r}} u(\vec{r} - \vec{v}_0 t - \vec{r}_0) + f_{k_0}^{(+)}(\hat{r}) \frac{e^{ik_0 r}}{r} u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \right].$$

Em qualquer direção diferente da direção frontal $\hat{r} = \hat{k}_0$ apenas o segundo termo é diferente de zero, e a componente radial do gradiente que aparece na expressão da corrente é

$$\hat{r} \cdot \nabla \psi(\vec{r}, t) \rightarrow e^{-i\frac{\hbar k_0^2 t}{2M}} f_{k_0}^{(+)}(\hat{r}) \frac{e^{ik_0 r}}{r} u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \times \left[ik_0 - \frac{1}{r} + \frac{1}{u} \frac{d}{dr} u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \right].$$

O segundo termo, $1/r$, não só é muito menor que o primeiro na região assintótica mas na realidade não contribui para a corrente por ser real. Se a amplitude u é real o terceiro termo tão pouco contribui para a corrente, mas mesmo no caso em que essa amplitude seja complexa a sua contribuição deve ser da ordem da incerteza $\sigma_k \ll k_0$ introduzida no momento do pacote pela sua localização espacial, e portanto muito menor que a do primeiro termo. Nessas condições apenas a contribuição desse primeiro termo é relevante, e é possível escrever

$$\hat{r} \cdot \vec{j}(\vec{r}, t) \rightarrow \frac{\hbar k_0}{M} \frac{|f_{k_0}^{(+)}(\hat{r})|^2}{r^2} |u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2,$$

o que leva à expressão para o numerador da seção de choque

$$N = \frac{\hbar k_0}{M} |f_{k_0}^{(+)}(\hat{r})|^2 d\Omega \int_{-\infty}^{+\infty} dt |u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2.$$

O cálculo do denominador D pode ser feito de uma forma semelhante ao do numerador. A probabilidade incidente por unidade de área é dada por

$$D = \int_{-\infty}^{t_{\max}} dt \hat{k}_0 \cdot \vec{j}(\vec{r}, t)$$

onde a densidade de corrente é agora calculada na região pré-colisão, em que $\vec{k}_0 \cdot \vec{r} < 0$, durante o tempo em que apenas o primeiro termo de $\psi(\vec{r}, t)$ é diferente de zero. A limitação da integral temporal a t_{\max} é introduzida para evitar a inclusão em D da contribuição “pós-colisão” do pacote espalhado para ângulos traseiros. A componente relevante do gradiente que aparece na expressão da corrente é portanto

$$\hat{k}_0 \cdot \nabla \psi(\vec{r}, t) \rightarrow e^{-i\frac{\hbar k_0^2 t}{2M}} e^{i\vec{k}_0 \cdot \vec{r}} u(\vec{r} - \vec{v}_0 t - \vec{r}_0) \left[ik_0 + \frac{1}{u} \hat{k}_0 \cdot \nabla u(\vec{r} - \vec{v}_0 t - \vec{r}_0) \right]$$

onde mais uma vez o termo envolvendo o gradiente de u é de ordem σ_k e portanto desprezível com relação ao primeiro termo. Isso leva à expressão para o denominador

$$D = \frac{\hbar k_0}{M} \int_{-\infty}^{t_{\max}} dt |u(\vec{r} - \vec{v}_0 t - \vec{r}_0)|^2.$$

Combinando os resultados obtidos para N e D resulta portanto

$$d\sigma = \frac{\frac{\hbar k_0}{M} |f_{k_0}^{(+)}(\hat{r})|^2 d\Omega \int_{-\infty}^{+\infty} dt |u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2}{\frac{\hbar k_0}{M} \int_{-\infty}^{t_{\max}} dt |u(\vec{r} - \vec{v}_0 t - \vec{r}_0)|^2} = |f_{k_0}^{(+)}(\hat{r})|^2 d\Omega$$

dado que as duas integrais temporais abrangem toda a extensão do pacote de ondas e são portanto iguais. O que se obtém para a seção de choque diferencial é portanto

$$\frac{d\sigma}{d\Omega} = |f_{k_0}^{(+)}(\hat{r})|^2$$

que reproduz o resultado (5.27), obtido anteriormente. Em termos práticos, portanto, o problema da obtenção de seções de choque para processos de espalhamento incoerente de feixes incoerentes se reduz ao problema do cálculo da amplitude de espalhamento $f_{k_0}^{(+)}(\hat{r})$.

10.2.3 Conservação de norma e o teorema óptico.

A evolução temporal do pacote de ondas $\psi_0(\vec{r} - \vec{r}_0)$ definido pela equação (10.8) é determinada pelo operador unitário de evolução $U(t) = e^{-\frac{i}{\hbar} H t}$, e esse fato por si só garante a preservação da norma do pacote ao longo do tempo. Por outro lado, para valores suficientemente grandes de t , foi possível escrever o pacote de ondas temporalmente evoluído sob a forma (10.14), que consiste da soma de um termo que representa uma amplitude transmitida, dada essencialmente como uma versão transladada do pacote inicial, $\psi_0(\vec{r} - \vec{v}_0 t - \vec{r}_0)$, com outro termo que representa uma amplitude espalhada, cuja dependência angular é determinada pela amplitude de espalhamento $f_{k_0}^{(+)}(\hat{r})$. A única região em que há superposição espacial dessas duas amplitudes é a região frontal, que corresponde ao ângulo de espalhamento $\theta \approx 0$ (v. figura 10.1). A forma da amplitude transmitida mostra imediatamente que, considerada isoladamente, ela tem a mesma normalização que o pacote inicial. Então, para que a densidade de probabilidade adicional, associada à onda espalhada (e portanto à seção de choque total de espalhamento) seja compatível com a conservação de norma, é preciso que ela seja compensada pelos termos de interferência entre as duas amplitudes, que existem apenas na região frontal $\theta \approx 0$. É preciso, portanto, que a amplitude de espalhamento, na região frontal em que ocorre a interferência, contenha informação suficiente acerca da seção de choque total de espalhamento para garantir a preservação da norma do estado evoluído completo. O que será feito a seguir é tratar de explicitar, em

termos quantitativos, a relação entre a amplitude de espalhamento na direção frontal e a seção de choque total que garante preservação de norma no processo de espalhamento.

A imposição da conservação de norma sobre a função de onda (10.14) leva à relação

$$1 = \int d^3r |\psi_0(\vec{r} - \vec{v}_0 t - \vec{r}_0)|^2 + \int d^3r \frac{|f_{k_0}^{(+)}(\hat{r})|^2}{r^2} |\psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2 + \\ + 2\text{Re} \int d^3r \psi_0^*(\vec{r} - \vec{v}_0 t - \vec{r}_0) \frac{f_{k_0}^{(+)}(\hat{r})}{r} \psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0),$$

em que o primeiro termo do lado direito é também igual a 1, enquanto o segundo termo pode ser escrito como

$$\int d\hat{r} |f_{k_0}^{(+)}(\hat{r})|^2 \int_0^\infty dr |\psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2 = \sigma_{\text{tot}} \int_0^\infty dr |\psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2$$

onde σ_{tot} é a seção de choque total. Juntando esses resultados a condição de conservação de norma fica reduzida a

$$\sigma_{\text{tot}} = - \frac{2\text{Re} \int_0^\infty dr \int d\hat{r} r f_{k_0}^{(+)}(\hat{r}) \psi_0^*(\vec{r} - \vec{v}_0 t - \vec{r}_0) \psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)}{\int_0^\infty dr |\psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)|^2} \quad (10.16)$$

que relaciona a seção de choque total (que é uma quantidade dada em termos do quadrado do módulo da amplitude de espalhamento) com uma expressão que é apenas *linear* em $f_{k_0}^{+}(\hat{r})$, por ser proveniente do termo de interferência entre a amplitude transmitida e a amplitude espalhada. Utilizando novamente a expressão (10.15) para a função $\psi_0(\vec{r} - \vec{r}_0)$ e fazendo uso da definição

$$F(\hat{r}, r) \equiv r e^{ik_0 r} f_{k_0}^{+}(\hat{r}) u(\vec{r} - \vec{v}_0 t - \vec{r}_0) u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0),$$

a integral que aparece no numerador pode ser escrita como

$$\int_0^\infty dr \int d\hat{r} e^{-i\vec{k}_0 \cdot \vec{r}} F(\hat{r}, r) = \int_0^\infty dr \int_0^{2\pi} d\varphi \int_{-1}^{+1} d(\cos \theta) e^{-ik_0 r \cos \theta} F(\hat{r}, r).$$

As contribuições relevantes dessa integral podem ser extraídas fazendo integrações por partes sobre a variável angular $\cos \theta$. Uma primeira integração dá

$$\int_{-1}^{+1} d(\cos \theta) e^{-ik_0 r \cos \theta} F(\hat{r}, r) = \frac{e^{-ik_0 r \cos \theta}}{-ik_0 r} F(\hat{r}, r) \Big|_{-1}^{+1} - \int_{-1}^{+1} d(\cos \theta) \frac{e^{-ik_0 r \cos \theta}}{-ik_0 r} \frac{dF(\hat{r}, r)}{d(\cos \theta)}.$$

O termo integrado se anula para $\cos \theta = -1$, que corresponde a $\theta = \pi$, pois nesse caso $\vec{r} = -\vec{k}_0 r$ e os argumentos dos dois fatores u são tais que eles nunca são simultaneamente diferentes de zero. Por outro lado, quando $\cos \theta = +1$ (o que corresponde a $\theta = 0$), $\vec{r} = \vec{k}_0 r$, a função $F(\hat{r}, r)$ se torna independente de φ e a contribuição para a integral é

$$2\pi \int_0^\infty dr \frac{e^{-ik_0 r}}{-ik_0 r} F(\hat{r} = \hat{k}_0, r) = \frac{2\pi i}{k_0} f_{k_0}^+(\theta = 0) \int_0^\infty dr \left[u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \right]^2. \quad (10.17)$$

A integral remanescente da primeira integração por partes pode ser novamente integrada por partes, dando lugar a termos integrados que contém um fator r^{-2} e que portanto se tornam desprezíveis para os valores grandes de r para os quais as amplitudes $u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0)$ são diferentes de zero. Do mesmo modo, a nova integral pode ser submetida a uma terceira integração por partes dando contribuições integradas com um fator r^{-3} , etc. Dessa forma a única contribuição relevante na região assintótica $r \rightarrow \infty$ é que é dada por (10.17), de modo que a relação (10.16) fica reduzida a

$$\sigma_{\text{tot}} = - \frac{2\text{Re} \left\{ \frac{2\pi i}{k_0} f_{k_0}^+(\theta = 0) \int_0^\infty dr \left[u(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \right]^2 \right\}}{\int_0^\infty dr \left| \psi_0(\hat{k}_0 r - \vec{v}_0 t - \vec{r}_0) \right|^2} = \frac{4\pi}{k_0} \text{Im} f_{k_0}^+(\theta = 0)$$

pois, em vista da relação (10.15), as integrais radiais que aparecem no numerador e no denominador são iguais. Esse resultado constitui o teorema óptico, já obtido anteriormente na seção 5.1.5, a partir da expansão em ondas parciais da amplitude de espalhamento. A peculiaridade que consiste em que esse resultado exprime uma “probabilidade” (a seção de choque total) em termos de uma expressão linear na *amplitude* de espalhamento pode ser vista agora, no entanto, como uma sinalização bastante clara de que ele envolve de forma essencial fenômenos quânticos de interferência.

10.3 Estados de espalhamento no espaço de momentos.

A representação no espaço de momentos da equação de Schrödinger (10.2), em termos da qual é descrito o espalhamento de partículas de massa M pelo potencial de curto alcance V pode ser obtida rapidamente projetando-a sobre o autovetor (impróprio) do momento $|\vec{k}\rangle$, relacionado com os autovetores da posição $|\vec{r}\rangle$ através da transformação

$$\langle \vec{r} | \vec{k} \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k} \cdot \vec{r}}.$$

Esses autovetores satisfazem as relações de completeza

$$\hat{1} = \int d^3 r |\vec{r}\rangle \langle \vec{r}| = \int d^3 k |\vec{k}\rangle \langle \vec{k}|$$

as quais implicam nas condições de normalização (impróprias) escritas em termos de funções delta de Dirac

$$\langle \vec{k} | \vec{k}' \rangle = \delta(\vec{k} - \vec{k}'), \quad \langle \vec{r} | \vec{r}' \rangle = \delta(\vec{r} - \vec{r}').$$

Desse modo se obtém, usando o fato de que a energia cinética é diagonal na representação de momentos,

$$\langle \vec{k} | \left(E - \frac{p^2}{2M} \right) | \phi_E \rangle = \left(E - \frac{\hbar^2 k^2}{2M} \right) \phi_E(\vec{k}) = \langle \vec{k} | V | \phi_E \rangle \quad (10.18)$$

onde $\phi_E(\vec{k}) \equiv \langle \vec{k} | \phi_E \rangle$ é a função de onda no espaço de momentos.

Usando a relação de completeza dos autovetores do momento, o segundo membro dessa equação pode ser escrito como

$$\langle \vec{k} | V | \phi_E \rangle = \int d^3 k' \langle \vec{k} | V | \vec{k}' \rangle \phi_E(\vec{k}');$$

o que mostra que a equação de Schrödinger no espaço de momentos tem a forma de uma equação integral *homogênea* para $\phi_E(\vec{k})$. A sua transformação em uma equação integral *inhomogênea*, equivalente à equação de Lippmann-Schwinger (10.4), envolve uma “inversão” do operador de multiplicação

$$\left(E - \frac{\hbar^2 k^2}{2M} \right) \equiv \frac{\hbar^2}{2M} (k_0^2 - k^2), \quad E \equiv \frac{\hbar^2 k_0^2}{2M} \quad (10.19)$$

que aparece no primeiro membro, e a adição de uma solução livre adequada às condições de contorno de espalhamento.

Estritamente falando, o operador (10.19) não pode ser invertido por se anular para $k = k_0$. Esse problema, no entanto, é a manifestação, na representação de momentos, da ambiguidade existente na determinação da função de Green para o operador diferencial que aparece no primeiro membro da equação diferencial correspondente (10.3), a qual é resolvida ali através da especificação de condições de contorno apropriadas. No espaço de momentos esse problema pode ser contornado fazendo uso do estratagema, devido a Dirac, que consiste em escrever

$$\left(E - \frac{\hbar^2 k^2}{2M} \right)^{-1} \longrightarrow \frac{\mathcal{P}}{E - \frac{\hbar^2 k^2}{2M}} + z \delta \left(E - \frac{\hbar^2 k^2}{2M} \right) \quad (10.20)$$

onde o símbolo \mathcal{P} significa que, numa integração sobre k , deve-se tomar o *valor principal* nas singularidades do integrando e o coeficiente z da função delta de Dirac que aparece no último termo é uma constante, em geral complexa, a ser determinada através da especificação de condições de contorno. Esse último termo é tal que

$$\left(E - \frac{\hbar^2 k^2}{2M}\right) z \delta\left(E - \frac{\hbar^2 k^2}{2M}\right) \psi(\vec{k}) = 0 \quad (10.21)$$

para quaisquer z e $\psi(\vec{k})$, pois o primeiro fator se anula no único ponto em que a função delta é diferente de zero.

Para escrever a equação correspondente à equação de Lippmann-Schwinger no espaço de momentos é preciso ainda especificar o termo inhomogêneo, que deve corresponder a uma solução do problema livre

$$\left(E - \frac{\hbar^2 k^2}{2M}\right) \varphi_{\vec{k}_0}(\vec{k}) = 0. \quad (10.22)$$

A escolha natural corresponde à solução de onda plana usada na representação de Schrödinger, que pode ser escrita no espaço de momentos como

$$\varphi_{\vec{k}_0}(\vec{k}) = (2\pi)^{3/2} \delta(\vec{k} - \vec{k}_0),$$

onde a direção do vetor \vec{k}_0 corresponde à direção da onda incidente da condição de contorno de espalhamento. A equação integral inhomogênea que corresponde à equação (10.18) é, portanto,

$$\phi_{\vec{k}_0}^{(z)}(\vec{k}) = (2\pi)^{3/2} \delta(\vec{k} - \vec{k}_0) + \left[\frac{\mathcal{P}}{E - \frac{\hbar^2 k^2}{2M}} + z \delta\left(E - \frac{\hbar^2 k^2}{2M}\right) \right] \langle \vec{k} | V | \phi_{\vec{k}_0}^{(z)} \rangle. \quad (10.23)$$

Manipulando os operadores singulares que aparecem nessa expressão com uso das relações (10.21) e (10.22), é fácil verificar que a função $\phi_{\vec{k}_0}^{(z)}(\vec{k})$ satisfaz a equação (10.18) para qualquer escolha de z . As escolhas que correspondem às condições de contorno de espalhamento usuais (10.1) podem ser identificadas transformando a equação (10.23) de volta para a representação Schrödinger, o que pode ser feito notando que

$$\phi_{\vec{k}_0}^{(z)}(\vec{r}) = \langle \vec{r} | \phi_{\vec{k}_0}^{(z)} \rangle = \int d^3k \langle \vec{r} | \vec{k} \rangle \phi_{\vec{k}_0}^{(z)}(\vec{k}) = \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\vec{k} \cdot \vec{r}} \phi_{\vec{k}_0}^{(z)}(\vec{k}).$$

O elemento de matriz que envolve o potencial V no segundo membro de (10.23) pode, por sua vez, ser escrito também em termos de $\phi_{\vec{k}_0}^{(z)}(\vec{r})$

$$\langle \vec{k} | V | \phi_{\vec{k}_0}^{(z)} \rangle = \int \frac{d^3 r'}{(2\pi)^{3/2}} e^{-i\vec{k} \cdot \vec{r}'} \langle \vec{r}' | V | \phi_{\vec{k}_0}^{(z)} \rangle = \int \frac{d^3 r'}{(2\pi)^{3/2}} e^{-i\vec{k} \cdot \vec{r}'} V(\vec{r}') \phi_{\vec{k}_0}^{(z)}(\vec{r}')$$

o que permite reduzir a equação à forma

$$\phi_{\vec{k}_0}^{(z)}(\vec{r}) = e^{i\vec{k}_0 \cdot \vec{r}} + \int \frac{d^3 r'}{(2\pi)^{3/2}} \int \frac{d^3 k}{(2\pi)^{3/2}} e^{i\vec{k} \cdot (\vec{r} - \vec{r}')} \left[\frac{\mathcal{P}}{E - \frac{\hbar^2 k^2}{2M}} + z \delta \left(E - \frac{\hbar^2 k^2}{2M} \right) \right] V(\vec{r}') \phi_{\vec{k}_0}^{(z)}(\vec{r}').$$

A transformação pode ser completada calculando as integrais sobre \vec{k} , o que pode ser feito sem dificuldade através de procedimentos usuais. A integral envolvendo o valor principal do operador invertido é

$$\mathcal{P} \int \frac{d^3 k}{(2\pi)^{3/2}} \frac{e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}}{E - \frac{\hbar^2 k^2}{2M}} = -\sqrt{2\pi} \frac{M}{\hbar^2} \frac{\cos k_0(|\vec{r} - \vec{r}'|)}{|\vec{r} - \vec{r}'|}$$

enquanto a integral que envolve a função delta de Dirac leva ao resultado

$$\int \frac{d^3 k}{(2\pi)^{3/2}} e^{i\vec{k} \cdot (\vec{r} - \vec{r}')} z \delta \left(E - \frac{\hbar^2 k^2}{2M} \right) = \sqrt{\frac{2}{\pi}} \frac{M}{\hbar^2} z \frac{\sin k_0(|\vec{r} - \vec{r}'|)}{|\vec{r} - \vec{r}'|}.$$

A equação (10.23), transformada para a representação de Schrödinger, tem então a forma

$$\phi_{\vec{k}_0}^{(z)}(\vec{r}) = e^{i\vec{k}_0 \cdot \vec{r}} - \frac{M}{2\pi\hbar^2} \int d^3 r' \left[\frac{\cos k_0(|\vec{r} - \vec{r}'|)}{|\vec{r} - \vec{r}'|} - \frac{z}{\pi} \frac{\sin k_0(|\vec{r} - \vec{r}'|)}{|\vec{r} - \vec{r}'|} \right] V(\vec{r}') \phi_{\vec{k}_0}^{(z)}(\vec{r}')$$

que coincide com a equação de Lippmann-Schwinger (10.4) com condições de contorno de espalhamento (10.1) com as escolhas

$$z = \mp i\pi \quad \text{para} \quad \phi_{\vec{k}_0}^{(z)}(\vec{r}) \rightarrow \phi_{\vec{k}_0}^{\pm}(\vec{r}).$$

A equação integral (10.23), com essas mesmas escolhas de z , constitui portanto a versão da equação de Lippmann-Schwinger no espaço de momentos. Ela pode ser escrita como

$$\phi_{\vec{k}_0}^{\pm}(\vec{k}) = (2\pi)^{3/2} \delta(\vec{k} - \vec{k}_0) - \frac{M}{2\pi\hbar^2} G_{\vec{k}_0}^{\pm}(k) \int d^3 k' \langle \vec{k} | V | \vec{k}' \rangle \phi_{\vec{k}_0}^{\pm}(\vec{k}'). \quad (10.24)$$

onde as funções de Green $G_{\vec{k}_0}^{\pm}(k)$ são

$$G_{k_0}^{\pm}(k) = -\frac{2\pi\hbar^2}{M} \left[\frac{\mathcal{P}}{E - \frac{\hbar^2 k^2}{2M}} \mp i\pi \delta \left(E - \frac{\hbar^2 k^2}{2M} \right) \right].$$

As integrais sobre momentos usadas para transformar a equação (10.23) de volta para a representação de Schrödinger podem ser usadas novamente aqui para verificar que essas funções de Green se relacionam com as funções de Green usadas para obter a equação de Lippmann-Schwinger (10.4) através da relação

$$\int \frac{d^3k}{(2\pi)^3} G_{k_0}^{\pm}(k) e^{i\vec{k} \cdot (\vec{r} - \vec{r}')} = \frac{e^{\pm ik_0 |\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} = G_{k_0}(\vec{r}, \vec{r}')$$

que pode ser vista simplesmente como a relação de transformação de volta para a representação de Schrödinger de um *operador* $G_{k_0}^{\pm}$ que é *diagonal* no espaço de momentos, onde sua forma é

$$\langle \vec{k} | G_{k_0}^{\pm} | \vec{k}' \rangle \equiv G_{k_0}^{\pm}(k) \delta(\vec{k} - \vec{k}'),$$

pois, de fato,

$$\langle \vec{r} | G_{k_0}^{\pm} | \vec{r}' \rangle = \int d^3k \int d^3k' \langle \vec{r} | \vec{k} \rangle \langle \vec{k} | G_{k_0}^{\pm} | \vec{k}' \rangle \langle \vec{k}' | \vec{r}' \rangle = \int \frac{d^3k}{(2\pi)^3} G_{k_0}^{\pm}(k) e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}.$$

Esse operador pode ser escrito sob a forma compacta

$$G_{k_0}^{\pm} = -\frac{2\pi\hbar^2}{M} \lim_{\epsilon \rightarrow 0} \frac{1}{E \pm i\epsilon - \frac{p^2}{2M}}$$

com $\epsilon > 0$, como pode ser visto separando as partes real e imaginária

$$\frac{1}{E + i\epsilon - \frac{p^2}{2M}} = \frac{E - \frac{p^2}{2M}}{\left(E - \frac{p^2}{2M}\right)^2 + \epsilon^2} \mp \frac{i\epsilon}{\left(E - \frac{p^2}{2M}\right)^2 + \epsilon^2}$$

e notando que, no limite em que $\epsilon \rightarrow 0$, os elementos de matriz desses objetos entre autovetores do momento se comportam como

$$\langle \vec{k} | \frac{E - \frac{p^2}{2M}}{\left(E - \frac{p^2}{2M}\right)^2 + \epsilon^2} | \vec{k}' \rangle \rightarrow \frac{\mathcal{P}}{E - \frac{\hbar^2 k^2}{2M}} \delta(\vec{k} - \vec{k}')$$

e

$$\langle \vec{k} | \frac{\epsilon}{\left(E - \frac{p^2}{2M}\right)^2 + \epsilon^2} | \vec{k}' \rangle \rightarrow \pi \delta \left(E - \frac{\hbar^2 k^2}{2M} \right) \delta(\vec{k} - \vec{k}').$$

Isso é freqüentemente resumido em termos da “identidade”

$$\lim_{\epsilon \rightarrow 0} \frac{1}{E \pm i\epsilon - \frac{p^2}{2M}} \equiv \frac{\mathcal{P}}{E - \frac{p^2}{2M}} \mp i\pi \delta \left(E - \frac{p^2}{2M} \right)$$

com o auxílio da qual a equação (10.24) pode ser posta sob a forma

$$\phi_{\vec{k}_0}^{\pm}(\vec{k}) = (2\pi)^{3/2} \delta(\vec{k} - \vec{k}_0) + \frac{1}{E \pm i\epsilon - \frac{\hbar^2 k^2}{2M}} \int d^3 k' \langle \vec{k} | V | \vec{k}' \rangle \phi_{\vec{k}_0}^{\pm}(\vec{k}'). \quad (10.25)$$

onde está subentendido o limite $\epsilon \rightarrow 0$ por valores positivos. Uma comparação com a equação equivalente (10.4) mostra que, no espaço de momentos, a função de Green aparece sob a forma de um operador singular mas *diagonal*, estreitamente relacionado com o hamiltoniano livre $H_0 \equiv p^2/2M$. O potencial, que foi tomado como um potencial local $V(\vec{r})$ na representação de Schrödinger para escrever a equação (10.4), aparece como um operador não diagonal na representação de momentos, eq. (10.25). Uma expansão de Born no espaço de momentos pode evidentemente ser obtida iterando essa equação, de forma análoga ao que foi feito na seção 10.1.1. A amplitude de espalhamento na primeira aproximação de Born é dada agora diretamente em termos dos elementos de matriz do potencial no espaço de momentos, enquanto aproximações de ordem superior envolvem integrações sobre estados intermediários de momento e denominadores de energia singulares, mas com uma prescrição definida para o tratamento das singularidades dada pela presença das quantidades $\pm i\epsilon$.

A equivalência das duas representações sugere, por outro lado, que seja possível também tratar a equação de Lippmann-Schwinger diretamente em termos de operadores (eventualmente singulares) de forma abstrata, i.e., sem introduzir explicitamente qualquer representação específica. O caráter diagonal da função de Green obtida na representação de momentos é, de fato, extremamente útil para determinar a forma abstrata da equação. Um tratamento desse tipo, a ser desenvolvido na seção seguinte, evita complicações técnicas inerentes a uma ou outra representação específica, e dessa forma facilita consideravelmente manipulações formais que permitem explicitar mecanismos importantes do processo de espalhamento.

10.4 Equação de Lippmann-Schwinger formal e operador de transição.

Com uso dos recursos desenvolvidos na seção anterior é muito simples obter uma versão da equação de Lippmann-Schwinger sem que seja necessário introduzir uma representação específica, como foram, anteriormente, a representação de Schrödinger e a representação de momentos. O ponto de partida é a versão formal da equação de Schrödinger

$$(E - H_0)|\phi_{\vec{k}_0}\rangle = V|\phi_{\vec{k}_0}\rangle \quad (10.26)$$

onde $H_0 = p^2/2M$, $E = \hbar^2 k_0^2/2M$ e V é o potencial de curto alcance responsável pelo espalhamento. O feixe incidente é representado por um autovetor $|\vec{k}_0\rangle$ de H_0 com momento bem definido $\hbar\vec{k}_0$, que satisfaz portanto a equação

$$(E - H_0)|\vec{k}_0\rangle = 0,$$

de modo que a equação integral formal que equivale à equação de Schrödinger (10.26) com condições de contorno de espalhamento pode ser escrita sob a forma

$$|\phi_{\vec{k}_0}^\pm\rangle = |\vec{k}_0\rangle + \frac{1}{E \pm i\epsilon - H_0} V |\phi_{\vec{k}_0}^\pm\rangle. \quad (10.27)$$

Estritamente falando, essa equação difere das equações (10.4) e (10.23) por um fator de normalização $(2\pi)^{3/2}$, como pode ser facilmente verificado notando, por exemplo, que o termo inhomogêneo é aqui dado, na representação de Schrödinger, por

$$\langle \vec{r} | \vec{k}_0 \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}_0 \cdot \vec{r}} \quad (10.28)$$

que difere do termo correspondente de (10.4) por esse fator. O mesmo fator aparece na representação de momentos, na qual o termo inhomogêneo de (10.27) é simplesmente $\langle \vec{k} | \vec{k}_0 \rangle = \delta(\vec{k} - \vec{k}_0)$, enquanto um fator adicional $(2\pi)^{3/2}$ aparece em (10.23). As soluções das equações integrais são, no entanto, soluções da equação *homogênea* (10.26), de modo que diferentes critérios de normalização podem ser trivialmente implementados a posteriori, e a forma adotada para (10.27) pode ser mantida, por simplicidade, sem qualquer perda de generalidade. A forma da função de Green que foi utilizada nessa equação se justifica facilmente a partir dos resultados obtidos na representação de momentos. Em particular, a notação usada subentende o limite $\epsilon \rightarrow 0$ por valores positivos, o que permite também utilizar a identidade formal

$$\frac{1}{E \pm i\epsilon - H_0} = \frac{\mathcal{P}}{E - H_0} \mp i\pi\delta(E - H_0). \quad (10.29)$$

As funções singulares de operadores que aparecem nessa identidade podem ser entendidas, por exemplo, utilizando os autovetores de H_0 que são autovetores simultâneos do momento, para os quais se pode escrever, por exemplo

$$\frac{\mathcal{P}}{E - H_0} |\vec{k}\rangle = \frac{\mathcal{P}}{E - \frac{\hbar^2 k^2}{2M}} |\vec{k}\rangle$$

e uma expressão análoga para o termo envolvendo a função delta, o que permite fazer contacto direto com as expressões obtidas na representação de momentos.

A série de Born pode agora ser obtida iterando a equação (10.27), o que leva facilmente ao resultado

$$\begin{aligned} |\phi_{\vec{k}_0}\rangle &= |\vec{k}_0\rangle + \frac{1}{E \pm i\epsilon - H_0} V |\vec{k}_0\rangle + \frac{1}{E \pm i\epsilon - H_0} V \frac{1}{E \pm i\epsilon - H_0} V |\vec{k}_0\rangle + \dots = \\ &= \left[1 + \frac{1}{E \pm i\epsilon - H_0} + \frac{1}{E \pm i\epsilon - H_0} V \frac{1}{E \pm i\epsilon - H_0} V + \dots \right] |\vec{k}_0\rangle. \end{aligned} \quad (10.30)$$

A série *de operadores* que aparece entre colchetes pode ser porém somada formalmente, dando

$$\left[1 + \frac{1}{E \pm i\epsilon - H_0} + \frac{1}{E \pm i\epsilon - H_0} V \frac{1}{E \pm i\epsilon - H_0} V + \dots \right] = \frac{1}{1 - \frac{1}{E \pm i\epsilon - H_0} V}$$

e este último objeto pode ser por sua vez re-escrito como

$$\begin{aligned} \frac{1}{1 - \frac{1}{E \pm i\epsilon - H_0} V} &= \left[1 - \frac{1}{E \pm i\epsilon - H_0} V \right]^{-1} = \left[\frac{1}{E \pm i\epsilon - H_0} (E \pm i\epsilon - H_0 - V) \right]^{-1} = \\ &= \frac{1}{E \pm i\epsilon - H_0 - V} (E \pm i\epsilon - H_0) \end{aligned} \quad (10.31)$$

onde, na última passagem, foi usado o fato de que o inverso de um produto de operadores é o produto dos operadores inversos em ordem contrária. O resultado deste cálculo representa a soma da série que aparece entre colchetes em (10.30), de modo que os vetores $|\phi_{\vec{k}_0}^{\pm}\rangle$ podem ser escritos sob as formas alternativas

$$|\phi_{\vec{k}_0}^{\pm}\rangle = \frac{1}{E \pm i\epsilon - H_0 - V} (E \pm i\epsilon - H_0) |\vec{k}_0\rangle = \frac{\pm i\epsilon}{E \pm i\epsilon - H_0 - V} |\vec{k}_0\rangle$$

pois $(E - H_0) |\vec{k}_0\rangle = 0$, e ainda

$$|\phi_{\vec{k}_0}^{\pm}\rangle = \frac{1}{E \pm i\epsilon - H_0 - V} (E \pm i\epsilon - H_0 - V + V) |\vec{k}_0\rangle = |\vec{k}_0\rangle + \frac{1}{E \pm i\epsilon - H} V |\vec{k}_0\rangle \quad (10.32)$$

com $H = H_0 + V$. Esses dois resultados exprimem os estados de espalhamento $|\phi_{\vec{k}_0}^{\pm}\rangle$ em termos de operadores que agem *sobre o estado livre* $|\vec{k}_0\rangle$, e devem ser vistos portanto como *soluções* formais da equação de Lippmann-Schwinger. A sua utilidade prática é limitada pelo fato de que nelas intervém a função de Green *completa*, na qual aparece não só o hamiltoniano livre H_0 mas o hamiltoniano completo $H = H_0 + V$, e a obtenção dessa função de Green é um problema cuja dificuldade é equivalente à do problema de resolver a equação de Schrödinger (10.26).

A solução formal (10.32) permite, no entanto, obter uma expressão interessante para a amplitude de espalhamento, e portanto para a seção de choque diferencial. De fato, o estudo da equação de Lippmann-Schwinger na representação de Schrödinger permitiu exprimir as amplitudes de espalhamento $f_{\vec{k}_0}^{\pm}$ em termos de elementos de matriz do potencial que envolvem os estados de espalhamento $|\phi_{\vec{k}_0}^{\pm}\rangle$ e ondas planas na direção final de observação, eq. (10.5). Essas expressões foram obtidas, no entanto, com a normalização dos estados de espalhamento determinada por um termo inhomogêneo escrito simplesmente como $e^{i\vec{k}_0 \cdot \vec{r}}$ e também com uma opção análoga para a onda plana que representa o estado final. Por é preciso ajustar os coeficientes numéricos de modo a ter em conta os fatores adicionais $1/(2\pi)^{3/2}$ que correspondem à normalização adotada aqui para o termo inhomogêneo $|\vec{k}_0\rangle$ da equação de Lippmann-Schwinger formal e para os estados finais $|\pm k_0 \hat{r}\rangle$. O ajuste consiste de um fator adicional $(2\pi)^3$, o que leva a

$$f_{k_0}^{(\pm)}(\hat{r}) = -\frac{(2\pi)^2 M}{\hbar^2} \langle \pm k_0 \hat{r} | V | \phi_{\vec{k}_0}^{\pm} \rangle.$$

Usando então a solução formal (10.32) resulta

$$f_{k_0}^{(\pm)}(\hat{r}) = -\frac{(2\pi)^2 M}{\hbar^2} \langle \pm k_0 \hat{r} | V + V \frac{1}{E \pm i\epsilon - H} V | \vec{k}_0 \rangle \quad (10.33)$$

que agora envolve elementos de matriz entre vetores de estado livres dos *operadores de transição*

$$T^{\pm}(k_0) \equiv V + V \frac{1}{E \pm i\epsilon - H} V. \quad (10.34)$$

Esses operadores dependem de k_0 através da energia E que aparece na função de Green, e a seção de choque diferencial pode ser obtida diretamente de $f_{k_0}^{(+)}(\hat{r})$ como

$$\frac{d\sigma}{d\Omega} = |f_{k_0}^{(+)}(\hat{r})|^2 = \frac{(2\pi)^4 M^2}{\hbar^4} \left| \langle k_0 \hat{r} | T^{+}(k_0) | \vec{k}_0 \rangle \right|^2. \quad (10.35)$$

Essa expressão para a seção de choque diferencial é *exata*. É interessante notar que ela pode ser obtida de uma probabilidade de transição por unidade de tempo, do estado inicial $|\vec{k}_0\rangle$ para estados finais $|\vec{k}\rangle$ escrita sob a forma da regra áurea de Fermi, mas na qual o potencial é substituído pelo operador de transição. Essa probabilidade de transição, tendo em conta que os estados finais de interesse são aqueles em que a direção do vetor de onda final \vec{k} está contida no elemento de ângulo sólido $d\Omega$ é

$$dW = \frac{2\pi}{\hbar} d\Omega \int_0^\infty dk k^2 \left| \langle \vec{k} | T^+(k_0) | \vec{k}_0 \rangle \right|^2 \delta \left(\frac{\hbar^2 k^2}{2M} - \frac{\hbar^2 k_0^2}{2M} \right) = \frac{2\pi}{\hbar} d\Omega \frac{k_0 M}{\hbar^2} \left| \langle k_0 \hat{r} | T^+(k_0) | \vec{k}_0 \rangle \right|^2.$$

De fato, a seção de choque correspondente, obtida da forma usual

$$\frac{d\sigma}{d\Omega} = \frac{1}{\text{fluxo incidente}} \frac{dW}{d\Omega} = \frac{(2\pi)^3 M}{\hbar k_0} \frac{dW}{d\Omega}$$

coincide com a expressão (10.35). O valor usado para o fluxo incidente corresponde à onda plana (10.28). É claro que, se o operador de transição é substituído simplesmente pelo potencial V , o que se obtém é simplesmente a seção de choque calculada em teoria de perturbação de primeira ordem, o que corresponde à aproximação de Born. O ponto que deve ser salientado aqui é o de que o resultado exato é compatível com a estrutura da regra áurea de Fermi, mediante uma definição apropriada do operador responsável pela transição. Neste caso a definição apropriada resulta da solução formal (10.32) da equação de Lippmann-Schwinger, da qual decorre imediatamente que

$$V|\phi_{\vec{k}_0}^\pm\rangle = T^\pm(k_0)|\vec{k}_0\rangle. \quad (10.36)$$

Uma consequência relevante dessa relação é a de que a amplitude de espalhamento $f_{k_0}^{(+)}(\hat{r})$ também pode ser escrita como

$$f_{k_0}^{(+)}(\hat{r}) = -\frac{(2\pi)^2 M}{\hbar^2} \langle \phi_{k_0 \hat{r}}^- | V | \vec{k}_0 \rangle. \quad (10.37)$$

De fato, para o estado final com a condição de contorno de ondas *imergentes* $|\phi_{k_0 \hat{r}}^- \rangle$ vale a relação

$$V|\phi_{k_0 \hat{r}}^- \rangle = T^-(k_0)|k_0 \hat{r}\rangle$$

que, tendo em vista que

$$(T^-(k_0))^\dagger = T^+(k_0),$$

mostra que (10.37) é equivalente à expressão (10.33) para $f_{k_0}^{(+)}(\hat{r})$.

Os operadores de transição (10.34) podem ser definidos, de forma geral, em termos das equações integrais formais equivalentes

$$T^\pm(k_0) = V + V \frac{1}{E \pm i\epsilon - H_0} T^\pm(k_0) \quad \text{ou} \quad T^\pm(k_0) = V + T^\pm(k_0) \frac{1}{E \pm i\epsilon - H_0} V.$$

Isso pode ser verificado iterando cada uma dessas equações, o que leva à série geométrica formal

$$\begin{aligned} T^\pm(k_0) &= V + V \frac{1}{E \pm i\epsilon - H_0} V + V \frac{1}{E \pm i\epsilon - H_0} V \frac{1}{E \pm i\epsilon - H_0} V + \dots = \\ &= V + V \left[1 + \frac{1}{E \pm i\epsilon - H_0} V + \dots \right] \frac{1}{E \pm i\epsilon - H_0} V = \\ &= V + V \left[1 - \frac{1}{E \pm i\epsilon - H_0} V \right]^{-1} \frac{1}{E \pm i\epsilon - H_0} V. \end{aligned}$$

O operador inverso que aparece nesta última expressão já foi calculado em (10.31)

$$\left[1 - \frac{1}{E \pm i\epsilon - H_0} V \right]^{-1} = \frac{1}{E \pm i\epsilon - H} (E \pm i\epsilon - H_0)$$

e esse resultado leva finalmente à expressão (10.34) para $T^\pm(k_0)$. As equações integrais formais poder ser também realizadas em termos de alguma representação específica. Usando, por exemplo, a representação de momentos para escrever a primeira das duas equações integrais, resulta

$$\langle \vec{k} | T^\pm(k_0) | \vec{k}' \rangle = \langle \vec{k} | V | \vec{k}' \rangle + \int d^3 k_1 \langle \vec{k} | V | \vec{k}_1 \rangle \frac{1}{E \pm i\epsilon - \frac{\hbar^2 k_1^2}{2M}} \langle \vec{k}_1 | T^\pm(k_0) | \vec{k}' \rangle$$

de cuja iteração podem ser obtidas “séries de Born” para as matrizes que representam os operadores de transição no espaço de momentos. É importante ressaltar o ponto de que a determinação de $T^\pm(k_0)$ envolve o conhecimento dessas matrizes *para todos os valores de \vec{k} e \vec{k}'* . A seção de choque de espalhamento (10.35), por outro lado, depende apenas dos elementos de matriz em que esses dois vetores têm o mesmo módulo, o qual é por sua vez igual ao argumento k_0 do operador de transição $T^\pm(k_0)$. Essa restrição provém em última análise da conservação de energia no processo de espalhamento, e por isso esses elementos de matriz são usualmente chamados “elementos de matriz do operador de transição na superfície de energia”, entendida como definida pela restrição $|\vec{k}| = |\vec{k}'| = k_0$. Portanto, a parte da matriz de transição $\langle \vec{k} | T^\pm(k_0) | \vec{k}' \rangle$ que está *fora da superfície de energia*, não é diretamente acessível à medida através de processos de espalhamento.

10.4.1 A “fórmula de dois potenciais” de Gell-Mann e Goldberger.

Um resultado que pode ser obtido de forma relativamente simples em termos da teoria formal desenvolvida acima e que é extremamente útil para tratar problemas de espalhamento é a chamada “fórmula de dois potenciais”, obtida originalmente por Gell-Mann e Goldberger⁵ Ela dá uma expressão para o operador de transição na superfície de energia, no caso em que o potencial V é tomado como dado pela soma de dois potenciais $V = V_1 + V_2$, em termos do operador de transição correspondente ao potencial V_1 apenas, mais um termo adicional que descreve os efeitos de V_2 no processo de espalhamento.

O estado $|\phi_{\vec{k}_0}^+\rangle$ que descreve o espalhamento apenas pelos dois potenciais $V_1 + V_2$ é dado em termos da solução formal (10.32) da equação de Lippmann-Schwinger como

$$|\phi_{\vec{k}_0}^+\rangle = |\vec{k}_0\rangle + \frac{1}{E + i\epsilon - H_0 - V_1 - V_2}(V_1 + V_2)|\vec{k}_0\rangle. \quad (10.38)$$

Essa expressão pode ser transformada com uso da identidade

$$\frac{1}{E + i\epsilon - H_0 - V_1 - V_2} \equiv \frac{1}{E + i\epsilon - H_0 - V_1} + \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1 - V_2}.$$

As manipulações envolvidas na verificação da identidade acima são típicas dos procedimentos utilizados no contexto da teoria formal do espalhamento. Para verificar essa identidade, o passo inicial consiste em escrever o primeiro membro sob a forma da série

$$\begin{aligned} \frac{1}{E + i\epsilon - H_0 - V_1 - V_2} &= \frac{1}{E + i\epsilon - H_0 - V_1} + \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1} \\ &+ \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1} \\ &+ \dots \end{aligned}$$

e, em seguida, re-somar a série como

$$\begin{aligned} &\frac{1}{E + i\epsilon - H_0 - V_1} + \\ &+ \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1} \left[1 + V_2 \frac{1}{E + i\epsilon - H_0 - V_1} + \dots \right] = \\ &= \frac{1}{E + i\epsilon - H_0 - V_1} + \\ &+ \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1} \left[1 - V_2 \frac{1}{E + i\epsilon - H_0 - V_1} \right]^{-1} \end{aligned} \quad (10.39)$$

⁵M. Gell-Mann e M. L. Goldberger, *Phys. Rev.* **91**, 398 (1953).

O operador inverso que aparece entre colchetes nesta última expressão pode agora ser re-escrito sob a forma

$$\begin{aligned} \left[1 - V_2 \frac{1}{E + i\epsilon - H_0 - V_1} \right]^{-1} &= \left[(E + i\epsilon - H_0 - V_1 - V_2) \frac{1}{E + i\epsilon - H_0 - V_1} \right]^{-1} \\ &= (E + i\epsilon - H_0 - V_1) \frac{1}{E + i\epsilon - H_0 - V_1 - V_2} \end{aligned}$$

que, substituída em (10.39), dá a identidade buscada.

Essa identidade tem o efeito imediato de permitir que a função $|\phi_{\vec{k}_0}^+\rangle$ seja escrita sob a forma

$$\begin{aligned} |\phi_{\vec{k}_0}^+\rangle &= |\vec{k}_0\rangle + \frac{1}{E + i\epsilon - H_0 - V_1} (V_1 + V_2) |\vec{k}_0\rangle + \\ &+ \frac{1}{E + i\epsilon - H_0 - V_1} V_2 \frac{1}{E + i\epsilon - H_0 - V_1 - V_2} (V_1 + V_2) |\vec{k}_0\rangle \end{aligned}$$

que, tendo em vista a relação (10.38), e também a sua correspondente para o espalhamento apenas pelo potencial V_1 , dada por

$$|\chi_{\vec{k}_0}^+\rangle = |\vec{k}_0\rangle + \frac{1}{E + i\epsilon - H_0 - V_1} V_1 |\vec{k}_0\rangle,$$

leva à equação integral para $|\phi_{\vec{k}_0}^+\rangle$

$$|\phi_{\vec{k}_0}^+\rangle = |\chi_{\vec{k}_0}^+\rangle + \frac{1}{E + i\epsilon - H_0 - V_1} V_2 |\phi_{\vec{k}_0}^+\rangle. \quad (10.40)$$

Esta última equação é uma equação de Lippmann-Schwinger cujo termo inhomogêneo é o estado com condições de contorno de espalhamento *pelo potencial V_1 apenas*. Os efeitos do potencial V_2 estão todos contidos no último termo, no sentido de que este termo se anula quando o operador $V_2 \rightarrow 0$. Ela pode também ser iterada dando origem a uma série análoga à série de Born, eventualmente chamada de *série de Born com ondas distorcidas* (pelo potencial V_1), que é uma série em potências de V_2 na qual os efeitos de V_1 são tratados exatamente.

A chamada fórmula de dois potenciais de Gell-Mann e Goldberger é o resultado correspondente a essa equação de Lippmann-Schwinger com ondas distorcidas formulado em termos do operador de transição. Ela pode ser obtida partindo da expressão para o elemento de matriz *na superfície de energia* do operador de transição para os dois potenciais

$$\langle k_0 \hat{r} | T^+(k_0) | \vec{k}_0 \rangle = \langle k_0 \hat{r} | (V_1 + V_2) | \phi_{\vec{k}_0}^+ \rangle$$

e usando a equação (10.40) para re-escreve-la como

$$\begin{aligned}
\langle k_0 \hat{r} | T^+(k_0) | \vec{k}_0 \rangle &= \langle k_0 \hat{r} | (V_1 + V_2) | \chi_{\vec{k}_0}^+ \rangle + \langle k_0 \hat{r} | (V_1 + V_2) \frac{1}{E + i\epsilon - H_0 - V_1} V_2 | \phi_{\vec{k}_0}^+ \rangle \\
&= \langle k_0 \hat{r} | V_1 | \chi_{\vec{k}_0}^+ \rangle + \langle k_0 \hat{r} | V_1 \frac{1}{E + i\epsilon - H_0 - V_1} V_2 | \phi_{\vec{k}_0}^+ \rangle + \\
&\quad + \langle k_0 \hat{r} | V_2 \left(| \chi_{\vec{k}_0}^+ \rangle + \frac{1}{E + i\epsilon - H_0 - V_1} V_2 | \phi_{\vec{k}_0}^+ \rangle \right).
\end{aligned}$$

Tendo em vista a equação (10.40), o último termo se reduz na realidade a $\langle k_0 \hat{r} | V_2 | \phi_{\vec{k}_0}^+ \rangle$. O segundo termo, porém pode ser reduzido introduzindo o estado com a condição de contorno assintótica de ondas *imergentes*

$$| \chi_{k_0 \hat{r}}^- \rangle = | k_0 \hat{r} \rangle + \frac{1}{E - i\epsilon - H_0 - V_1} V_1 | k_0 \hat{r} \rangle$$

pois dessa relação segue que

$$\langle k_0 \hat{r} | V_1 \frac{1}{E + i\epsilon - H_0 - V_1} = \langle \chi_{k_0 \hat{r}}^- | - \langle k_0 \hat{r} |$$

o que leva finalmente à expressão para o operador de transição total na superfície de energia

$$\begin{aligned}
\langle k_0 \hat{r} | T^+(k_0) | \vec{k}_0 \rangle &= \langle k_0 \hat{r} | V_1 | \chi_{\vec{k}_0}^+ \rangle + \langle \chi_{k_0 \hat{r}}^- | V_2 | \phi_{\vec{k}_0}^+ \rangle = \\
&= \langle k_0 \hat{r} | T_1^+(k_0) | \vec{k}_0 \rangle + \langle \chi_{k_0 \hat{r}}^- | V_2 | \phi_{\vec{k}_0}^+ \rangle
\end{aligned} \tag{10.41}$$

onde $T_1^+(k_0)$ é o operador de transição para o potencial V_1 apenas. É muito importante notar que essa expressão exata envolve, no último termo, um estado de espalhamento pelo potencial V_1 com condições de contorno assintóticas de ondas imergentes. Em particular, a chamada *aproximação de Born com ondas distorcidas* é obtida simplesmente substituindo nesse último termo o estado de espalhamento exato $| \phi_{\vec{k}_0}^+ \rangle$ por sua aproximação de ordem zero em V_2 , isto é, $| \chi_{\vec{k}_0}^+ \rangle$. A expressão resultante

$$\langle k_0 \hat{r} | T^+(k_0) | \vec{k}_0 \rangle \approx \langle k_0 \hat{r} | T_1^+(k_0) | \vec{k}_0 \rangle + \langle \chi_{k_0 \hat{r}}^- | V_2 | \chi_{\vec{k}_0}^+ \rangle$$

também envolve de modo essencial um estado de espalhamento com ondas imergentes no termo linear em V_2 .

10.5 Operador de espalhamento S. Propriedades formais.

Embora os operadores de transição $T^\pm(k_0)$, eq. (10.34), sejam suficientes para determinar todas as quantidades que podem ser medidas em processos de espalhamento, eles não são, em geral, determinados por essas quantidades. Isso significa que a restrição do operador $T^+(k_0)$ à superfície de energia é em geral insuficiente para determinar completamente esse operador. Devido a esse fato, para a análise de processos de espalhamento é conveniente dispor de um operador alternativo que seja completamente determinado pela restrição de $T^+(k_0)$ à superfície de energia. Isso é feito definindo um *operador de espalhamento* S que pode ser caracterizado formalmente através da relação, válida para qualquer par de vetores de onda \vec{k} e \vec{k}' ,

$$\langle \vec{k} | S | \vec{k}' \rangle = \langle \phi_{\vec{k}}^- | \phi_{\vec{k}'}^+ \rangle.$$

Uma justificação, em termos qualitativos, dessa definição formal, pode ser dada em termos das relações (10.11) e (10.12), segundo as quais os autovetores $|\phi_{\vec{k}'}^+\rangle$ e $|\phi_{\vec{k}}^-\rangle$ são apropriados para expandir respectivamente pacotes de onda *pré-colisão* (isto é, pacotes de onda que se aproximam do alvo) e *pós-colisão* (pacotes que se afastam do alvo). Desse modo, superposições apropriadas dos elementos de matriz do operador S $\langle \vec{k} | S | \vec{k}' \rangle$ correspondem à probabilidade de que um dado pacote pré-colisão evolua para um também dado pacote pós colisão.

A relação formal existente entre os operadores S e T pode ser obtida a partir da solução (10.32) da equação de Lippmann-Schwinger escrita com o auxílio de uma identidade semelhante à (10.29) sob a forma

$$|\phi_{\vec{k}}^\pm\rangle = |\vec{k}\rangle + \frac{\mathcal{P}}{E_k - H} V |\vec{k}\rangle \mp i\pi\delta(E_k - H) V |\vec{k}\rangle$$

onde $E_k \equiv \hbar^2 k^2 / 2M$ e $H = p^2 / 2M + V$. Essa expressão permite relacionar os autovetores do espectro contínuo de H com condições de contorno assintóticas de ondas emergentes e de ondas imergentes através de

$$|\phi_{\vec{k}}^-\rangle = |\phi_{\vec{k}}^+\rangle + 2\pi i\delta(E_k - H) V |\vec{k}\rangle$$

que, levada à definição de S , dá

$$\langle \vec{k} | S | \vec{k}' \rangle = \langle \phi_{\vec{k}}^+ | \phi_{\vec{k}'}^+ \rangle - 2\pi i\delta(E_k - E_{k'}) \langle \vec{k} | V | \phi_{\vec{k}'}^+ \rangle.$$

Tendo em conta a ortonormalidade dos autovetores $|\phi_{\vec{k}}^+\rangle$ de H e a relação (10.36) resulta finalmente

$$\langle \vec{k} | S | \vec{k}' \rangle = \delta(\vec{k} - \vec{k}') - 2\pi i \delta(E_k - E_{k'}) \langle \vec{k} | T^+(k') | \vec{k}' \rangle. \quad (10.42)$$

A presença do fator $\delta(E_k - E_{k'})$ mostra explicitamente que os elementos de matriz de S são definidos pelos elementos de matriz de T^+ na *superfície de energia* apenas.

A propriedade, geralmente admitida, de completeza dos estados de espalhamento com condições de contorno de ondas emergentes $|\phi_k^+\rangle$ (acrescidos, quando for o caso, dos estados ligados $|\phi_i\rangle$ de H), no sentido que existe a resolução da unidade

$$\hat{1} = \int d^3k |\phi_k^+\rangle \langle \phi_k^+| + \sum_i |\phi_i\rangle \langle \phi_i|,$$

bem como da afirmação correspondente para os estados com condições de contorno de ondas imergentes $|\phi_k^-\rangle$, que leva à resolução alternativa da unidade

$$\hat{1} = \int d^3k |\phi_k^-\rangle \langle \phi_k^-| + \sum_i |\phi_i\rangle \langle \phi_i|,$$

implica em que o operador de espalhamento S é um operador *unitário*. Nessas duas resoluções da unidade os estados ligados $|\phi_i\rangle$ são estados ortonormais

$$\langle \phi_i | \phi_j \rangle = \delta_{ij}$$

e ortogonais a qualquer dos estados do espectro contínuo

$$\langle \phi_i | \phi_k^+ \rangle = \langle \phi_i | \phi_k^- \rangle = 0,$$

para os quais valem, por outro lado, as relações de ortogonalidade e normalização

$$\langle \phi_k^+ | \phi_{k'}^+ \rangle = \delta(\vec{k} - \vec{k}'), \quad \langle \phi_k^- | \phi_{k'}^- \rangle = \delta(\vec{k} - \vec{k}').$$

Para obter a unitariedade de S basta calcular, de fato,

$$\begin{aligned} \langle \vec{k} | S S^\dagger | \vec{k}' \rangle &= \int d^3k_1 \langle \vec{k} | S | \vec{k}_1 \rangle \langle \vec{k}_1 | S^\dagger | \vec{k}' \rangle = \int d^3k_1 \langle \vec{k} | S | \vec{k}_1 \rangle \langle \vec{k}' | S | \vec{k}_1 \rangle^* \\ &= \int d^3k_1 \langle \phi_k^- | \phi_{k_1}^+ \rangle \langle \phi_{k'}^- | \phi_{k_1}^+ \rangle^* = \int d^3k_1 \langle \phi_k^- | \phi_{k_1}^+ \rangle \langle \phi_{k_1}^+ | \phi_{k'}^- \rangle \\ &= \langle \phi_k^- | \hat{1} | \phi_{k'}^- \rangle = \delta(\vec{k} - \vec{k}'). \end{aligned}$$

A substituição

$$\int d^3 k_1 |\phi_{k_1}^{\pm}\rangle \langle \phi_{k_1}^{\pm}| \rightarrow \hat{1}$$

é possível devido à ortogonalidade dos estados ligados aos estados do espectro contínuo, e corresponde à chamada “completeza assintótica” dos estados de espalhamento $|\phi_k^{\pm}\rangle$. Essa denominação faz referência ao fato de que as funções de onda correspondentes aos estados ligados devem se aproximar de zero suficientemente depressa a grandes distâncias do potencial. De modo inteiramente análogo, e fazendo uso da “completeza assintótica” dos estados $|\phi_k^{\pm}\rangle$, resulta também que

$$\langle \vec{k} | S^{\dagger} S | \vec{k}' \rangle = \delta(\vec{k} - \vec{k}')$$

de modo que $SS^{\dagger} = S^{\dagger}S = \hat{1}$ e portanto $S^{\dagger} = S^{-1}$.

A relação de unitariedade para o operador de espalhamento pode também ser expressa em termos do operador de transição $T^+(k)$ utilizando a relação (10.42). De fato

$$\begin{aligned} \delta(\vec{k} - \vec{k}') &= \int d^3 k_1 \langle \vec{k} | S^{\dagger} | \vec{k}_1 \rangle \langle \vec{k}_1 | S | \vec{k}' \rangle = \int d^3 k_1 \langle \vec{k}_1 | S | \vec{k} \rangle^* \langle \vec{k}_1 | S | \vec{k}' \rangle = \\ &= \int d^3 k_1 \left[\delta(\vec{k} - \vec{k}_1) + 2\pi i \delta(E_k - E_{k_1}) \langle \vec{k}_1 | T^+(k) | \vec{k} \rangle^* \right] \times \\ &\quad \times \left[\delta(\vec{k}_1 - \vec{k}') - 2\pi i \delta(E_{k_1} - E_{k'}) \langle \vec{k}_1 | T^+(k') | \vec{k}' \rangle \right] = \\ &= \delta(\vec{k} - \vec{k}') - 2\pi i \delta(E_k - E_{k'}) \left[\langle \vec{k} | T^+(k') | \vec{k}' \rangle - \langle \vec{k}' | T^+(k) | \vec{k} \rangle^* \right] + \\ &\quad + (2\pi)^2 \int d^3 k_1 \delta(E_k - E_{k_1}) \delta(E_{k_1} - E_{k'}) \langle \vec{k}_1 | T^+(k) | \vec{k} \rangle^* \langle \vec{k}_1 | T^+(k') | \vec{k}' \rangle. \end{aligned}$$

Nessas expressões o operador de transição é calculado sempre na superfície de energia. O último termo pode ser calculado fazendo uso da identidade

$$\delta(E_k - E_{k_1}) \delta(E_{k_1} - E_{k'}) \equiv \delta(E_k - E_{k'}) \delta(E_{k_1} - E_{k'})$$

o que leva finalmente a

$$i \left[\langle \vec{k} | T^+(k') | \vec{k}' \rangle - \langle \vec{k}' | T^+(k) | \vec{k} \rangle^* \right] = 2\pi \int d^3 k_1 \delta(E_{k_1} - E_{k'}) \langle \vec{k}_1 | T^+(k) | \vec{k} \rangle^* \langle \vec{k}_1 | T^+(k') | \vec{k}' \rangle \quad (10.43)$$

que é a relação de unitariedade do operador de espalhamento expressa em termos do operador de transição. No caso particular em que $\vec{k} = \vec{k}'$, essa expressão se reduz a

$$-2 \operatorname{Im} \langle \vec{k} | T^+(k) | \vec{k}' \rangle = 2\pi \int d^3 k_1 \delta(E_{k_1} - E_k) \left| \langle \vec{k}_1 | T^+(k) | \vec{k} \rangle \right|^2$$

cujo significado pode ser identificado utilizando a relação (10.33) entre os elementos de matriz do operador de transição na superfície de energia e a amplitude de espalhamento. Segundo essa relação, de fato

$$\langle \vec{k} | T^+(k) | \vec{k} \rangle = -\frac{\hbar^2}{(2\pi)^2 M} f_k^{(+)}(0)$$

e também

$$\left| \langle \vec{k}_1 | T^+(k) | \vec{k} \rangle \right|^2 = \left[\frac{\hbar^2}{(2\pi)^2 M} \right]^2 \left| f_k^{(+)}(\hat{k}_1) \right|^2$$

onde $f_k^{(+)}(0)$ e $\left| f_k^{(+)}(\hat{k}_1) \right|^2$ são respectivamente a amplitude de espalhamento na direção frontal e a seção de choque diferencial na direção \hat{k}_1 . O caso particular $\vec{k} = \vec{k}'$ da relação (10.43) corresponde portanto ao teorema óptico

$$2 \operatorname{Im} f_k^{(+)}(0) = \frac{\hbar^2}{2\pi M} \int d^3 k_1 \delta(E_{k_1} - E_k) \left| f_k^{(+)}(\hat{k}_1) \right|^2 = \frac{k}{2\pi} \sigma_{\text{tot}}.$$

Na seção 10.2.3 o teorema óptico foi relacionado à conservação de norma no problema do espalhamento de um pacote de ondas, e este resultado permite então relacionar essa conservação de norma à completeza assintótica dos estados de espalhamento e à conseqüente unitariedade do operador S .

10.6 Diagonalização do operador de espalhamento para potenciais invariantes sob rotações. Ondas parciais e defasagens.

Até aqui tanto o operador de espalhamento S como o operador de transição $T(k)$ foram tratados na representação de momentos, e.g. $\langle \vec{k} | S | \vec{k}' \rangle$, na qual eles não são diagonais. A utilidade da representação de momentos está em que ela permite um contacto direto com a parametrização usual da amplitude de espalhamento em termos do módulo e da orientação relativa dos momentos assintóticos inicial e final (ou seja, em termos da energia incidente e do ângulo de

espalhamento). Outras representações para esses operadores são também úteis e convenientes para certas aplicações, especialmente a representação que diagonaliza o operador de espalhamento S . Sendo esse um operador unitário, seus autovalores são números complexos de módulo 1 (fatores de fase), de modo que a equação de autovalores pode ser escrita como

$$S|\varphi_\alpha\rangle = e^{2i\delta_\alpha}|\varphi_\alpha\rangle.$$

O ângulo δ_α é usualmente chamado *defasagem característica* associada ao *canal característico* $|\varphi_\alpha\rangle$. Na representação dos canais característicos o operador S é portanto diagonal, e caracterizado pelas defasagens características para cada um dos canais característicos

$$\langle\varphi_\alpha|S|\varphi_\beta\rangle = e^{2i\delta_\alpha}\delta_{\alpha\beta}.$$

O objetivo desta seção é explicitar a representação dos canais característicos no caso do espalhamento por um potencial esfericamente simétrico de curto alcance. Como será visto, eles se reduzem neste caso às *ondas parciais* com momento angular bem definido tratadas na seção 5.1.5, enquanto as defasagens características se reduzem às defasagens correspondentes às ondas parciais. Isso servirá para estabelecer uma conexão entre o desenvolvimento formal envolvendo o operador S e o tratamento anterior, em termos da separação de variáveis em coordenadas esféricas da versão diferencial da equação de Schrödinger estacionária, associada ao problema de espalhamento por um potencial esfericamente simétrico.

10.6.1 Ingredientes cinemáticos.

O hamiltoniano que descreve o processo de espalhamento será tomado como sendo

$$H = \frac{p^2}{2M} + V$$

onde, por simplicidade, o potencial V será tomado como sendo independente de spins. Desse modo, a simetria esférica se exprime pela comutatividade de V com o operador \vec{L} correspondente ao momento angular orbital, isto é,

$$[V, \vec{L}] = 0$$

donde segue também que

$$[S, \vec{L}] = 0 \quad \text{e} \quad [T(k), \vec{L}] = 0$$

pois esses operadores são definidos em termos de V e de outros operadores invariantes por rotações, como a energia cinética $p^2/2M$, apenas. Isso significa que S e $T(k)$ podem ser diagonalizados simultaneamente com os operadores L^2 e L_3 , que são também compatíveis entre si. Desse modo, uma base conveniente que pode ser considerada em substituição à base de autovetores do momento $\{|\vec{k}\rangle\}$, é a “base esférica”, formada pelos autovetores simultâneos dos operadores L^2 , L_3 e p^2 , $\{|klm\rangle\}$, que satisfazem as equações de autovalores

$$\begin{aligned} L^2|klm\rangle &= \hbar^2 l(l+1)|klm\rangle \\ L_3|klm\rangle &= \hbar m|klm\rangle \\ p^2|klm\rangle &= \hbar^2 k^2|klm\rangle. \end{aligned} \quad \text{e}$$

As autofunções simultâneas desses três operadores, $\langle \vec{r}|klm\rangle$, já foram encontradas na expansão em ondas parciais de uma onda plana, no capítulo 5, eq. (5.18). Deixando por ora em aberto uma constante de normalização C_{kl} , a ser fixada convenientemente em seguida, elas são

$$\langle \vec{r}|klm\rangle = C_{kl} j_l(kr) Y_{lm}(\hat{r}).$$

Ao contrário do que acontece com L^2 e L_3 , o espectro de p^2 é contínuo. Um critério de normalização conveniente para essas autofunções consiste em escolher C_{kl} de modo que se tenha

$$\langle klm|k'l'm'\rangle = \delta_{ll'} \delta_{mm'} \delta(E - E') \quad (10.44)$$

com $E = \hbar^2 k^2/2M$ e $E' = \hbar^2 k'^2/2M$. Usando as propriedades de ortonormalidade das harmônicas esféricas $Y_{lm}(\hat{r})$, o produto escalar acima fica expresso como

$$\begin{aligned} \langle klm|k'l'm'\rangle &= C_{kl}^* C_{k'l'} \int d^3r j_l(kr) j_{l'}(k'r) Y_{lm}^*(\hat{r}) Y_{l'm'}(\hat{r}) = \\ &= C_{kl}^* C_{k'l'} \delta_{ll'} \delta_{mm'} \int_0^\infty dr r^2 j_l(kr) j_{l'}(k'r). \end{aligned}$$

Usando agora a relação

$$\frac{2}{\pi} \int_0^\infty dr r^2 j_l(kr) j_{l'}(k'r) = \frac{\delta(k - k')}{k^2}$$

resulta

$$\langle klm|k'l'm'\rangle = C_{kl}^* C_{k'l'} \delta_{ll'} \delta_{mm'} \frac{\pi}{2k^2} \delta(k - k')$$

o que mostra que as constantes de normalização devem ser escolhidas de forma que se tenha

$$|C_{kl}|^2 \frac{\pi}{2k^2} \delta(k - k') = \delta(E - E') = \frac{dk}{dE} \delta(k - k').$$

Portanto

$$|C_{kl}|^2 = \frac{2k^2}{\pi} \frac{dk}{dE} = \frac{2k^2}{\pi} \frac{M}{\hbar^2 k} = \frac{2kM}{\pi \hbar^2}.$$

Adotando a escolha mais simples de fases, as autofunções simultâneas de L^2 , L_3 e p^2 com a normalização (10.44) são, explicitamente,

$$\langle \vec{r} | klm \rangle = \sqrt{\frac{2Mk}{\pi \hbar^2}} j_l(kr) Y_{lm}(\hat{r}).$$

A resolução da unidade em termos dos autovetores $|klm\rangle$ com essa normalização é então escrita como

$$\hat{1} = \sum_{lm} \int dE |klm\rangle \langle klm|.$$

Essa relação pode ser transcrita para a representação de Schrödinger fazendo uso da forma explícita das funções de onda. Nesse caso ela assume a forma conhecida da representação da função delta de Dirac em termos de harmônicas esféricas e de funções de Bessel esféricas

$$\begin{aligned} \delta(\vec{r} - \vec{r}') &= \sum_{lm} \int dE \langle \vec{r} | klm \rangle \langle klm | \vec{r}' \rangle = \\ &= \frac{2M}{\pi \hbar^2} \sum_{lm} \int dE j_l(kr) j_l(kr') Y_{lm}(\hat{r}) Y_{lm}^*(\hat{r}') = \\ &= \frac{2}{\pi} \sum_{lm} \int_0^\infty dk k^2 j_l(kr) j_l(kr') Y_{lm}(\hat{r}) Y_{lm}^*(\hat{r}') = \frac{\delta(r - r')}{r^2} \delta(\hat{r} - \hat{r}') \end{aligned}$$

As funções de onda dos estados da base esférica *no espaço de momentos*, $\langle \vec{k}' | klm \rangle$, também são úteis e podem ser calculadas facilmente a partir dos resultados anteriores e da expansão da onda plana em ondas parciais, eq. (5.18). De fato

$$\langle \vec{k}' | klm \rangle = \int d^3r \langle \vec{k}' | \vec{r} \rangle \langle \vec{r} | klm \rangle = \frac{1}{(2\pi)^{3/2}} C_{kl} \int d^3r e^{-i\vec{k}' \cdot \vec{r}} j_l(kr) Y_{lm}(\hat{r})$$

que, usando a expansão

$$e^{i\vec{k}\cdot\vec{r}} = 4\pi \sum_{lm} i^l j_l(kr) Y_{lm}(\hat{r}) Y_{lm}^*(\hat{k})$$

e também as relações de ortonormalidade das harmônicas esféricas bem como a integral encontrada anteriormente sobre o produto de duas funções de Bessel esféricas, leva a

$$\langle \vec{k}' | klm \rangle = (-i)^l \sqrt{\frac{2}{\pi}} C_{kl} \frac{\pi}{2k^2} \delta(k - k') Y_{lm}(\hat{k}') = (-i)^l \sqrt{\frac{\hbar^2}{Mk}} \delta(E - E') Y_{lm}(\hat{k}').$$

Esse resultado pode ser controlado verificando a expressão

$$\sum_{lm} \int dE \langle \vec{k}' | klm \rangle \langle klm | \vec{k}'' \rangle = \delta(\vec{k}' - \vec{k}''),$$

que é a resolução da unidade expressa na representação de momentos.

Cabe notar aqui que é possível, e mesmo freqüente, usar outros critérios de normalização para os estados da base esférica que não aquele expresso na condição (10.44). Notadamente, é comum o uso do critério alternativo

$$\langle klm | k'l'm' \rangle = \delta_{ll'} \delta_{mm'} \delta(k - k') \quad (10.45)$$

que leva à constante de normalização $C'_{kl} = \sqrt{\frac{2}{\pi}} k$. Devido à presença de funções delta de Dirac nos diversos critérios de normalização envolvendo o espectro contínuo, diferentes critérios levam em geral a estados $|klm\rangle$ com propriedades *dimensionais* também diferentes. Em particular, os estados normalizados segundo o critério (10.44) têm dimensão $[E]^{-1/2}$, ao passo que a dimensão dos estados normalizados segundo o critério (10.45) é $[L]^{-1/2}$.

10.6.2 Transformação de S e $T(k)$ para a base esférica.

Os elementos de matriz do operador S na base esférica podem ser obtidos a partir da representação desse operador no espaço de momentos usando como elementos de transformação as funções de onda dos vetores da base esférica no espaço de momentos. De fato

$$\langle klm | S | k'l'm' \rangle = \int d^3k_1 \int d^3k_2 \langle klm | \vec{k}_1 \rangle \langle \vec{k}_1 | S | \vec{k}_2 \rangle \langle \vec{k}_2 | k'l'm' \rangle.$$

A relação geral entre os elementos de matriz de S e de $T(k)$ no espaço de momentos, eq. (10.42), no entanto, pode ser escrita como

$$\langle \vec{k}_1 | (S - \hat{1}) | \vec{k}_2 \rangle = -2\pi i \delta(E_1 - E_2) \langle \vec{k}_1 | T(k_2) | \vec{k}_2 \rangle \quad (10.46)$$

que, em vista do fator $\delta(E_1 - E_2)$, mostra explicitamente que $\langle \vec{k}_1 | (S - \hat{1}) | \vec{k}_2 \rangle$, e portanto também $\langle \vec{k}_1 | S | \vec{k}_2 \rangle$, se anulam para $|\vec{k}_1| \neq |\vec{k}_2|$. Em outras palavras, o operador S é diagonal no módulo dos momentos, e como é também diagonal nos números quânticos associados ao momento angular (como consequência da invariância rotacional do potencial V), é diagonal na base esférica. Os vetores dessa base são portanto os canais característicos, neste caso. Em vista da unitariedade de S é possível escrever então, usando a normalização (10.44) para os vetores da base esférica,

$$\langle klm | S | k'l'm' \rangle = e^{2i\delta_{lm}(k)} \delta_{ll'} \delta_{mm'} \delta(E - E').$$

Na realidade, a invariância rotacional de S , expressa pela relação de comutação $[S, \vec{L}] = 0$, identifica esse operador como um tensor de Racah de ordem zero, o que permite utilizar o teorema de Wigner-Eckart (cf. eq. (9.21) para escrever

$$\langle klm | S | k'l'm' \rangle = \frac{1}{\sqrt{2l+1}} C_{m'0m}^{l'0l} \langle kl || S || k'l' \rangle.$$

O coeficiente de Clebsch-Gordan que aparece nessa expressão é simplesmente $\delta_{ll'} \delta_{mm'}$, e portanto o elemento de matriz não nulo é *independente de m* . Isso implica que a defasagem característica $\delta_{lm}(k)$ é na realidade independente de m , e pode portanto ser escrita simplesmente como $\delta_l(k)$. Desse modo a relação (10.46), transformada para a base esférica, fica reduzida a

$$\langle klm | (S - \hat{1}) | k'l'm' \rangle = (e^{2i\delta_l(k)} - 1) \delta_{ll'} \delta_{mm'} \delta(E - E') = -2\pi i \delta(E - E') \langle klm | T(k') | k'l'm' \rangle$$

o que significa que

$$\langle klm | T(k') | k'l'm' \rangle = T_l(k) \delta_{ll'} \delta_{mm'}$$

com

$$T_l(k) = -\frac{e^{2i\delta_l(k)} - 1}{2\pi i} = -\frac{1}{\pi} e^{i\delta_l(k)} \sin \delta_l(k).$$

10.6.3 Reconstrução da amplitude de espalhamento e da forma assintótica da função de onda.

A amplitude de espalhamento $f_k^{(+)}(\hat{r})$ é dada em termos dos elementos de matriz do operador $T(k)$ na base de momentos e na superfície de energia através da expressão (10.33). Essa

expressão pode ser relacionada com os elementos de matriz $T_l(k)$ através de uma transformação simples para a base esférica usando mais uma vez as funções de onda dos vetores dessa base no espaço de momentos, o que dá

$$\begin{aligned}
f_k^{(+)}(\hat{r}) &= -\frac{4\pi^2 M}{\hbar^2} \sum_{lm} \sum_{l'm'} \int dE_1 \int dE_2 \langle k\hat{r}|k_1 l m\rangle \langle k_1 l m|T(k)|k_2 l' m'\rangle \langle k_2 l' m'|\vec{k}\rangle = \\
&= -\frac{4\pi^2 M}{\hbar^2} \sum_{lm} \sum_{l'm'} \int dE_1 \int dE_2 \frac{\hbar^2}{M\sqrt{k_1 k_2}} \delta(E_1 - E) \delta(E_2 - E) Y_{lm}(\hat{r}) Y_{l'm'}^*(\hat{k}) \times \\
&\quad \times \delta_{ll'} \delta_{mm'} \langle k_1 l m|T(k)|k_2 l m\rangle = \\
&= -\frac{4\pi^2 M}{\hbar^2} \sum_{lm} \frac{\hbar^2}{Mk} Y_{lm}(\hat{r}) Y_{lm}^*(\hat{k}) T_l(k) = -\frac{4\pi^2}{k} \sum_l T_l(k) \sum_m Y_{lm}(\hat{r}) Y_{lm}^*(\hat{k}).
\end{aligned}$$

A soma sobre m pode ser feita usando o chamado teorema da adição para as harmônicas esféricas

$$\sum_m Y_{lm}(\hat{r}) Y_{lm}^*(\hat{k}) = \frac{2l+1}{4\pi} P_l(\hat{r} \cdot \hat{k})$$

onde $\hat{r} \cdot \hat{k} \equiv \cos \theta$, sendo θ o ângulo de espalhamento. Introduzindo a expressão dos $T_l(k)$ em termos das defasagens $\delta_l(k)$ resulta finalmente

$$f_k^{(+)}(\theta) = \frac{1}{k} \sum_l (2l+1) e^{i\delta_l(k)} \sin \delta_l(k) P_l(\cos \theta)$$

que reproduz a expressão (5.25), obtida no capítulo 5 a partir da forma assintótica da função de onda $\phi_k^+(\vec{r})$. Essa forma assintótica pode ser por sua vez reconstruída a partir de

$$\phi_k^+(\vec{r}) \longrightarrow e^{i\vec{k} \cdot \vec{r}} + f_k^{(+)}(\theta) \frac{e^{ikr}}{r}, \quad r \rightarrow \infty$$

e utilizando a expansão da onda plana em ondas parciais juntamente com a expressão de $f_k^{(+)}(\theta)$ em termos das defasagens características, o que leva a

$$\phi_k^+(\vec{r}) \longrightarrow \sum_l i^l (2l+1) e^{i\delta_l(k)} \frac{\sin(kr - l\pi/2 + \delta_l(k))}{kr} P_l(\cos \theta).$$

Se a função de onda $\phi_k^+(\vec{r})$ é decomposta em ondas parciais da forma usual, como

$$\phi_k^+(\vec{r}) = \sum_l i^l (2l+1) R_l^+(r) P_l(\cos \theta)$$

o comportamento assintótico das funções radiais $R_l^+(r)$ é dado *em termos das defasagens características* $\delta_l(k)$ como

$$R_l^+(r) \longrightarrow e^{i\delta_l(k)} \frac{\sin(kr - l\pi/2 + \delta_l(k))}{kr} = -\frac{1}{2ikr} \left(e^{-i(kr - \frac{l\pi}{2})} - e^{2i\delta_l(k)} e^{i(kr - \frac{l\pi}{2})} \right) \quad (10.47)$$

mostrando que as defasagens características coincidem com as obtidas anteriormente através da solução das equações diferenciais para cada uma das ondas parciais. O fato de que as diferentes ondas parciais satisfazem equações diferenciais *independentes*, em particular, é correlato ao caráter diagonal do operador S na base esférica, e essas duas características decorrem, por sua vez, da invariança rotacional do potencial V .

Em termos práticos, a forma de obter o operador S para um potencial rotacionalmente invariante V *dado* é resolver a equação de Schrödinger para o espectro contínuo com condições de contorno de espalhamento, extraíndo as defasagens do comportamento assintótico da solução. Isso expõe também o fato de que o potencial determina o operador S , e portanto os observáveis assintóticos associados ao processo de espalhamento. Por outro lado, a partir do conhecimento desses observáveis (associados em última análise às defasagens $\delta_l(k)$) é possível, em geral e em princípio, construir o operador S e o comportamento assintótico das funções de onda de espalhamento, mas não o potencial, que depende de informação adicional, referente notadamente ao espectro discreto.

Problemas.

10.1 - Considere o espalhamento de ondas s ($L = 0$) pelo potencial central (que corresponde a uma camada repulsiva muito fina na superfície de uma esfera de raio a)

$$V(r) = \frac{\hbar^2 \lambda}{2M} \delta(r - a)$$

onde λ é uma constante com dimensão de inverso de comprimento, a função $\delta(r - a)$ é a função de Dirac, cujo argumento envolve a coordenada $r \equiv |\vec{r}|$ e $a > 0$. A equação de Schrödinger estacionária a ser resolvida é portanto

$$\left(\frac{d^2}{dr^2} + k^2 \right) u(r) = \lambda \delta(r - a) u(r)$$

com as condições de contorno $u(0) = 0$ e, para $r \rightarrow \infty$ (a rigor neste caso simples basta $r > a$),

$$u(r > a) = \sin kr + f_k^{(0)} e^{ikr}$$

(certifique-se disto em caso de qualquer dúvida!).

a) Verifique que a função

$$g_k(r, r') = -\frac{1}{k} \sin kr_{<} e^{ikr_{>}}$$

onde $r_{<}$ e $r_{>}$ correspondem respectivamente ao menor e ao maior dos dois argumentos r, r' satisfaz a equação

$$\left(\frac{d^2}{dr^2} + k^2 \right) g_k(r, r') = \delta(r - r')$$

e é portanto uma função de Green da equação de Schrödinger a ser resolvida. **Sugestão:** se $u_0(r)$ e $v_0(r)$ são duas soluções da equação homogênea $(d^2/dr^2 + k^2)w(r) = 0$, o determinante da matriz 2×2 cujas linhas são respectivamente $\{u_0(r), v_0(r)\}$ e $\{du_0(r)/dr, dv_0(r)/dr\}$ é chamado Wronskiano $W(r)$ dessas duas soluções; verifique em primeiro lugar que para essa particular equação o Wronskiano, visto como função de r , na realidade não depende de r . Verifique então que os fatores de $g(r, r')$ correspondem a duas soluções da equação homogênea com Wronskiano diferente de zero; e em seguida que essa função, vista como função de r com parâmetro r' , satisfaz a equação homogênea sempre que $r \neq r'$, e que para $r = r'$ há uma descontinuidade na derivada com valor 1, independentemente do valor de r' , devido à propriedade de que o Wronskiano não depende de r .

b) Use a função de Green $g_k(r, r')$ para reescrever a equação de Schrödinger para a onda s como uma equação de Lippmann-Schwinger, usando $\sin kr$ como termo inhomogêneo. Verifique que a solução dessa equação satisfaz as condições de contorno para $r = 0$ e $r > a$ (por exemplo, estudando o comportamento da função de Green na região apropriada) e mostre que a amplitude $f_k^{(0)}$ da onda espalhada admite a representação formal

$$f_k^{(0)} = -\frac{\lambda}{k} \sin ka u(a)$$

em termos da solução $u(r)$ no ponto $r = a$.

c) Resolva completamente a equação de Lippmann-Schwinger. **Sugestão:** note que a integração pode ser feita usando as propriedades da função delta, dando $u(r)$ em termos do termo inhomogêneo e de $u(a)$. O resultado permite exprimir $u(a)$ em termos do termo inhomogêneo calculado em $r = a$, o que fornece a solução da equação.

d) Use a solução da equação de Lippmann-Schwinger para obter $f_k^{(0)}$ substituindo o fator $u(a)$ que aparece na relação formal citada no item **b)**. Mostre em seguida que a *solução* na região assintótica $u(r > a)$ pode ser posta sob as formas alternativas (cf. eq. (10.47))

$$u(r > a) = \text{const.} \times (e^{-ikr} - S_0 e^{ikr}) = \text{const.}' \times \sin(kr + \delta_0)$$

onde $S_0 \equiv e^{2i\delta_0}$ é um fator de fase (a matriz S para a onda s) e δ_0 é a defasagem da onda s . Escreva expressões para S_0 e δ_0 como funções de k . Mostre ainda que a amplitude $f_k^{(0)}$ é dada em termos de δ_0 como

$$f_k^{(0)} = e^{i\delta_0} \sin \delta_0.$$

10.2 - O efeito de uma ressonância na onda parcial $l = 0$ pode ser representado por um potencial não local, separável, dependente da energia, da forma

$$V_R(\vec{r}, \vec{r}') = \frac{v(|\vec{r}|)v^*(|\vec{r}'|)}{E - E_R}$$

onde E é a energia do sistema, $E_R > 0$ é um parâmetro com dimensão de energia que determina a energia em que a ressonância ocorre e $v(r)$ é um “fator de forma” com dimensão de energia tal que

$$\int d^3r |v(r)| < \infty.$$

Para verificar essas afirmações considere o problema de espalhamento associado ao hamiltoniano

$$H = \frac{p^2}{2M} + V_R \quad \text{isto é} \quad H \phi_k^\pm = E \phi_k^\pm, \quad E = \frac{\hbar^2 k^2}{2M}.$$

a) Use a expansão de ϕ_k^\pm em ondas parciais para mostrar que apenas a componente $l = 0$ é afetada pelo potencial V_R . **Sugestão:** a ação do potencial não local sobre a função de onda é dada por

$$(V_R \phi_k^\pm)(\vec{r}) = \int d^3r' V_R(\vec{r}, \vec{r}') \phi_k^\pm(\vec{r}').$$

Em vista disso, qual é o valor das defasagens $\delta_l(k)$ para $l \neq 0$?

b) Escreva a equação diferencial para a onda parcial $l = 0$ e use a função de Green $g_k(r, r')$ do problema anterior para convertê-la numa equação integral do tipo da equação de Lippmann-Schwinger.

c) Resolva essa equação utilizando o caráter separável do potencial V_R para calcular a quantidade $\int d^3r v^*(r)u_0(r)$, onde $u_0(r)$ é a amplitude radial correspondente à onda parcial $l = 0$.

d) A partir da solução obtenha a defasagem $\delta_0(k)$ para a onda parcial $l = 0$. Verifique que ela aumenta rapidamente com a energia nas vizinhanças de $E = E_R$. Calcule a seção de choque diferencial como função da energia nessa região de energia.

10.3 - Este problema é uma versão “formal” do problema anterior. Considere o problema de espalhamento de partículas de massa M por um potencial *separável*, dependente da energia E , da forma

$$V \equiv \frac{v|u\rangle\langle u|v}{E - E_R}$$

onde $|u\rangle$ é um vetor de estado normalizado ($\langle u|u\rangle = 1$) e E_R é uma constante real. A equação de Schrödinger estacionária para os estados de espalhamento é portanto

$$\left[E - H_0 - \frac{v|u\rangle\langle u|v}{E - E_R} \right] |\phi_k^+\rangle = 0$$

onde $H_0 = p^2/2M$ é o hamiltoniano livre.

a) Escreva a equação de Lippmann-Schwinger para esse problema em termos do operador de Green

$$G_E = \frac{1}{E + i\epsilon - H_0}.$$

b) Utilizando a forma especial do potencial, mostre que vale a relação

$$\langle u|v|\psi_k^+\rangle = (2\pi)^{3/2} \frac{E - E_R}{E - E_R - \langle u|v \frac{1}{E + i\epsilon - H_0} v|u\rangle} \langle u|v|\vec{k}\rangle.$$

c) Use esta última expressão para obter a *solução* formal da equação de Lippmann-Schwinger

$$|\phi_k^+\rangle = (2\pi)^{3/2} \left[|\vec{k}\rangle + \frac{1}{E + i\epsilon - H_0} \frac{v|u\rangle\langle u|v|\vec{k}\rangle}{E - E_R - \langle u|v \frac{1}{E + i\epsilon - H_0} v|u\rangle} \right]$$

onde $|\vec{k}\rangle$ é um autovetor do momento \vec{p} e portanto também do hamiltoniano livre H_0 .

d) Usando a expressão formal

$$f_k^+(\hat{k}') = (2\pi)^{3/2} \langle \vec{k}' | V | \psi_k^+ \rangle = (2\pi)^{3/2} \langle \vec{k}' | \frac{v|u\rangle \langle u|v}{E - E_R} | \phi_k^+ \rangle$$

para a amplitude de espalhamento mostre ainda que

$$f_k^+(\hat{k}') = (2\pi)^3 \frac{\langle \vec{k}' | v | u \rangle \langle u | v | \vec{k} \rangle}{E - E_R - \langle u | v \frac{1}{E + i\epsilon - H_0} v | u \rangle}.$$

e) Verifique que se $\langle \vec{r} | u \rangle = u(r)$, independente dos ângulos de \vec{r} , e se o operador v é invariante por rotações, então a seção de choque

$$\frac{d\sigma}{d\Omega} = |f_k^+(\hat{k}')|^2$$

é independente de ângulos. **Sugestão:** a possível dependência angular provém do numerador da amplitude de espalhamento (certifique-se disto!); escreva os elementos de matriz que aparecem nesse denominador como (por exemplo)

$$\langle u | v | \vec{k} \rangle = \int d^3r \int d^3r' \langle u | \vec{r} \rangle \langle \vec{r} | v | \vec{r}' \rangle \langle \vec{r}' | \vec{k} \rangle = \int d^3r \int d^3r' u^*(r) \langle \vec{r} | v | \vec{r}' \rangle \frac{1}{(2\pi)^{3/2}} e^{i\vec{k} \cdot \vec{r}'}$$

notando que a invariância rotacional de v implica que $\langle \vec{r} | v | \vec{r}' \rangle$ só pode depender de r , r' e de $\hat{r} \cdot \hat{r}'$ onde \hat{r} e \hat{r}' são vetores de módulo 1 na direção de \vec{r} e de \vec{r}' respectivamente.

Referências gerais.

Os textos gerais sobre mecânica quântica geralmente trazem capítulos específicos sobre a teoria do espalhamento, especialmente espalhamento por potenciais. Um tratamento mais extenso e detalhado que a média se encontra em

★ **A. Messiah**, *op. cit.*, Capítulos X e XIX.

O primeiro desses capítulos inclui uma discussão dependente do tempo do processo de espalhamento usando pacotes de onda, enquanto o segundo trata da teoria formal. Outros textos são, por exemplo

★ **L. I. Schiff**, *Quantum Mechanics* (Third Edition), McGraw Hill Book Co. 1968, Capítulo 9.

★ **C. Cohen-Tanoudji, B. Diu e F. Laloë**, *op. cit.*, Capítulo VIII (trata-se, neste caso, de uma apresentação breve, na escala de dois volumes do texto).

Existem dois textos substanciais e de certo modo “clássicos”, dedicados inteiramente à teoria do espalhamento. Eles são

★ **R. G. Newton**, *Scattering of Waves and Particles*, McGraw-Hill Book Co., 1966.

★ **M. L. Goldberger e K. M. Watson**, *Collision Theory*, John Wiley, 1967 (third printing).

Outros textos tratando especialmente das propriedades de analiticidade dos operadores de transição e de espalhamento e de suas conseqüências incluem

★ **V. de Alfaro e T. Regge**, *Potential Scattering*, North-Holland Publ. Co., 1965.

★ **H. M. Nussenzveig**, *Causality and Dispersion Relations*, Academic Press, 1972.

Capítulo 11

Equações relativísticas.

Depois do trabalho de Einstein de 1905 que, através da teoria da relatividade restrita, estabeleceu a estrutura cinemática básica do espaço-tempo associando mudanças de referencial inercial a transformações de Lorentz, um requisito mínimo para uma “boa teoria” passou a ser o de que seus resultados físicos não dependam do particular referencial que tenha sido usado para obtê-los, tendo em conta que as coordenadas relativas a diferentes referenciais inerciais se relacionam entre si através de tais transformações. A mecânica quântica “não relativística”, tal como formulada por Heisenberg, Schrödinger e Dirac, não satisfaz esse requisito por razões bastante profundas. De fato, posições e momentos são realizados nessa teoria em termos de operadores agindo sobre um espaço linear, complexo, de amplitudes de probabilidade, enquanto o tempo é tratado apenas como um parâmetro “externo”, associado à evolução dinâmica das amplitudes de probabilidade. A diferença de “status” de variáveis espaciais e da variável temporal é na realidade irreconciliável com o princípio da relatividade restrita. Uma “boa” teoria quântica deve tratar variáveis espaciais e temporais de forma mais simétrica, o que é conseguido notadamente no contexto de uma teoria de campos, como observado de passagem no capítulo 8, com relação à quantização do campo de radiação livre, tratando as variáveis espaciais também como parâmetros.

Existe, no entanto, um ponto de vista menos radical, dentro do qual é possível buscar modificações da mecânica quântica não relativística no sentido de aproximá-la dos requisitos da relatividade restrita. Ele consiste em observar que a equação de movimento dessa teoria, para uma partícula livre, por exemplo, pode ser escrita, na representação de Schrödinger, como uma equação a derivadas parciais que é de segunda ordem com relação às coordenadas espaciais mas de primeira ordem com relação ao tempo. Essa assimetria entre coordenadas espaciais e temporais é incompatível com os requisitos mínimos para o que poderia ser visto como uma “boa” equação de movimento, e leva diretamente à questão da possibilidade de substituí-la por outra que trate variáveis espaciais e temporais de forma mais simétrica. O

objetivo central deste capítulo é justamente o de formular equações de movimento que sejam compatíveis, nesse sentido, com a cinemática da relatividade restrita, e discutir propriedades de suas soluções. Em particular, é possível desse modo formular um modelo para o átomo de hidrogênio que incorpora automaticamente os efeitos relativísticos introduzidos apenas sob a forma de correções perturbativas nas seções 6.3 e 6.4.

É claro que as dificuldades mais básicas, associadas ao tratamento diferenciado das variáveis de posição por um lado, e do tempo por outro, não são resolvidas com a formulação de equações de movimento relativísticas, em substituição à equação de Schrödinger. Essa formulação, no entanto, é também relevante para o contexto mais geral, pois as equações relativisticamente bem formuladas podem eventualmente ser reinterpretadas como teorias de campo “clássicas” (isto é, pertencentes ao mesmo nível, por exemplo, que as equações de Maxwell para o campo eletromagnético) e podendo portanto ser, por sua vez, quantizadas, levando a uma formulação mais consistente em termos de teorias quânticas de campo.

11.1 Sumário da cinemática da relatividade restrita.

No contexto da relatividade restrita, um “evento”, localizado em um ponto do espaço especificado pelo vetor de posição $\vec{r} \equiv \{x^1, x^2, x^3\}$, onde os x^i , $i = 1, 2, 3$ são componentes cartesianas de \vec{r} , e ocorrendo no tempo t , é rotulado pelo objeto de quatro componentes $x \equiv \{x^\mu\}$, $\mu = 0, 1, 2, 3$, com $x^0 = ct$. Esse objeto contém as quatro componentes *contravariantes* (o que é indicado pela posição *superior* dos índices) de um *quadrivetor* sob transformações de Lorentz, que deixam invariante a quantidade

$$x \cdot x \equiv \sum_{\mu\nu} x^\mu g_{\mu\nu} x^\nu \equiv \sum_{\mu} x^\mu \sum_{\nu} g_{\mu\nu} x^\nu \equiv \sum_{\mu} x^\mu x_\mu \quad \text{invariante} \quad (11.1)$$

onde a métrica é definida pelo tensor

$$\{g_{\mu\nu}\} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

É comum, neste contexto, a omissão dos símbolos de soma em expressões como a (11.1), com o entendimento de que, salvo menção expressa em contrário, índices repetidos são sempre somados, e índices gregos variam sempre de 0 a 3. Essa convenção será portanto utilizada daqui por diante. Ela permite re-escrever a (11.1) simplesmente como

$$x \cdot x \equiv x^\mu g_{\mu\nu} x^\nu \equiv x^\mu x_\mu.$$

As componentes $x_\mu \equiv g_{\mu\nu}x^\nu$ são chamadas componentes *covariantes* do quadrivetor x . Tendo em vista a forma do tensor métrico $g_{\mu\nu}$, elas se relacionam com as componentes contravariantes através de

$$\{x_0, x_1, x_2, x_3\} \equiv \{x^0, -x^1, -x^2, -x^3\}$$

de modo que o invariante $x \cdot x$ é dado explicitamente como

$$x \cdot x = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = c^2t^2 - \vec{r} \cdot \vec{r}.$$

A versão contravariante do tensor métrico, $g^{\mu\nu}$, pode ser obtida de $x^\mu = g^{\mu\nu}x_\nu$ que, em vista do caráter diagonal desse objeto, leva simplesmente a $g^{\mu\nu} = g_{\mu\nu}$. Existe também uma versão “mista” do tensor métrico, dada por

$$g^\mu_\nu = g^{\mu\lambda}g_{\lambda\nu} = \delta^\mu_\nu$$

onde o último objeto é simplesmente um delta de Kronecker usual, igual a 1 se $\mu = \nu$ e igual a zero em outros casos.

As transformações de Lorentz Λ são transformações lineares, reais (para preservar o caráter real das coordenadas espacio-temporais) que mantêm a quantidade $x \cdot x$ invariante, isto é

$$x^\mu \xrightarrow{\Lambda} x'^\mu = \Lambda^\mu_\nu x^\nu, \quad x^\mu x_\mu = x'^\mu x'_\mu.$$

Uma consequência importante dessa condição pode ser obtida escrevendo

$$x'^\mu x'_\mu = x'^\mu g_{\mu\lambda} x'^\lambda = \Lambda^\mu_\nu x^\nu g_{\mu\lambda} \Lambda^\lambda_\sigma x^\sigma = g_{\nu\tau} \Lambda^{\mu\tau} x^\nu \Lambda_{\mu\sigma} x^\sigma$$

onde foram usadas as relações

$$\Lambda^\mu_\nu = g_{\nu\tau} \Lambda^{\mu\tau} \quad \text{e} \quad \Lambda_{\mu\sigma} = g_{\mu\lambda} \Lambda^\lambda_\sigma.$$

Para que $x'^\mu x'_\mu$ seja igual a $x^\nu g_{\nu\sigma} x^\sigma$ é preciso que se tenha, portanto,

$$\Lambda^{\mu\tau} \Lambda_{\mu\sigma} = \delta^\tau_\sigma. \quad (11.2)$$

Essa relação tem implicações diretas para o determinante da transformação $\{\Lambda^\mu_\nu\}$. De fato

$$\det \{\Lambda^{\mu\tau} \Lambda_{\mu\sigma}\} = \det \{(\Lambda^\dagger)^\tau{}^\mu \Lambda_{\mu\sigma}\} = \det \{(\Lambda^\dagger)^\tau{}_\mu \Lambda^\mu{}_\sigma\} = 1$$

donde segue que $\det \{\Lambda^\mu_\nu\} = \pm 1$, pois o determinante do produto de duas matrizes é o produto dos determinantes de cada uma delas e os determinantes de $\{(\Lambda^\dagger)^\mu{}_\nu\}$ e de $\{\Lambda^\mu{}_\nu\}$ são iguais.

As transformações de Lorentz formam um grupo (no sentido das definições dadas na seção 9.1) cujos elementos mantêm invariante a quantidade (11.1), ou seja, $x \cdot x \equiv c^2 t^2 - r^2$. A relação geral (11.2) fica, no caso particular em que $\sigma = \tau = 0$,

$$1 = \Lambda^{\mu 0} \Lambda_{\mu 0} = \Lambda^{\mu}_{0} g_{\mu\nu} \Lambda^{\nu}_0 = (\Lambda^0_0)^2 - (\Lambda^1_0)^2 - (\Lambda^2_0)^2 - (\Lambda^3_0)^2$$

donde segue que $(\Lambda^0_0)^2 \geq 1$. Isso permite classificar as transformações em dois tipos, que correspondem respectivamente a $\Lambda^0_0 \geq 1$ e a $\Lambda^0_0 \leq -1$. Para cada um desses dois tipos é possível ainda ter $\det \{\Lambda^{\mu}_{\nu}\} = \pm 1$, de modo que há a considerar na realidade quatro classes distintas de transformações, identificadas pelos sinais de Λ^0_0 e de $\det \{\Lambda^{\mu}_{\nu}\}$. As transformações do tipo $++$ são as chamadas *transformações próprias ortócronas*, e incluem a transformação *idêntica* $\Lambda^{\mu}_{\nu} = \delta^{\mu}_{\nu}$. Essas transformações constituem na realidade um sub-grupo do grupo de Lorentz completo. Os outros tres tipos podem ser obtidos das transformações próprias ortócronas através de uma *inversão temporal* $T : x'^0 = -x^0, x'^i = x^i, i = 1, 2, 3$ (que troca o sinal de tanto de Λ^0_0 como também o sinal do determinante, levando ao tipo $--$), de uma *inversão espacial* $P : x'^0 = x^0, x'^i = -x^i, i = 1, 2, 3$ (que troca o sinal do determinante sem mudar o sinal de Λ^0_0 , levando ao tipo $+-$) ou de ambas (o que troca o sinal de Λ^0_0 sem trocar o sinal do determinante, levando ao tipo $-+$). As transformações T e P estão associadas respectivamente à inversão temporal e à inversão espacial. No que segue, salvo menção expressa em contrário, as transformações de Lorentz que serão consideradas serão transformações próprias ortócronas.

A forma explícita de uma transformação de Lorentz própria e ortócrona, para um referencial que se move com velocidade v , na direção x^1 , com relação ao referencial de partida é

$$x'^0 = \frac{x^0 - \frac{v}{c} x^1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad x'^1 = \frac{x^1 - \frac{v}{c} x^0}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad x'^2 = x^2, \quad x'^3 = x^3$$

ou, em forma matricial

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \cosh \varphi & -\sinh \varphi & 0 & 0 \\ -\sinh \varphi & \cosh \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \quad (11.3)$$

com as definições

$$\cosh \varphi \equiv \gamma \equiv \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \sinh \varphi \equiv \beta \gamma \equiv \frac{v/c}{\sqrt{1 - \frac{v^2}{c^2}}}$$

de modo que $\tanh \varphi = \beta \equiv v/c$. Os elementos dessa matriz ligam as componentes x'^{μ} às componentes x^{ν} , e correspondem portanto aos Λ^{μ}_{ν}. O caráter próprio e ortócrono dessa transformação é evidente, pois $\Lambda^0_0 = \cosh \varphi \geq 1$ e $\det \{\Lambda^{\mu}_{\nu}\} = \cosh^2 \varphi - \sinh^2 \varphi = 1$. As matrizes

$\{\Lambda_{\mu\nu}\} \equiv \{g_{\mu\lambda}\Lambda^\lambda_\nu\}$ e $\{\Lambda^{\mu\nu}\} \equiv \{g^{\nu\lambda}\Lambda^\mu_\lambda\}$ para essas transformações particulares são então dadas por

$$\{\Lambda_{\mu\nu}\} = \{(\Lambda^\dagger)^{\mu\nu}\} = \begin{pmatrix} \cosh \varphi & -\sinh \varphi & 0 & 0 \\ \sinh \varphi & -\cosh \varphi & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

para as quais é imediato verificar explicitamente a validade da relação (11.2).

O caráter quadrivetorial de $\{x^\mu\}$ com relação às transformações de Lorentz implica em propriedades de transformação bem definidas também para uma série de outras quantidades diretamente relacionadas com esse objeto, notadamente:

a) a derivada parcial $\partial/\partial x^\mu$ se comporta como a componente μ de um quadrivetor *covariante*, enquanto $\partial/\partial x_\mu$ se comporta como a componente μ de um quadrivetor *contravariante*. Isso é usualmente expresso através das notações abreviadas

$$\frac{\partial}{\partial x^\mu} \equiv \partial_\mu \quad \text{e} \quad \frac{\partial}{\partial x_\mu} \equiv \partial^\mu$$

e pode ser verificado notando, por exemplo, que $\partial_\mu x^\mu = \partial^\mu x_\mu = 4$ são invariantes. A relação entre essas duas derivadas parciais é, por outro lado $\partial_\mu = g_{\mu\nu}\partial^\nu$.

b) Como consequência direta desse último ponto, o operador diferencial

$$\partial^\mu \partial_\mu = \partial_\mu \partial^\mu = \partial_\mu g^{\mu\nu} \partial_\nu = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2$$

é invariante sob transformações de Lorentz.

Objetos com n índices que se transformam com relação a cada um deles como vetores, e.g., para $n = 2$

$$A'^{\mu\nu} = \Lambda^\mu_\lambda \Lambda^\nu_\sigma A^{\lambda\sigma}$$

são *tensores* de ordem n (de segunda ordem, para o exemplo acima). Do ponto **a)** acima decorre portanto também que

c) os conjuntos de derivadas parciais segundas $\{\partial_\mu \partial_\nu\}$ e $\{\partial^\mu \partial^\nu\}$ se comportam respectivamente como tensores de segunda ordem. O primeiro desses conjuntos é um tensor *covariante* de segunda ordem, enquanto o segundo é *contravariante*. É possível considerar também o tensor de segunda ordem *misto* $\{\partial_\mu \partial^\nu\}$, que é covariante com relação ao índice μ e contravariante com relação ao índice ν . As componentes desses diferentes tipos de tensores se relacionam entre si também através do tensor métrico $g_{\mu\nu}$, por exemplo $\partial^\mu \partial^\nu = g^{\mu\lambda} \partial_\lambda \partial^\nu = g^{\mu\lambda} g^{\nu\sigma} \partial_\lambda \partial_\sigma$.

d) Se uma partícula (com massa de repouso M) se move em um dado referencial com velocidade \vec{v} , a sua posição \vec{r} varia, num intervalo de tempo infinitesimal dt , de $d\vec{r} = \vec{v}dt$. Como $dx^0 \equiv c dt$ e as componentes dx^i de $d\vec{r}$ se comportam sob transformações de Lorentz como componentes de um quadrivetor, a quantidade $dx^\mu dx_\mu$ é um invariante, e portanto também

$$d\tau^2 \equiv dt^2 - \frac{(d\vec{r})^2}{c^2} = dt^2 \left(1 - \frac{v^2}{c^2}\right).$$

O tempo τ definido dessa forma, que é um invariante de Lorentz, é chamado *tempo próprio* da partícula. A invariança de τ garante, por outro lado que

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \left\{ \frac{c}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right\}$$

se transforma como um quadrivetor (contravariante), chamado *quadrivelocidade* da partícula. É imediato verificar que as quatro componentes da quadrivelocidade não são independentes, pois $u^\mu u_\mu = c^2$. O produto Mu^μ da massa pela quadrivelocidade é, dada a invariança de M , também um quadrivetor (o quadrivetor momento), cujas componentes são

$$Mu^\mu = \left\{ \frac{Mc}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{M\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right\} \equiv \left\{ \frac{E}{c}, \vec{p} \right\} \quad (11.4)$$

onde E e \vec{p} são as expressões relativísticas para a energia e para o momento da partícula, respectivamente.

e) Uma transformação de Lorentz transforma o elemento de volume quadridimensional $d^4x = dx^0 dx^1 dx^2 dx^3$ se acordo com

$$d^4x' = dx'^0 dx'^1 dx'^2 dx'^3 = \frac{\partial(x'^0, x'^1, x'^2, x'^3)}{\partial(x^0, x^1, x^2, x^3)} d^4x$$

onde o jacobiano é o determinante da transformação de Lorentz

$$x'^\mu = \Lambda^\mu_\nu x^\nu,$$

o que identifica o elemento de volume como um invariante sob transformações de Lorentz próprias.

11.1.1 O eletromagnetismo de Maxwell como exemplo de “boa” teoria.

Um ponto de partida conveniente para formular o eletromagnetismo de Maxwell como uma “boa” teoria, isto é, como uma teoria corretamente adaptada ao princípio da relatividade restrita, consiste em considerar a carga elétrica total Q como um invariante, ou como uma quantidade escalar. Como ela se exprime em termos da densidade de carga $\rho(\vec{r}, t)$ através da integral

$$Q = \int d^3r \rho(\vec{r}, t),$$

segue que a densidade de carga $\rho(\vec{r}, t)$ deve se transformar como dx^0 , e a densidade de corrente $\vec{j}(\vec{r}, t) = \rho(\vec{r}, t)\vec{v}(\vec{r}, t)$ deve se transformar como as três componentes espaciais de um quadrivetor. Com os necessários ajustes dimensionais, isso leva a considerar o objeto

$$\{j^\mu\} \equiv \{c\rho, \vec{j}\} \quad (11.5)$$

como um quadrivetor (contravariante), chamado *quadrivetor densidade de corrente*. A conservação da carga total, expressa em forma diferencial pela equação de continuidade

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0$$

pode então ser escrita simplesmente como a condição de anulamento de um invariante, como

$$\partial_\mu j^\mu = 0. \quad (11.6)$$

As equações de Maxwell, que relacionam o campo eletromagnético com as cargas e correntes podem ser resumidas convenientemente em termos das equações para o potencial escalar ϕ e para o potencial vetor \vec{A}

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = -4\pi\rho \quad \text{e} \quad \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} = -\frac{4\pi}{c} \vec{j}$$

quando esses potenciais satisfazem a *condição de Lorentz*

$$\frac{1}{c} \frac{\partial \phi}{\partial t} + \vec{\nabla} \cdot \vec{A} = 0.$$

Como o operador diferencial que aparece nas equações que relacionam os potenciais com as fontes do campo (cargas e correntes) é invariante, as propriedades de transformação de ϕ e

de \vec{A} dever ser respectivamente iguais às de ρ e de \vec{j}/c . Desse modo é possível introduzir um *potencial quadrivetor* $\{A^\mu\}$ definido como

$$\{A^\mu\} = \{\phi, \vec{A}\} \quad (11.7)$$

em termos do qual essas equações podem ser escritas, usando também a definição do quadrivetor corrente, como

$$\partial^\mu \partial_\mu A^\nu = -\frac{4\pi}{c} j^\nu \quad (11.8)$$

enquanto a condição de Lorentz assume a forma

$$\partial_\mu A^\mu = 0. \quad (11.9)$$

Desse modo, o eletromagnetismo de Maxwell fica reduzido simplesmente às equações (11.8) com a condição de Lorentz (11.9) e com as definições (11.5) e (11.7). A conservação da carga elétrica total, por outro lado, é expressa pela condição (11.6), a ser satisfeita pelas componentes do quadrivetor densidade de corrente. As propriedades de transformação dos ingredientes dessas equações sob transformações de Lorentz garante *explicitamente* a sua validade, nessa mesma forma, em qualquer referencial de Lorentz.

Vale a pena ainda re-escrever nessa linguagem as equações de Maxwell para os campos \vec{E} e \vec{B} . Esses campos são dados em termos do potencial vetor e do potencial escalar através das relações

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad \text{e} \quad \vec{E} = -\vec{\nabla} \phi - \frac{1}{c} \frac{\partial \vec{A}}{\partial t}.$$

Como ϕ e \vec{A} estão relacionados com as componentes de um quadrivetor, as derivadas espaciais e temporais desses potenciais estão relacionados com as componentes de um tensor de segunda ordem. De fato, definindo o tensor de segunda ordem, anti-simétrico

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu = -F^{\nu\mu}$$

com $\{A^\mu\}$ dado por (11.7), é fácil verificar que

$$\{F^{\mu\nu}\} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \quad (11.10)$$

Em termos desse tensor, as equações de Maxwell

$$\vec{\nabla} \times \vec{B} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c} \vec{j} \quad \text{e} \quad \vec{\nabla} \cdot \vec{E} = 4\pi\rho$$

podem ser escritas sob a forma compacta

$$\partial_\mu F^{\mu\nu} = -\frac{4\pi}{c} j^\mu.$$

Para re-escrever as outras duas equações de Maxwell

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \quad \text{e} \quad \vec{\nabla} \cdot \vec{B} = 0,$$

que são semelhantes à versão homogênea das duas equações anteriores, com a substituição de \vec{B} por $-\vec{E}$ e vice-versa, é conveniente definir o tensor *dual*

$$\tilde{F}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\lambda\sigma} F_{\lambda\sigma} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & -E^1 & 0 \end{pmatrix}$$

que leva diretamente a

$$\partial_\mu \tilde{F}^{\mu\nu} = 0.$$

O objeto $\epsilon^{\mu\nu\lambda\sigma}$ envolvido na transformação de $F^{\mu\nu}$ no seu dual $\tilde{F}^{\mu\nu}$ é o símbolo anti-simétrico definido como sendo 1 ou -1 conforme $\mu\nu\lambda\sigma$ seja uma permutação par ou ímpar de 0,1,2,3. Esse objeto se comporta como um tensor contravariante de quarta ordem, e é fácil verificar que $\epsilon_{\mu\nu\lambda\sigma} = -\epsilon^{\mu\nu\lambda\sigma}$.

11.2 A equação de Klein-Gordon.

A equação de movimento de Schrödinger para a amplitude de probabilidade $\psi(\vec{r}, t)$ que descreve o estado de uma partícula livre

$$\left(-\frac{\hbar^2}{2M} \nabla^2 - i\hbar \frac{\partial}{\partial t} \right) \psi(\vec{r}, t) = 0$$

não define uma “boa teoria”, no sentido da seção anterior, pois o operador diferencial envolvido nela não é tensorialmente consistente sob transformações de Lorentz. Uma equação consistente

sob esse ponto de vista pode ser obtida notando que os valores do momento \vec{p} e da energia E para a onda plana que descreve o estado de uma partícula livre quântica

$$\psi(\vec{r}, t) = e^{i\frac{\vec{p}}{\hbar} \cdot \vec{r} - i\frac{E}{\hbar} t} \quad (11.11)$$

podem ser obtidos como autovalores dos operadores associados a esses observáveis através de

$$\vec{p} \longrightarrow \frac{\hbar}{i} \vec{\nabla} \quad \text{e} \quad E \longrightarrow i\hbar \frac{\partial}{\partial t}.$$

Com essa reinterpretação de \vec{p} e E , a relação relativística entre a energia e o momento

$$p^2 c^2 + M^2 c^4 = E^2 \quad (11.12)$$

que exprime a invariança do quadrado da norma do quadrvetor $\{E, \vec{p}c\} = Mcu^\mu$, onde u^μ é a quadrirelocidade, leva à equação

$$\left(-\hbar^2 c^2 \nabla^2 + M^2 c^4\right) \psi(\vec{r}, t) = -\hbar^2 \frac{\partial^2 \psi(\vec{r}, t)}{\partial t^2}$$

que pode ser também escrita como

$$\left(\partial^\mu \partial_\mu + \frac{M^2 c^2}{\hbar^2}\right) \psi(x) = 0 \quad (11.13)$$

com $x \equiv \{ct, \vec{r}\}$. A equação (11.13), que é claramente uma “boa” equação, envolvendo um operador diferencial invariante e um operador de multiplicação também invariante, expresso em termos da massa de repouso M , além das constantes c e \hbar , é conhecida como *equação de Klein-Gordon* (ou, eventualmente, *equação de Klein-Fock* para os russos)¹. Ela foi porém considerada pela primeira vez por Schrödinger, antes mesmo da formulação da teoria não relativística, sob uma forma mais interessante, que inclui os efeitos de um campo eletromagnético externo descrito pelos potenciais ϕ e \vec{A} , para estudar o espectro do hidrogênio. Essa forma parte da relação invariante, que estende a relação (11.12),

$$\left(\vec{p} + \frac{e}{c} \vec{A}\right)^2 c^2 + M^2 c^4 = (E + e\phi)^2$$

onde a carga do elétron foi escrita como $-e$. O caráter invariante dessa relação resulta do fato de que $\{\phi, \vec{A}\}$ também se transformam como as componentes de um quadrvetor contravariante. Mediante as mesmas substituições que levam à equação (11.13), esta última relação leva a

¹O. Klein, *Zeit. f. Physik* **37**, 895 (1926); V. Fock, *Zeit. f. Physik* **38**, 242 (1926) e *ibid* **39**, 226 (1926); W. Gordon, *Zeit. f. Physik* **40**, 117 (1926).

$$\left[\left(-i\hbar \vec{\nabla} + \frac{e}{c} \vec{A} \right)^2 c^2 + M^2 c^4 \right] \psi(\vec{r}, t) = \left(i\hbar \frac{\partial}{\partial t} + e\phi \right)^2 \psi(\vec{r}, t). \quad (11.14)$$

Os estados estacionários do átomo de hidrogênio poderiam então ser obtidos fazendo $\vec{A} = 0$ e $\phi = e/r$. Supondo que a dependência temporal das funções de onda estacionárias seja $e^{-iEt/\hbar}$, a equação a ser resolvida (com condições de contorno apropriadas) é

$$(-c^2 \hbar^2 \nabla^2 + M^2 c^4) \psi_E(\vec{r}) = \left(E + \frac{e^2}{r} \right)^2 \psi_E(\vec{r}).$$

As soluções obtidas por Schrödinger para essa equação dão resultados corretos para o espectro do átomo de hidrogênio na ordem α^2 (α sendo a constante de estrutura fina $e^2/\hbar c$), que corresponde ao espectro de Bohr (pois a constante de Rydberg pode ser escrita como $\alpha^2 m_e c^2/2$), mas os efeitos de ordem α^4 , associados à estrutura fina do espectro, são grandes demais por um fator próximo de três. Schrödinger atribuiu corretamente essa discrepância ao fato de que a equação usada não inclui o spin do elétron (proposto por Uhlenbeck e Goudsmit em 1925), tendo sido eventualmente levado a abandonar essa equação e a explorar o seu sucesso na ordem α^2 desenvolvendo a equação não relativística. O fato de que as equações (11.13) e (11.14) descrevem partículas *sem spin* está diretamente ligado à natureza da função de onda $\psi(\vec{r}, t)$, tomada como uma função complexa de \vec{r} e t , e portanto sem qualquer dependência possível de graus de liberdade internos. Um dos pontos centrais do trabalho de Dirac foi, de fato, generalizar essa equação para funções de onda com mais de uma componente, de forma a possibilitar a inclusão do spin.

Existe ainda outra peculiaridade associada à equação (11.13). Uma condição mínima para poder relacionar a interpretação das soluções $\psi(\vec{r}, t)$ com a adotada com sucesso na teoria não relativística consiste em obter em termos delas uma densidade $\rho(\vec{r})$ (interpretada na teoria não relativística como uma densidade de probabilidade) e uma densidade de corrente de $\vec{j}(\vec{r}, t)$ (interpretada na teoria não relativística como uma densidade de corrente de probabilidade) que satisfaçam uma equação de continuidade, exprimindo a conservação da integral espacial de $\rho(\vec{r}, t)$. Uma tal equação de continuidade pode ser obtida facilmente a partir da equação (11.13) e de sua complexo-conjugada

$$\left(\partial^\mu \partial_\mu + \frac{M^2 c^2}{\hbar^2} \right) \psi^*(\vec{r}, t) = 0.$$

De fato, multiplicando esta última equação por $\psi(\vec{r}, t)$, e subtraindo o resultado da equação (11.13) multiplicada por $\psi^*(\vec{r}, t)$, se obtém

$$\psi \partial^\mu \partial_\mu \psi^* - \psi^* \partial^\mu \partial_\mu \psi = \partial^\mu [\psi (\partial_\mu \psi^*) - \psi^* (\partial_\mu \psi)] = 0$$

que é uma equação de continuidade para uma densidade de corrente quadri-vetorial cujas componentes covariantes são

$$j_\mu = N [(\partial_\mu \psi^*)\psi - \psi^*(\partial_\mu \psi)]$$

onde N é um fator invariante introduzido a título de normalização. As componentes contravariantes espaciais dessa quadri-corrente são então

$$\vec{j} = N [\psi^*(\vec{\nabla}\psi) - (\vec{\nabla}\psi^*)\psi].$$

Essa densidade de corrente coincide com a expressão para a densidade de corrente de probabilidade da teoria não relativística para a escolha $N = \hbar/2Mi$, o que dá, para a quantidade correspondente à densidade de probabilidade, associada à componente 0 da quadri-corrente

$$\rho = \frac{j^0}{c} = \frac{\hbar}{2Mc^2i} \left[\left(\frac{\partial \psi^*}{\partial t} \right) \psi - \psi^* \left(\frac{\partial \psi}{\partial t} \right) \right]$$

A equação de continuidade pode então ser escrita como

$$\frac{\partial \rho(\vec{r}, t)}{\partial t} + \vec{\nabla} \cdot \vec{j}(\vec{r}, t) = 0$$

e exprime a conservação da quantidade definida por

$$P \equiv \int d^3r \rho(\vec{r}, t), \quad \text{isto é,} \quad \frac{dP}{dt} = 0.$$

Aqui, no entanto, a densidade $\rho(\vec{r}, t)$ não é, como no caso não relativístico, uma quantidade não negativa, o que *exclui a possibilidade de interpretá-la como uma densidade de probabilidade*, e portanto também a possibilidade de interpretar P como a probabilidade total.

A origem dessa peculiaridade pode ser identificada calculando $\rho(\vec{r}, t)$ para as soluções estacionárias com a forma de ondas planas. eq. (11.11). De fato, o que se obtém nesse caso é

$$\rho(\vec{r}, t) = \frac{E}{Mc^2}$$

onde a energia E e a massa de repouso M estão relacionadas através de (11.12), ou seja

$$E = \pm \sqrt{p^2 c^2 + M^2 c^4}$$

Essas relações mostram que a existência de valores negativos de ρ está ligada ao caráter quadrático da relação (11.12), que introduz a possibilidade de soluções de energia negativa. É ainda

esse mesmo caráter quadrático que leva a uma equação de segunda ordem em t , para a qual $\psi(\vec{r}, t = 0)$ e $\partial\psi(\vec{r}, t = 0)/\partial t$ são condições iniciais *independentes*. Isso significa que o estado da partícula é determinado dando, num determinado instante, não apenas uma, mas *duas* funções de \vec{r} , que podem ser tomadas como sendo ψ e $\partial\psi/\partial t$. Alternativamente, é possível usar para esse fim duas combinações lineares independentes quaisquer dessas duas funções. Um particular par de combinações lineares que permite fazer contacto com a teoria não relativística é

$$\psi_{\pm}(\vec{r}, t) \equiv \psi(\vec{r}, t) \pm \frac{i\hbar}{Mc^2} \frac{\partial\psi(\vec{r}, t)}{\partial t}$$

em termos do qual se pode definir o estado através da função de duas componentes

$$\Psi(\vec{r}, t) \equiv \begin{pmatrix} \psi_+(\vec{r}, t) \\ \psi_-(\vec{r}, t) \end{pmatrix} = \begin{pmatrix} \psi(\vec{r}, t) + \frac{i\hbar}{Mc^2} \frac{\partial\psi(\vec{r}, t)}{\partial t} \\ \psi(\vec{r}, t) - \frac{i\hbar}{Mc^2} \frac{\partial\psi(\vec{r}, t)}{\partial t} \end{pmatrix}.$$

A equação de Klein-Gordon adquire assim a forma de um par de equações acopladas com estrutura semelhante à da equação de Schrödinger não relativística

$$\left[(\sigma_3 + i\sigma_2) \frac{p^2}{2M} + \sigma_3 Mc^2 \right] \Psi(\vec{r}, t) = i\hbar \frac{\partial\Psi(\vec{r}, t)}{\partial t},$$

onde σ_2 e σ_3 são matrizes de Pauli. No limite não relativístico a energia da partícula difere muito pouco de Mc^2 (ou, em outros termos, $p^2 c^2 \ll M^2 c^4$), de modo que, tendo em conta que $i\hbar\partial/\partial t \rightarrow E$, $i\hbar\partial\psi(\vec{r}, t)/\partial t$ difere muito pouco de $Mc^2\psi(\vec{r}, t)$, e conseqüentemente ψ_- é tipicamente muito menor que ψ_+ . Nesse caso as equações tendem a se desacoplar, com ψ_+ satisfazendo

$$\left(\frac{p^2}{2M} + Mc^2 \right) \psi_+(\vec{r}, t) \approx i\hbar \frac{\partial\psi_+(\vec{r}, t)}{\partial t}.$$

Apenas nesse limite, portanto, a quantidade $\psi_+^* \psi_+$ pode ser interpretada como uma densidade de probabilidade. Por outro lado fica claro também que, no caso geral, a consideração de uma partícula relativística *sem spin* leva a uma situação envolvendo *dois* tipos de graus de liberdade, que podem ser associados, por exemplo, a $\psi_+(\vec{r}, t)$ e a $\psi_-(\vec{r}, t)$. Como visto, isso se relaciona com o fato de que a equação de movimento é de segunda ordem no tempo, e à conseqüente introdução de soluções com energia negativa.

Esse fato precisa ser levado em conta para obter uma interpretação para a constante do movimento P . A presença de dois tipos de graus de liberdade sugere, desde logo, que a equação de Klein-Gordon *não* possa em geral ser vista como uma equação de movimento para o estado quântico de uma única partícula, diferentemente do que ocorre no caso da equação de

Schrödinger não relativística. A solução geral da equação pode ser escrita como uma soma com coeficientes arbitrários das ondas planas (11.11). Usando condições de contorno periódicas num volume L^3 e separando explicitamente as componentes com energia positiva e negativa, uma solução desse tipo pode ser escrita como

$$\psi(\vec{r}, t) = \sum_{\vec{k}} \frac{1}{\sqrt{\omega_k L^3}} \left[a_{\vec{k}} e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + b_{\vec{k}} e^{i(\vec{k} \cdot \vec{r} + \omega_k t)} \right]$$

onde $\omega_k = +\sqrt{\hbar^2 k^2 c^2 + M^2 c^4} / \hbar$. Os termos envolvendo os coeficientes $a_{\vec{k}}$ e $b_{\vec{k}}$ correspondem portanto a estados de energia positiva e negativa, respectivamente, e o fator $\omega_k^{-1/2}$ foi introduzido explicitamente nesses coeficientes por conveniência. Usando as propriedades de ortogonalidade das ondas planas com condições de contorno periódicas é então fácil verificar que a quantidade conservada P é dada por

$$P = \frac{\hbar}{2Mc^2 i} \int_{L^3} d^3r \left(\frac{\partial \psi^*}{\partial t} \psi - \psi^* \frac{\partial \psi}{\partial t} \right) = \frac{\hbar}{Mc^2} \sum_{\vec{k}} (|a_{\vec{k}}|^2 - |b_{\vec{k}}|^2).$$

A característica importante dessa expressão é a de que ela é dada como a *diferença* de dois termos positivos, associados respectivamente à parte de energia positiva e à parte de energia negativa de $\psi(\vec{r}, t)$. O termo correspondente aos estados de energia negativa pode ser interpretado em termos da presença de *anti-partículas* no sistema, com o que P passa a ter o sentido do número de partículas *menos* o número de anti-partículas presentes no sistema. Com essa interpretação, no entanto, a função $\psi(\vec{r}, t)$ deixa de poder ser associada a um sistema constituído de uma partícula apenas, como no caso das soluções da equação de Schrödinger não relativística. Um tratamento mais satisfatório desta equação deve portanto consistir em trata-la como uma teoria envolvendo um *campo* escalar, complexo, $\psi(\vec{r}, t)$ que deve ser trabalhado com técnicas de segunda quantização em um espaço de Fock, de forma análoga à que foi utilizada para tratar o potencial vetor do campo eletromagnético livre na seção 8.1.1.

11.3 A equação de Dirac.

Para incluir a descrição de graus de liberdade internos como o spin do elétron é preciso, pelo menos, considerar soluções com mais de uma componente da equação de Klein-Gordon. Ao tratar dessa questão, Dirac procurou resolver também o problema das densidades negativas postulando que as soluções relevantes, além de satisfazerem a equação de Klein-Gordon, devem satisfazer *também* uma equação linear, invariante e *de primeira ordem no tempo*. A invariância sob transformações de Lorentz exige então que ela deva ser também de primeira ordem com relação às variáveis espaciais. A forma mais simples de uma tal equação é

$$\left(i\gamma^\mu \frac{\partial}{\partial x^\mu} - K\right) \psi(\vec{r}, t) \equiv (i\gamma^\mu \partial_\mu - K) \psi(\vec{r}, t) = 0 \quad (11.15)$$

onde os $\{\gamma^\mu\} \equiv \{\gamma^0, \vec{\gamma}\}$ devem ter propriedades de componentes de um quadrivetor contravariante e K é invariante.

A condição de que $\psi(\vec{r}, t)$ deve satisfazer também a equação de Klein-Gordon impõe certas restrições sobre os coeficientes γ^μ e sobre K . Essas restrições podem ser determinadas considerando a equação de segunda ordem que resulta da aplicação do operador $-i\gamma^\mu \partial_\mu - K$ sobre a equação (11.15), o que leva a

$$\begin{aligned} 0 &= (-i\gamma^\mu \partial_\mu - K)(i\gamma^\mu \partial_\mu - K) \psi(\vec{r}, t) = (\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + K^2) \psi(\vec{r}, t) = \\ &= \left(\frac{1}{2}(\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) \partial_\mu \partial_\nu + K^2\right) \psi(\vec{r}, t) \end{aligned}$$

onde foi usado o fato de que $\partial_\nu \partial_\mu = \partial_\mu \partial_\nu$. Essa equação reproduz a equação de Klein-Gordon se K for escolhido como Mc/\hbar e

$$\frac{1}{2}(\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) = g^{\mu\nu}. \quad (11.16)$$

Com essa condição, portanto, o operador que aparece na equação de Klein-Gordon pode ser fatorizado como

$$\partial^\mu \partial_\mu + \frac{M^2 c^2}{\hbar^2} = g^{\mu\nu} \partial_\mu \partial_\nu + \frac{M^2 c^2}{\hbar^2} = \left(-i\gamma^\mu \partial_\mu - \frac{Mc}{\hbar}\right) \left(i\gamma^\nu \partial_\nu - \frac{Mc}{\hbar}\right).$$

As soluções de (11.15) serão então automaticamente soluções também da equação de Klein-Gordon, embora esta possa admitir soluções adicionais que não satisfazem a equação (11.15).

A condição (11.16) mostra porém que os γ^μ não podem ser tomados como sendo números, mas que podem ser implementados em termos de matrizes. De fato, ela implica na *anti-comutatividade* desses objetos para $\mu \neq \nu$, isto é

$$\mu \neq \nu \quad \rightarrow \quad \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 0 \quad (11.17)$$

além das relações correspondentes aos casos em que $\mu = \nu$

$$(\gamma^0)^2 = 1 \quad \text{e} \quad (\gamma^k)^2 = -1, \quad k = 1, 2, 3.$$

É também imediato que $\gamma_0 = \gamma^0$ e $\gamma_k = -\gamma^k$, $k = 1, 2, 3$, e portanto $\gamma^\mu \gamma_\mu = \gamma_\mu \gamma^\mu = 1$ (sem soma sobre μ) o que significa, em particular, que $(\gamma^\mu)^{-1} = \gamma_\mu$. Multiplicando então a relação (11.17) por γ_μ (ainda sem soma sobre o índice repetido μ)

$$0 = (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) \gamma_\mu = \gamma^\mu \gamma^\nu \gamma_\mu + \gamma^\nu$$

e, tomando o traço,

$$\text{Tr } \gamma^\nu = -\text{Tr } (\gamma^\mu \gamma^\nu \gamma_\mu) = -\text{Tr } (\gamma^\nu \gamma_\mu \gamma^\mu) = -\text{Tr } \gamma^\nu$$

donde segue que as matrizes γ^ν devem ter *traço nulo*. Por outro lado, como $\gamma_0 = \gamma^0$, $(\gamma^0)^2 = 1$, os autovalores de γ^0 devem ser ± 1 . Para que o seu traço possa ser nulo, é preciso então que a sua dimensão, e portanto também a dimensão das demais matrizes γ^μ , seja *par*.

É fácil ver que os γ^μ não podem ser realizados em termos de matrizes de dimensão 2. De fato, como visto na seção 6.1.2, as três matrizes de Pauli e mais a matriz unidade de dimensão 2 constituem um conjunto completo de matrizes em termos das quais é possível escrever qualquer matriz de dimensão 2. As três matrizes de Pauli obedecem, de fato, a relações de anti-comutação iguais às exigidas para os γ^μ , mas a única outra matriz que é independente dessas três matrizes *comuta* com qualquer uma delas e tem traço diferente de zero. Isso mostra que não existem *quatro* matrizes de dimensão 2 satisfazendo a condição (11.17). Os γ^μ devem portanto ser realizados em termos de matrizes de dimensão pelo menos igual a 4. A escolha dessa dimensão é a que leva à equação relativística de primeira ordem adotada por Dirac. Existe uma infinidade de representações possíveis das matrizes γ^μ nesse caso. Uma das representações possíveis, e que é de uso freqüente, é a que é dada por

$$\gamma^0 = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix}, \quad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \quad (11.18)$$

onde 1_2 é a matriz identidade de dimensão 2, 0 representa uma matriz nula também de dimensão 2 e $\vec{\sigma}$ são as matrizes de Pauli. A amplitude $\psi(\vec{r}, t)$ é portanto um objeto de quatro componentes e a equação de Dirac (11.15) é na realidade um sistema de quatro equações acopladas para as componentes de $\psi(\vec{r}, t)$.

11.3.1 Forma hamiltoniana e soluções da equação de Dirac livre.

Separando explicitamente as componentes espaciais da componente temporal na equação (11.15) ela assume a forma, de estrutura semelhante à da equação de Schrödinger não relativística

$$\left(i\vec{\gamma} \cdot \vec{\nabla} + i\gamma^0 \frac{1}{c} \frac{\partial}{\partial t} - \frac{Mc}{\hbar} \right) \psi(\vec{r}, t) = 0.$$

Multiplicando ainda essa equação pela esquerda pela matriz $\hbar c \gamma^0$, e lembrando que $(\gamma^0)^2 = 1$ resulta

$$(c \vec{\alpha} \cdot \vec{p} + \beta M c^2) \psi(\vec{r}, t) = i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} \quad (11.19)$$

onde foram introduzidas as definições

$$\beta = \gamma^0, \quad \vec{\alpha} = \gamma^0 \vec{\gamma} = \beta \vec{\gamma}$$

e $\vec{p} = \hbar \vec{\nabla}/i$ é o operador associado ao momento. A equação (11.19) é a forma hamiltoniana da equação de Dirac para uma partícula livre. O operador hamiltoniano, que tem o papel de gerador da evolução temporal, é *linear* no momento. Ele é dado por

$$H = c \vec{\alpha} \cdot \vec{p} + \beta M c^2 \quad (11.20)$$

ou, usando explicitamente a representação (11.18) para as matrizes γ^μ ,

$$H = \begin{pmatrix} M c^2 \times 1_2 & c \vec{\sigma} \cdot \vec{p} \\ c \vec{\sigma} \cdot \vec{p} & -M c^2 \times 1_2 \end{pmatrix}.$$

Esse hamiltoniano é evidentemente uma matriz hermiteana de dimensão 4. Na representação adotada, ela foi escrita em termos de blocos de dimensão 2, que envolvem seja a matriz identidade de dimensão 2, 1_2 , seja as matrizes de Pauli $\vec{\sigma}$. Por outro lado, a hermiticidade de H deve ser preservada independentemente da particular representação que seja adotada para as matrizes γ^μ . Isso significa que devem ser satisfeitas *em geral* as condições

$$\vec{\alpha} = \gamma^0 \vec{\gamma} = \vec{\alpha}^\dagger = \vec{\gamma}^\dagger \gamma^{0\dagger} \quad \text{e} \quad \beta = \gamma^0 = \beta^\dagger = \gamma^{0\dagger}$$

ou seja,

$$\gamma^{0\dagger} = \gamma^0 \quad \text{e} \quad \gamma^{k\dagger} = \gamma^0 \gamma^k \gamma^0 = -\gamma^k. \quad (11.21)$$

Uma interpretação simples para as três matrizes hermiteanas $\vec{\alpha}$ (de dimensão 4) resulta da equação de Heisenberg

$$i\hbar \frac{d\vec{r}}{dt} = [\vec{r}, H] = i\hbar c \vec{\alpha},$$

que mostra que $\vec{\alpha}$ funciona como o operador associado à velocidade da partícula, em unidades de c , isto é

$$\vec{\alpha} = \frac{1}{c} \frac{d\vec{r}}{dt} = \frac{\vec{v}}{c}. \quad (11.22)$$

Como as componentes de $\vec{\alpha}$ não comutam entre si, e também não comutam com H , as diferentes componentes da velocidade não podem ser simultaneamente bem definidas, e também as soluções estacionárias para a partícula de Dirac livre, que são da forma

$$\psi_E(\vec{r}, t) = \phi_E(\vec{r})e^{-i\frac{E}{\hbar}t}$$

com

$$(c \vec{\alpha} \cdot \vec{p} + \beta Mc^2) \phi_E(\vec{r}) = E \phi_E(\vec{r})$$

em geral não podem ser associadas a um valor bem definido da velocidade. Na realidade, como os autovalores das componentes de $\vec{\alpha}$ são ± 1 , os autovalores das componentes do operador associado à velocidade são $\pm c$. Uma interpretação para a matriz β , por outro lado, pode ser obtida a partir de

$$\beta = \frac{H - c\vec{\alpha} \cdot \vec{p}}{Mc^2}.$$

Tomando o valor médio dessa relação entre operadores numa função de onda normalizada

$$\langle \beta \rangle = \frac{\langle E \rangle}{Mc^2} - \frac{\langle \vec{v} \cdot \vec{p} \rangle}{Mc^2} \longrightarrow \frac{1}{Mc^2} \left(\frac{Mc^2}{\sqrt{1 - v^2/c^2}} - \frac{Mv^2}{\sqrt{1 - v^2/c^2}} \right) = \sqrt{1 - \frac{v^2}{c^2}} \quad (11.23)$$

onde os valores médios da energia E e de $\vec{v} \cdot \vec{p}$ foram interpretados em termos das expressões relativísticas clássicas correspondentes (v. eq. (11.4)). Isso indica que no limite não relativístico $v \ll c$ a matriz β tende a funcionar como a matriz identidade para estados de energia positiva. Tendo em vista a forma de $\beta = \gamma^0$ na representação particular (11.18), deve-se esperar que a terceira e quarta componentes de $\psi(\vec{r}, t)$ sejam pouco importantes (i.e., “pequenas”) nesse limite, quando seja usada essa representação.

Não é difícil obter soluções estacionárias explícitas para a equação de Dirac livre, escrita na forma (11.20), utilizando a representação (11.18). As soluções devem ser objetos de quatro componentes, com valores bem definidos para o momento, $\hbar \vec{k}$ e para a energia, E . Introduzindo o *ansatz*

$$\psi(\vec{r}) = e^{i\vec{k} \cdot \vec{r} - iEt/\hbar} u(\vec{k}, E)$$

onde $u(\vec{k}, E)$ é entendido como um objeto com quatro componentes, possivelmente dependentes dos valores do momento e da energia, resulta para esse objeto a equação

$$\begin{pmatrix} Mc^2 \times 1_2 & c\hbar\vec{\sigma} \cdot \vec{k} \\ c\hbar\vec{\sigma} \cdot \vec{k} & -Mc^2 \times 1_2 \end{pmatrix} u(\vec{k}, E) = E u(\vec{k}, E).$$

Escrevendo também $u(\vec{k}, E)$ em termos de duas partes de duas componentes como

$$u(\vec{k}, E) = \begin{pmatrix} U \\ V \end{pmatrix}$$

e explicitando as operações matriciais envolvendo os blocos de dimensão 2 resultam as duas relações

$$Mc^2 U + c\hbar\vec{\sigma} \cdot \vec{k} V = E U \quad \text{e} \quad c\hbar\vec{\sigma} \cdot \vec{k} U - Mc^2 V = E V. \quad (11.24)$$

Isolando V na segunda equação resulta

$$V = \frac{c\hbar\vec{\sigma} \cdot \vec{k}}{E + Mc^2} U$$

que, após substituição na primeira equação, dá

$$(E - Mc^2)U = \frac{c^2\hbar^2(\vec{\sigma} \cdot \vec{k})^2}{E + Mc^2} U = \frac{c^2\hbar^2\vec{k}^2}{E + Mc^2} U.$$

Note que nesta última passagem foi usada a identidade $(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) \equiv \vec{A} \cdot \vec{B} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B})$. Da última relação segue que

$$(E^2 - M^2c^4 - c^2\hbar^2\vec{k}^2) U = 0$$

e portanto *qualquer* objeto de duas componentes U será solução desde que se tenha

$$E^2 = M^2c^4 + c^2\hbar^2\vec{k}^2 \quad \text{ou seja} \quad E = \pm\sqrt{M^2c^4 + c^2\hbar^2\vec{k}^2}.$$

Isso mostra que há soluções de energia positiva e também soluções de energia negativa. Por outro lado, da arbitrariedade de U segue que para cada um dos sinais de E há *duas* soluções independentes, que podem ser escolhidas, por exemplo, tomando U como

$$U_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{ou} \quad U_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Em cada um desses casos V é determinado pela escolha de U , de modo que há em geral *quatro* soluções independentes para cada valor $\hbar\vec{k}$ do momento.

Para escrever explicitamente (a menos de um eventual fator de normalização) as duas soluções de energia positiva basta obter as componentes V_1 e V_2 correspondentes a U_1 e a U_2 respectivamente. Isso é feito simplesmente usando a relação entre U e V

$$V_i = \frac{c\hbar\vec{\sigma} \cdot \vec{k}}{E + Mc^2} U_i = \frac{c\hbar}{E + Mc^2} \begin{pmatrix} k^3 & k^- \\ k^+ & -k^3 \end{pmatrix} U_i, \quad \vec{k} \equiv \{k^1, k^2, k^3\}, \quad k^\pm \equiv k^1 \pm ik^2$$

onde $E = +\sqrt{M^2c^4 + c^2\hbar^3\vec{k}^2}$, que dá

$$V_1 = \frac{c\hbar}{E + Mc^2} \begin{pmatrix} k^3 \\ k^+ \end{pmatrix} \quad \text{e} \quad V_2 = \frac{c\hbar}{E + Mc^2} \begin{pmatrix} k^- \\ -k^3 \end{pmatrix}.$$

Essas expressões mostram, em particular, que V_1 e V_2 tendem a zero quando $\hbar\vec{k} \rightarrow 0$. Esse mesmo procedimento pode ser usado para obter (também a menos de um eventual fator de normalização) as duas soluções de energia negativa, para o que na realidade basta substituir E por $-E$ no resultado final. Desse modo, no entanto, as componentes V *divergem* no limite $\hbar\vec{k} \rightarrow 0$. Essa divergência é apenas um efeito de normalização, associado ao fato de que as componentes U das soluções de energia negativa se tornam desprezíveis nesse limite. Para evitá-la, no entanto, é possível adotar um procedimento de solução simétrico ao usado para as soluções de energia positiva, isto é, resolver a primeira das equações (11.24) para U , o que leva a

$$U = -\frac{c\hbar\vec{\sigma} \cdot \vec{k}}{Mc^2 - E} V$$

obtendo em seguida, por substituição na segunda equação,

$$(M^2c^4 - E^2 + c^2\hbar^2\vec{k}^2)V = 0.$$

Fazendo então as escolhas para V

$$\bar{V}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{e} \quad \bar{V}_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

resulta, para as correspondentes componentes \bar{U} ,

$$\bar{U}_3 = -\frac{c\hbar}{Mc^2 - E} \begin{pmatrix} k^3 \\ k^+ \end{pmatrix} \quad \text{e} \quad \bar{U}_4 = \frac{c\hbar}{Mc^2 - E} \begin{pmatrix} -k^- \\ k^3 \end{pmatrix}$$

que desta forma se anulam no limite $\hbar \vec{k} \rightarrow 0$. É fácil verificar explicitamente que as soluções U_1, V_1 e U_2, V_2 , bem como as soluções \bar{U}_3, \bar{V}_3 e \bar{U}_4, \bar{V}_4 são mutuamente ortogonais, e que o mesmo vale para qualquer das primeiras duas, calculadas para $E = +\sqrt{M^2 c^4 + \hbar^2 k^2 c^2}$ com relação a qualquer das segundas, calculadas para $E = -\sqrt{M^2 c^4 + \hbar^2 k^2 c^2}$. Nessas mesmas condições, essas soluções formam um conjunto completo de soluções, no sentido que qualquer objeto de quatro componentes pode ser expresso como uma combinação linear delas.

Existe forma rápida de obter as soluções da equação de Dirac livre em termos de uma representação *qualquer* das matrizes γ^μ . Partindo da equação escrita sob a forma (em unidades tais que $\hbar = c = 1$)

$$(i\gamma^\mu \partial_\mu - M) \psi(x) = 0, \quad x \equiv \{x^\mu\} = \{ct, \vec{r}\} \quad (11.25)$$

e fazendo para $\psi(x)$ o ansatz

$$\psi(x) = e^{-ip \cdot x} u(p) = e^{-ip_\mu x^\mu} u(p), \quad \{p_\mu\} = \{E, -\vec{p}\}$$

resulta para $u(p)$ a equação

$$(\gamma^\mu p_\mu - M) u(p) \equiv (\not{p} - M) u(p) = 0 \quad (11.26)$$

onde foi introduzida a notação (devida a Feynman) $\not{p} \equiv \gamma^\mu p_\mu$. As soluções $u(p)$ dessa equação podem ser obtidas notando que

$$(\not{p} - M)(\not{p} + M) = \not{p}\not{p} - M^2 = \gamma^\mu \gamma^\nu p_\mu p_\nu - M^2 = g^{\mu\nu} p_\mu p_\nu - M^2 = E^2 - \vec{p}^2 - M^2.$$

É importante não esquecer que esse objeto é uma matriz de dimensão 4. O último resultado acima indica portanto um múltiplo da matriz identidade de dimensão 4. Para $E^2 - \vec{p}^2 - M^2 = 0$, essa matriz se anula, e as soluções de (11.26) são as colunas da matriz $(\not{p} + m)$, pois

$$(\not{p} - M)(\not{p} + M) = 0 \quad \text{se} \quad E^2 - \vec{p}^2 - M^2 = 0.$$

Esse resultado geral pode ser particularizado para a representação (11.18). Nesse caso

$$(\not{p} + M) = \begin{pmatrix} (E + M)1_2 & -\vec{\sigma} \cdot \vec{p} \\ \vec{\sigma} \cdot \vec{p} & (M - E)1_2 \end{pmatrix}$$

cujas colunas reproduzem (a menos de fatores de normalização) as soluções obtidas anteriormente de outro modo. Nesta última expressão a matriz de dimensão 4 foi escrita em termos de blocos de dimensão 2, envolvendo a matriz identidade 1_2 e as matrizes de Pauli $\vec{\sigma}$.

As soluções obtidas acima mostram que o hamiltoniano livre de Dirac tem um espectro contínuo bipartido, de $E = Mc^2$ a $E \rightarrow +\infty$ e de $E = -Mc^2$ a $E \rightarrow -\infty$, e não é portanto limitado inferiormente. A exigência de que, além de satisfazer a equação de Klein-Gordon, as soluções aceitáveis de quatro componentes satisfaçam também à equação de primeira ordem (11.19) (ou equivalentemente, à equação (11.15)) não é portanto suficiente para evitar as soluções de energia negativa.

A equação de primeira ordem, no entanto, de fato leva a uma equação de continuidade envolvendo uma densidade positiva definida. De fato, escrevendo a solução $\psi(\vec{r}, t)$ explicitamente como

$$\psi(\vec{r}, t) = \begin{pmatrix} \psi_1(\vec{r}, t) \\ \psi_2(\vec{r}, t) \\ \psi_3(\vec{r}, t) \\ \psi_4(\vec{r}, t) \end{pmatrix},$$

tomando o produto matricial da equação (11.19) com

$$\psi^\dagger(\vec{r}, t) = (\psi_1^*(\vec{r}, t) \quad \psi_2^*(\vec{r}, t) \quad \psi_3^*(\vec{r}, t) \quad \psi_4^*(\vec{r}, t))$$

pela esquerda, e subtraído a relação resultante de sua conjugada complexa, resulta a equação de continuidade

$$\vec{\nabla} \cdot \vec{j}(\vec{r}, t) + \frac{\partial \rho(\vec{r}, t)}{\partial t} = 0$$

com

$$\rho(\vec{r}, t) = \psi^\dagger(\vec{r}, t)\psi(\vec{r}, t) = \sum_{j=1}^4 |\psi_j(\vec{r}, t)|^2 \geq 0 \quad \text{e} \quad \vec{j}(\vec{r}, t) = \psi^\dagger(\vec{r}, t) c \vec{\alpha} \psi(\vec{r}, t).$$

Essa equação pode ser expressa de uma forma mais simétrica definindo a solução *adjunta*

$$\bar{\psi}(\vec{r}, t) \equiv \psi^\dagger(\vec{r}, t)\beta = \psi^\dagger(\vec{r}, t)\gamma^0$$

em termos da qual, usando o fato de que $(\gamma^0)^2 = 1$, a densidade é dada por

$$\rho(\vec{r}, t) = \frac{1}{c} \bar{\psi}(\vec{r}, t) c \gamma^0 \psi(\vec{r}, t) \equiv \frac{1}{c} j^0(\vec{r}, t) \quad (11.27)$$

enquanto as três componentes da corrente são

$$j^k(\vec{r}, t) = \bar{\psi}(\vec{r}, t) c \gamma^k \psi(\vec{r}, t), \quad k = 1, 2, 3 \quad (11.28)$$

de modo que a equação de continuidade fica reduzida a

$$\partial_\mu j^\mu = 0.$$

Para que essa equação possa ser vista como invariante é necessário ainda verificar que as quatro componentes $\{j^\mu\}$ efetivamente se transformam como as componentes de um quadrivetor contravariante. Isso será feito na seção 11.3.3 abaixo.

11.3.2 Propriedades algébricas das matrizes gama.

A expansão de uma matriz qualquer de dimensão 2 como combinação linear das três matrizes de Pauli e mais a matriz identidade foi tratada na seção 6.1.2. O resultado análogo para matrizes de dimensão 4 deve envolver $4 \times 4 = 16$ matrizes, e o ponto a ser examinado aqui consiste em obter um conjunto de matrizes que permita essa expansão a partir das matrizes γ^μ , utilizando apenas as suas propriedades gerais dadas pela relação (11.16) e de traço nulo. As matrizes a serem utilizadas para esse fim podem ser construídas do seguinte modo:

- a) A matriz identidade de dimensão 4, 1_4 .
- b) As quatro matrizes γ^μ .
- c) Seis matrizes $\sigma^{\mu\nu} \equiv \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)$, com $\mu \neq \nu$.
- d) Os quatro diferentes produtos de três matrizes gama. Tendo em conta que $(\gamma^0)^2 = 1_4$ e que $(\gamma^k)^2 = -1_4$, $k = 1, 2, 3$, bem como a anticomutatividade de matrizes diferentes, essas quatro matrizes podem ser substituídas pelos produtos $i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^\mu \equiv \gamma_5\gamma^\mu$, com a definição

$$\gamma_5 \equiv \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3.$$

Assim, por exemplo, $\gamma^0\gamma^2\gamma^3 = i\gamma_5\gamma^1$, etc.

- e) A matriz γ_5 , proporcional ao produto das quatro matrizes γ^μ .

Com a exceção evidente da matriz 1_4 , essas matrizes têm traço zero. Isso já foi mostrado para as matrizes γ^μ , e é evidente para as matrizes $\sigma^{\mu\nu}$, em vista da propriedade cíclica do traço, $\text{Tr}(AB) = \text{Tr}(BA)$. Usando ainda a propriedade cíclica do traço e a anticomutatividade das matrizes γ^μ segue ainda que

$$\text{Tr}(\gamma^0\gamma^1\gamma^2\gamma^3) = \text{Tr}(\gamma^1\gamma^2\gamma^3\gamma^0) = -\text{Tr}(\gamma^0\gamma^1\gamma^2\gamma^3) = 0$$

e portanto γ_5 tem traço zero. Finalmente é fácil ver que $\gamma_5\gamma^\mu = -\gamma^\mu\gamma_5$, donde segue que as matrizes **d)** também têm traço zero, usando mais uma vez a propriedade cíclica do traço.

Outra propriedade relevante é a de que um produto qualquer de duas matrizes *diferentes da matriz* 1_4 e também *diferentes entre si* também tem traço zero. Esta propriedade pode ser

verificada a partir do fato de que um tal produto é sempre proporcional a uma das 15 matrizes de traço zero. Por outro lado,

$$\text{Tr}(1_4)^2 = 4, \quad \text{Tr}(\gamma^\mu)^2 = 4g^{\mu\mu}, \quad \text{Tr}(\sigma^{\mu\nu})^2 = 4g^{\mu\mu}g^{\nu\nu}, \quad \text{Tr}(\gamma_5\gamma^\mu)^2 = -4g^{\mu\mu}, \quad \text{Tr}(\gamma_5)^2 = 4.$$

Dessas propriedades decorre, então, que as 16 matrizes consideradas são *linearmente independentes*, no sentido que o anulamento de uma combinação linear do tipo

$$a 1_4 + b_\mu \gamma^\mu + c_{\mu\nu} \sigma^{\mu\nu} + d_\mu \gamma_5 \gamma^\mu + e \gamma_5 = A = 0 \quad (11.29)$$

implica no anulamento dos coeficientes, isto é,

$$a = b_\mu = c_{\mu\nu} = d_\mu = e = 0,$$

e portanto qualquer matriz de dimensão 4 pode ser escrita como uma combinação linear delas. A independência linear pode ser verificada multiplicando (11.29) sucessivamente por cada uma das 16 matrizes e tomando o traço. No caso em que a combinação linear seja uma matriz *não nula* A , de dimensão 4, resulta que

$$A = \frac{\text{Tr } A}{\text{Tr}(1_4)^2} 1_4 + \frac{\text{Tr}(\gamma^\mu A)}{\text{Tr}(\gamma^\mu)^2} \gamma^\mu + \frac{\text{Tr} \sigma^{\mu\nu} A}{\text{Tr}(\sigma^{\mu\nu})^2} \sigma^{\mu\nu} + \frac{\text{Tr}(\gamma_5 \gamma^\mu A)}{\text{Tr}(\gamma_5 \gamma^\mu)^2} \gamma_5 \gamma^\mu + \frac{\text{Tr}(\gamma_5 A)}{\text{Tr}(\gamma_5)^2} \gamma_5.$$

A conveniência de expansões em termos das 16 matrizes acima, relativamente a expansões em termos de algum outro conjunto de matrizes linearmente independentes, se deve a que ela se relaciona de forma natural com objetos que têm propriedades de transformação tensoriais bem definidas sob transformações de Lorentz. Esse ponto será elaborado na seção seguinte.

11.3.3 Transformações de Lorentz e rotações.

Dois tipos de transformação interessantes para a equação de Dirac livre são transformações de Lorentz e rotações, que estão relacionadas respectivamente com o caráter invariante da equação sob transformações de Lorentz e ao momento angular da partícula de Dirac.

Para uma transformação de Lorentz infinitesimal (isto é, com $\delta v \ll c$) do tipo especial considerado na equação (11.3), em que a velocidade relativa δv dos dois referenciais tem a direção x^1 , é possível escrever

$$(\Lambda_1)^\mu{}_\nu = \delta^\mu_\nu + \delta\varphi(\Theta_1)^\mu{}_\nu + \mathcal{O}(\delta\varphi^2)$$

com $\delta\varphi = \delta v/c + \mathcal{O}(\delta v^2)$ resultante da expansão até primeira ordem de $\sinh \delta\varphi$, e onde foi introduzido o gerador

$$\{(\Theta_1)^\mu{}_\nu\} \equiv \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Essa expressão pode ser generalizada imediatamente para uma transformação infinitesimal associada a uma velocidade relativa de direção arbitrária $\delta\vec{v} \equiv \{\delta v_1, \delta v_2, \delta v_3\}$ escrevendo

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad \text{com} \quad \Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \delta\vec{\varphi} \cdot \vec{\Theta}^\mu{}_\nu + \mathcal{O}(\delta\varphi^2)$$

onde $\delta\vec{\varphi} = \{\delta v_1/c, \delta v_2/c, \delta v_3/c\} + \mathcal{O}(\delta v^2)$ e $\vec{\Theta} = \{\Theta_1, \Theta_2, \Theta_3\}$, sendo os dois geradores adicionais dados por

$$\{(\Theta_2)^\mu{}_\nu\} \equiv \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{e} \quad \{(\Theta_3)^\mu{}_\nu\} \equiv \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

Por outro lado, a matriz de $r^\mu{}_\nu$, de dimensão 4, que corresponde a uma rotação espacial infinitesimal $\delta\vec{\omega}$, escrita como

$$x'^\mu = r^\mu{}_\nu x^\nu$$

pode ser posta sob a forma

$$r^\mu{}_\nu = \delta^\mu{}_\nu + \delta\vec{\omega} \cdot \vec{\lambda}^\mu{}_\nu + \mathcal{O}(\delta\omega^2)$$

onde as matrizes $\vec{\lambda} = \{\lambda_1, \lambda_2, \lambda_3\}$ são tais que $(\lambda_i)^0{}_\nu = (\lambda_i)^\mu{}_0 = 0$ e as sub-matrizes de dimensão 3 cujos elementos são $(\lambda_i)^j{}_k = -(\lambda_i)^k{}_j$ são os geradores anti-simétricos que foram escritos em (9.1) e (9.8).

Portanto, a combinação de uma transformação de Lorentz infinitesimal com uma rotação espacial, também infinitesimal, pode ser escrita de forma compacta como

$$x'^\mu = T^\mu{}_\nu x^\nu$$

onde a matriz $\{T^\mu{}_\nu\}$ pode ser escrita, *a menos de termos de segunda ordem*, como

$$T^\mu{}_\nu = \delta^\mu{}_\nu + \delta\vec{\varphi} \cdot \vec{\Theta}^\mu{}_\nu + \delta\vec{\omega} \cdot \vec{\lambda}^\mu{}_\nu \equiv \delta^\mu{}_\nu + L^\mu{}_\nu + R^\mu{}_\nu \quad (11.30)$$

onde a matriz L^μ_ν é simétrica, com $L^0_0 = 0$ e $L^i_j = 0$ para $i, j = 1, 2, 3$ e a matriz R^μ_ν é anti-simétrica, com $R^0_\nu = R^\mu_0 = 0$. Alternativamente, essa transformação pode ser caracterizada pela matrizes

$$L^{\mu\nu} = g^{\nu\lambda} L^\mu_\lambda \quad \text{ou} \quad L_{\mu\nu} = g_{\mu\lambda} L^\lambda_\nu$$

que, tendo em vista a definição do tensor métrico $g^{\mu\nu} = g_{\mu\nu}$, são matrizes anti-simétricas de dimensão 4.

Para estudar as propriedades de transformação das soluções da equação de Dirac livre (11.25) basta estudar as propriedades de transformação do conjunto completo de soluções elementares da forma

$$\psi(x) = e^{-ip \cdot x} u(p),$$

obtidas anteriormente, sob transformações de Lorentz e rotações infinitesimais, dado que transformações finitas podem ser obtidas dessas usando a lei de composição de grupo. Além disso, como o fator exponencial é explicitamente invariante tanto sob transformações de Lorentz como sob rotações, é suficiente, na realidade, estudar as soluções $u(p)$ da equação (11.26)

$$(\gamma^\mu p_\mu - M) u(p) = 0.$$

Dado que os p_μ são componentes covariantes de um quadrivetor, essa equação é invariante se os γ^μ se transformam como as componentes contravariantes também de um quadrivetor, isto é

$$\gamma'^\mu = T^\mu_\nu \gamma^\nu.$$

Com isso $\gamma'^\mu p'_\mu = \gamma^\mu p_\mu$ e a equação transformada é satisfeita por $u'(p') = u(p)$, isto é

$$(\gamma'^\mu p'_\mu - M) u(p) = (\gamma^\mu p_\mu - M) u(p) = 0. \quad (11.31)$$

No entanto, as matrizes γ'^μ a serem usadas no novo referencial são *diferentes* das matrizes originais γ^μ . Para evitar esse inconveniente, o que se faz usualmente é associar à transformação T^μ_ν uma *mudança de representação* das matrizes γ^μ , escolhida de modo tal que, *no novo referencial* as matrizes transformadas tenham a mesma forma que as matrizes originais no referencial antigo. Chamando S a matriz associada a essa mudança de representação, supondo que S seja inversível (o que poderá ser verificado explicitamente *a posteriori*) e impondo a condição

$$S^{-1} \gamma^\mu S = \gamma'^\mu = T^\mu_\nu \gamma^\nu \quad (11.32)$$

resulta da equação (11.31) que

$$\left(S^{-1} \gamma^\mu S p'_\mu - M \right) u(p) = S^{-1} \left(\gamma^\mu p'_\mu - M \right) S u(p) = 0$$

de modo que

$$\left(\gamma^\mu p'_\mu - M \right) S u(p) = 0.$$

Nessas condições, as soluções transformadas são dadas por $u'(p') = S u(p)$, e

$$\psi'(x') = e^{-ip' \cdot x'} S u(p) = S \psi(x). \quad (11.33)$$

As matrizes S constituem portanto uma representação do grupo de transformações considerado (transformações de Lorentz e rotações espaciais) no espaço linear das funções de onda de Dirac.

Para transformações infinitesimais $T^\mu{}_\nu$, a matriz S deve também diferir infinitesimalmente da matriz identidade de dimensão 4, e pode portanto ser escrita como

$$S = 1_4 + \epsilon + \mathcal{O}(\epsilon^2) \quad \text{onde segue que} \quad S^{-1} = 1_4 - \epsilon + \mathcal{O}(\epsilon^2),$$

e a relação (11.32) mostra que a matriz ϵ se relaciona com a matriz de transformação $T^\mu{}_\nu$ através de (a menos de termos de segunda ordem)

$$\gamma^\mu - [\epsilon, \gamma^\mu] = T^\mu{}_\nu \gamma^\nu = \gamma^\mu + L^\mu{}_\nu \gamma^\nu + R^\mu{}_\nu \gamma^\nu, \quad \text{ou seja,} \quad [\gamma^\mu, \epsilon] = (L^\mu{}_\nu + R^\mu{}_\nu) \gamma^\nu.$$

Da mesma forma que qualquer matriz de dimensão 4, a matriz ϵ pode ser escrita fazendo uso da expansão dada na seção 11.3.2, (v. eq. (11.29)). Como, no entanto, das 16 matrizes linearmente independentes ali consideradas, apenas as seis matrizes $\sigma^{\mu\nu}$ dão resultados proporcionais às próprias matrizes γ^μ quando comutadas com elas, isto é,

$$[\gamma^\lambda, \sigma^{\mu\nu}] = 2i \left(g^{\lambda\mu} \gamma^\nu - g^{\lambda\nu} \gamma^\mu \right),$$

é desde logo claro que apenas essas matrizes intervirão na expansão, que pode por isso ser escrita

$$\epsilon = c_{\mu\nu} \sigma^{\mu\nu} \quad \text{com} \quad c_{\mu\nu} = \frac{1}{4i} g_{\mu\lambda} \left(L^\lambda{}_\nu + R^\lambda{}_\nu \right) = \frac{1}{4i} (L_{\mu\nu} + R_{\mu\nu}).$$

Note, em particular, que da anti-simetria de $L_{\mu\nu}$ e de $R_{\mu\nu}$ resulta também que $c_{\mu\nu} = -c_{\nu\mu}$, consistentemente com $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$. A transformação S infinitesimal pode portanto ser escrita como

$$\mathbf{S} = 1_4 - \frac{i}{4} (\mathbf{L}_{\mu\nu} + \mathbf{R}_{\mu\nu}) \sigma^{\mu\nu} + \mathcal{O}(\epsilon^2). \quad (11.34)$$

Essa expressão pode ser especializada para transformações de Lorentz “puras” anulando a parte correspondente às rotações espaciais, $\mathbf{R}_{\mu\nu} = 0$. Nesse caso, lembrando que os elementos não nulos de $\mathbf{L}_{\mu\nu}$ são $\mathbf{L}_{0j} = -\mathbf{L}_{j0} = -\delta v_j/c$, $j = 1, 2, 3$,

$$\mathbf{S}_{\text{Lor}} \simeq 1_4 + \frac{i}{2} \left(\frac{\delta v_j}{c} \sigma^{0j} \right) = 1_4 + \frac{i}{2} \left(\frac{\delta v_j}{c} \frac{i}{2} (\gamma^0 \gamma^j - \gamma^j \gamma^0) \right) = 1_4 - \frac{|\delta \vec{v}|}{2c} \hat{v} \cdot \vec{\alpha}$$

com $\vec{\alpha} = \{\gamma^0 \gamma^j\}$ e \hat{v} sendo um vetor unitário com a direção de $\delta \vec{v}$. Isso mostra que a matriz $\hat{v} \cdot \vec{\alpha}$, sendo $\{\vec{\alpha}\}$ as matrizes introduzidas para escrever a equação de Dirac sob forma hamiltoniana, eq (11.19), é (a menos de um fator numérico 1/2) o gerador da mudança de representação \mathbf{S}_{Lor} associada a transformações de Lorentz com velocidade na direção \hat{v} . A mudança de representação que corresponde a uma transformação de Lorentz finita, caracterizada pela velocidade \vec{v} , é então dada por

$$\mathbf{S}_{\text{Lor}} = e^{-\frac{1}{2} \varphi \hat{v} \cdot \vec{\alpha}}$$

com $\tanh \varphi = |\vec{v}|/c$. Essa transformação tem claramente uma inversa, que é dada por uma das formas equivalentes

$$\mathbf{S}_{\text{Lor}}^{-1} = e^{\frac{1}{2} \varphi \hat{v} \cdot \vec{\alpha}} = \gamma^0 \mathbf{S}_{\text{Lor}} \gamma^0 = \gamma^0 \mathbf{S}_{\text{Lor}}^\dagger \gamma^0.$$

É também importante analisar o caso de uma rotação espacial “pura”, fazendo $\mathbf{L}_{\mu\nu} = 0$ em (11.34). Este caso permitirá identificar os geradores da mudança de representação \mathbf{S}_{Rot} associada a esse tipo de transformação. Em vista da relação geral existente entre geradores de rotações e momento angular, e dado que a presente análise exclui a dependência espacio-temporal $e^{-ip \cdot x}$ das soluções elementares da equação de Dirac, esses geradores podem ser vistos como associados a uma componente *intrínseca* (spin) do momento angular da partícula de Dirac.

Usando os resultados obtidos na equação (11.30) para as rotações infinitesimais, e tendo em conta a forma das matrizes $\vec{\lambda}$, os $\mathbf{R}_{\mu\nu}$ não nulos podem ser identificados como sendo dados por

$$\mathbf{R}_{23} = -\mathbf{R}_{32} = -\delta\omega_1, \quad \mathbf{R}_{13} = -\mathbf{R}_{31} = \delta\omega_2 \quad \text{e} \quad \mathbf{R}_{12} = -\mathbf{R}_{21} = -\delta\omega_3$$

onde os $\delta\omega_j$ representam rotações infinitesimais em torno do eixo j . Com esses valores a relação (11.34) dá, a menos de termos de segunda ordem,

$$\mathbf{S}_{\text{Rot}} \simeq 1_4 + \frac{i}{2} (\delta\omega_1 \sigma^{23} + \delta\omega_2 \sigma^{31} + \delta\omega_3 \sigma^{12}).$$

Essa expressão pode ser re-escrita de uma forma mais compacta e sugestiva definindo três novas matrizes $\underline{\sigma} = \{\underline{\sigma}_1, \underline{\sigma}_2, \underline{\sigma}_3\}$ como

$$\underline{\sigma}_j \equiv \frac{1}{2} \epsilon_{jkl} \sigma^{kl}, \quad j, k, l = 1, 2, 3 \quad (11.35)$$

onde ϵ_{jkl} é o símbolo anti-simétrico em três dimensões e se subentende a soma sobre os índices repetidos k e l . Note que o fator $1/2$ juntamente com a propriedade de anti-simetria das matrizes $\sigma^{\mu\nu}$, isto é, $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$, permitem escrever essa definição simplesmente como

$$\underline{\sigma}_1 = \sigma^{23}, \quad \underline{\sigma}_2 = \sigma^{31} \quad \text{e} \quad \underline{\sigma}_3 = \sigma^{12}.$$

É fácil verificar explicitamente, usando as relações de anti-comutação das matrizes γ^μ juntamente com as várias definições envolvidas, que $\underline{\sigma}_j^2 = 1_4$, e ainda que essas três matrizes satisfazem relações de comutação iguais às que valem para as matrizes de Pauli, isto é

$$[\underline{\sigma}_j, \underline{\sigma}_k] = 2i\epsilon_{jkl} \underline{\sigma}_l \quad (11.36)$$

Usando a definição das matrizes $\underline{\sigma}$ é possível escrever

$$S_{\text{Rot}} \simeq 1_4 + \frac{i\delta\omega}{2} (\hat{u} \cdot \underline{\sigma})$$

onde \hat{u} é um vetor unitário na direção do eixo da rotação infinitesimal $\delta\omega$, e o gerador da mudança de representação associada a uma rotação infinitesimal em torno de \hat{u} fica identificado como sendo $\hat{u} \cdot \underline{\sigma}/2$. A mudança de representação que corresponde a uma rotação *finita* ω em torno do eixo \hat{u} é portanto dada por

$$R_{\text{Rot}} = e^{\frac{i\omega}{2} \hat{u} \cdot \underline{\sigma}}.$$

Como no caso de transformações de Lorentz puras, essa transformação tem também uma inversa, que pode ser escrita sob as formas equivalentes

$$R_{\text{Rot}}^{-1} = e^{-\frac{i\omega}{2} \hat{u} \cdot \underline{\sigma}} = R_{\text{Rot}}^\dagger = \gamma^0 R_{\text{Rot}}^\dagger \gamma^0.$$

A utilidade desta última forma, algo redundante para o caso de rotações puras, vem de que ela vale igualmente para as transformações de Lorentz puras, consideradas anteriormente.

É fácil verificar explicitamente que as matrizes de dimensão 4 $\underline{\alpha}$ e $\underline{\sigma}$ *não* comutam. Isso significa que, em geral, transformações de Lorentz finitas e rotações finitas *não* comutam. Usando a representação particular (11.18) para as matrizes de Dirac, um cálculo simples mostra que, nesse caso, as matrizes $\underline{\sigma}$, de dimensão 4, que devem ser associadas ao spin da partícula de Dirac são dadas por

$$\underline{\vec{\sigma}} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \quad (11.37)$$

onde $\vec{\sigma}$ são as matrizes de Pauli, de dimensão 2, e 0 representa a matriz nula, também de dimensão 2. Usando essa mesma representação, as matrizes $\vec{\alpha}$ são

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}. \quad (11.38)$$

O uso desses resultados pode ser ilustrado com o cálculo das propriedades de transformação da quadri-corrente $\{j^\mu\}$ definida na seção anterior, eqs. (11.27) e (11.28). Essas duas equações podem ser resumidas pondo

$$j^\mu(x) = \bar{\psi}(x) c \gamma^\mu \psi(x)$$

onde $\bar{\psi}(x) = \psi^\dagger(x) \gamma^0$ é a *adjunta* da solução $\psi(x)$ da equação de Dirac. A corrente transformada será então

$$j'^\mu(x') = \bar{\psi}'(x') c \gamma^\mu \psi'(x')$$

que, usando a relação (11.33) e o fato de que $(\gamma^0)^2 = 1$, pode ser re-escrita como

$$j'^\mu(x') = \psi^\dagger(x) \gamma^0 \gamma^0 \mathbf{S}^\dagger \gamma^0 c \gamma^\mu \mathbf{S} \psi(x) = \bar{\psi}(x) \mathbf{S}^{-1} c \gamma^\mu \mathbf{S} \psi(x).$$

Usando agora a relação de definição da transformação \mathbf{S} , eq. (11.31) se obtém

$$j'^\mu(x') = \bar{\psi}(x) \mathbf{T}^\mu_\nu c \gamma^\nu \psi(x) = \mathbf{T}^\mu_\nu j^\nu(x)$$

mostrando que as componentes da corrente se transformam como as componentes de um quadri-vetor.

De uma forma inteiramente análoga é possível mostrar também que as quantidades $\bar{\psi}(x) \psi(x)$ e $\bar{\psi}(x) \sigma^{\mu\nu} \psi(x)$ se transformam respectivamente como um escalar e como um tensor de segunda ordem. As quantidades $\bar{\psi}(x) \gamma_5 \gamma^\mu \psi(x)$ e $\bar{\psi}(x) \gamma_5 \psi(x)$, por outro lado, também se transformam como componentes de um quadri-vetor e como um escalar sob transformações próprias (sem inversão espacial). Quando inversões espaciais são permitidas, no entanto, contrariamente ao que acontece com as quantidades análogas em que não intervém a matriz γ_5 , elas ganham um fator adicional ± 1 proveniente do determinante da transformação considerada, e são por isso chamadas pseudo-(quadri)vectores e pseudo-escalares, respectivamente.

11.3.4 Inversão espacial.

Outra transformação relevante, que pode ser tratada de forma análoga à usada para as transformações de Lorentz e rotações no que diz respeito às matrizes de Dirac, é a inversão espacial, definida explicitamente por

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

ou seja,

$$x'^\mu = I^\mu_\nu x^\nu \quad \text{com} \quad I^0_0 = 1, \quad I^j_j = -1, \quad j = 1, 2, 3 \quad \text{e} \quad I^\mu_{\nu \neq \mu} = 0.$$

Se as matrizes γ^μ são tratadas como componentes de um quadrivetor, elas devem se transformar sob essa operação de acordo com

$$\gamma'^0 = \gamma^0, \quad \gamma'^j = -\gamma^j, \quad j = 1, 2, 3$$

e, para manter a sua forma inalterada, é necessário introduzir uma mudança de representação \mathbf{S} satisfazendo as relações (11.32), que neste caso são

$$\mathbf{S}^{-1} \gamma^0 \mathbf{S} = \gamma^0 \quad \text{e} \quad \mathbf{S}^{-1} \gamma^j \mathbf{S} = -\gamma^j, \quad j = 1, 2, 3.$$

Em vista das propriedades gerais das matrizes γ^μ , escolhas possíveis para \mathbf{S} satisfazendo essas condições são

$$\mathbf{S} = \mathbf{S}^{-1} = \gamma^0 \quad \text{ou} \quad \mathbf{S} = \mathbf{S}^{-1} = -\gamma^0. \quad (11.39)$$

A duplicidade de sinal está relacionada com a paridade intrínseca (v. seção 9.5), e é irrelevante num contexto em que o número de partículas seja fixo (em particular, quando se trate de uma única partícula). Desse modo, o efeito de uma inversão espacial sobre uma função de onda (com quatro componentes) para uma partícula de Dirac poderá ser representado por

$$\psi(\vec{r}, t) \longrightarrow \gamma^0 R_P \psi(\vec{r}, t) = \gamma^0 \psi(-\vec{r}, t)$$

onde $R_P = R_P^{-1}$ representa o operador que troca o sinal das componentes espaciais dos quadri-vetores. É imediato verificar que o hamiltoniano de Dirac é invariante sob essa transformação, isto é

$$\gamma^0 R_P H R_P \gamma^0 = \gamma^0 R_P (c \gamma^0 \vec{\gamma} \cdot \vec{p} + \gamma^0 M c^2) R_P \gamma^0 = H$$

o que significa que suas autofunções podem ser escolhidas de forma a serem autofunções simultâneas do operador paridade $\mathcal{P} \equiv \gamma^0 R_P$.

11.3.5 Momento angular na equação de Dirac livre.

O hamiltoniano não relativístico de uma partícula livre com spin é um operador invariante sob rotações que age sobre as variáveis de spin como o operador identidade

$$H_{\text{nr}} = \frac{\vec{p}^2}{2M} \otimes \hat{1}_S.$$

Em conseqüência disso, tanto as componentes do momento angular intrínseco (spin) \vec{S} como as do momento angular orbital \vec{L} comutam com H . O hamiltoniano de Dirac livre, embora seja igualmente um operador invariante sob rotações, apresenta no entanto uma estrutura menos trivial com relação às diferentes componentes do estado cuja transformação sob rotações deve ser associada, como observado na seção anterior, a uma componente intrínseca do momento angular, e isso modifica o resultado não relativístico de uma forma importante. De fato, calculando o comutador do hamiltoniano livre de Dirac

$$H = c\vec{\alpha} \cdot \vec{p} + \beta Mc^2$$

com o momento angular orbital $\vec{L} = \vec{r} \times \vec{p}$, o que se obtém é o resultado *não nulo*

$$[H, \vec{L}] = [c\vec{\alpha} \cdot \vec{p}, \vec{r} \times \vec{p}] = -i\hbar c \vec{\alpha} \times \vec{p} \quad (11.40)$$

que mostra que as componentes de \vec{L} , e portanto também L^2 , *não* são compatíveis com H .

Por outro lado, é possível calcular também as relações de comutação de H com as matrizes $\underline{\sigma}$, obtidas na seção anterior como geradores dos efeitos de uma rotação sobre as matrizes de Dirac. Como a matriz $\beta \equiv \gamma^0$ comuta com as matrizes σ^{jk} , $j, k = 1, 2, 3$ que intervêm na construção das matrizes $\underline{\sigma}$, o cálculo dessas relações de comutação se reduz na realidade ao cálculo de

$$[H, \underline{\sigma}_j] = c\vec{p} \cdot [\vec{\alpha}, \underline{\sigma}_j].$$

O ingrediente básico a ser estudado é portanto o comutador das componentes de $\vec{\alpha}$ com as de $\underline{\sigma}$

$$[\alpha_k, \underline{\sigma}_j] = [\gamma^0 \gamma^k, \frac{1}{2} \epsilon_{lmj} \sigma^{lm}] = \frac{i}{4} \epsilon_{lmj} [\gamma^0 \gamma^k, (\gamma^l \gamma^m - \gamma^m \gamma^l)] = \frac{i}{2} \epsilon_{lmj} [\gamma^0 \gamma^k, \gamma^l \gamma^m]$$

onde $k, l, m = 1, 2, 3$. Os comutadores envolvendo quatro matrizes de Dirac podem ser calculados usando as relações de anti-comutação satisfeitas por essas matrizes

$$\begin{aligned}
[\gamma^0 \gamma^k, \gamma^l \gamma^m] &= \gamma^0 \gamma^k \gamma^l \gamma^m - \gamma^l \gamma^m \gamma^0 \gamma^k = \gamma^0 (\gamma^k \gamma^l \gamma^m - \gamma^l \gamma^m \gamma^k) = \\
&= \gamma^0 (\{\gamma^k, \gamma^l\} \gamma^m - \gamma^l \gamma^k \gamma^m - \gamma^l \{\gamma^m, \gamma^k\} + \gamma^l \gamma^k \gamma^m) = \\
&= 2\gamma^0 (g^{kl} \gamma^m - g^{mk} \gamma^l)
\end{aligned}$$

Desse resultado segue então que

$$[\alpha_k, \underline{\sigma}_j] = i\epsilon_{lmj} \gamma^0 (g^{kl} \gamma^m - g^{mk} \gamma^l) = i\epsilon_{lmj} (\delta_{mk} \alpha_l - \delta_{kl} \alpha_m)$$

e finalmente, fazendo a soma sobre os índices l e m ,

$$[H, \underline{\sigma}_j] = ic\epsilon_{lmj} (\alpha_l p_m - \alpha_m p_l) = 2ic (\vec{\alpha} \times \vec{p})_j. \quad (11.41)$$

Portanto as matrizes $\underline{\sigma}$, proporcionais aos geradores dos efeitos de rotações sobre a estrutura intrínseca das soluções elementares da equação de Dirac, também *não* são compatíveis com H . Comparando as equações (11.40) e (11.41), no entanto, é evidente que

$$\left[H, \vec{L} + \frac{\hbar}{2} \underline{\sigma} \right] = 0$$

isto é, o operador

$$\vec{J} = \vec{r} \times \vec{p} + \frac{\hbar}{2} \underline{\sigma} \equiv \vec{L} + \vec{S}$$

é compatível com H . Em vista das relações de comutação (11.36), as componentes do segundo termo satisfazem as mesmas relações de comutação padrão de momento angular que as componentes do momento angular orbital. Esse operador pode portanto ser identificado como o *momento angular total* da partícula de Dirac, e fornece as duas constantes de movimento usuais, mutuamente compatíveis, J^2 e J_3 . Finalmente, do fato de que

$$S^2 = \frac{\hbar^2}{4} (\underline{\sigma}_1^2 + \underline{\sigma}_2^2 + \underline{\sigma}_3^2) = \frac{3\hbar^2}{4} 1_4 = \hbar^2 \frac{1}{2} \left(\frac{1}{2} + 1 \right) 1_4$$

segue que S^2 é uma constante de movimento, cujo valor mostra que a equação de Dirac descreve partículas de spin $1/2$.

11.4 Propriedades eletromagnéticas de uma partícula de Dirac carregada.

Como a equação de Dirac descreve partículas de spin 1/2, no caso em que a partícula seja carregada ela pode em particular ser dotada de um momento de dipolo magnético. Dado que o spin da partícula está de fato incorporado na estrutura geral da equação, deve-se esperar que uma forma de examinar essa questão seja submeter a partícula a um campo eletromagnético de prova, externo, fazendo isso simplesmente através do acoplamento mínimo, e estudar a natureza da interação da partícula com esse campo.

O acoplamento mínimo de uma partícula de Dirac com carga elétrica e com um campo eletromagnético externo descrito em termos do potencial quadrivetor A_μ é efetuado através da substituição

$$\partial_\mu \longrightarrow \partial_\mu + i\frac{e}{\hbar c}A_\mu \quad \text{ou} \quad i\partial_\mu \longrightarrow i\partial_\mu - \frac{e}{\hbar c}A_\mu$$

ou seja, explicitando a componente temporal e as componentes espaciais e introduzindo em cada caso as constantes multiplicativas apropriadas,

$$i\hbar\frac{\partial}{\partial t} \longrightarrow i\hbar\frac{\partial}{\partial t} - e\phi \quad \text{e} \quad \frac{\hbar}{i}\frac{\partial}{\partial x^k} \longrightarrow \frac{\hbar}{i}\frac{\partial}{\partial x^k} - \frac{e}{c}A^k.$$

Note, com relação às componentes espaciais, que $A^k = -A_k$. Com essa substituição a equação de Dirac fica (cf. eq. (11.25))

$$\left(i\gamma^\mu\partial_\mu - \frac{e}{\hbar c}\gamma^\mu A_\mu - \frac{Mc}{\hbar}\right)\psi(x) = 0. \quad (11.42)$$

Uma forma simples de explicitar os termos de interação envolvendo os campos externos \vec{E} e \vec{B} consiste em transformar essa equação numa equação de segunda ordem, aplicando a ela o operador

$$i\gamma^\mu\partial_\mu - \frac{e}{\hbar c}\gamma^\mu A_\mu + \frac{Mc}{\hbar}$$

o que dá a equação

$$\left[\left(i\gamma^\mu\partial_\mu - \frac{e}{\hbar c}\gamma^\mu A_\mu\right)^2 - \frac{M^2c^2}{\hbar^2}\right]\psi(x) = 0.$$

É claro que todas as soluções da equação de primeira ordem (11.42) são também soluções desta equação, que pode no entanto admitir soluções adicionais. Para explicitar nela os termos de acoplamento eletromagnéticos basta efetuar as transformações

$$\begin{aligned}
\left(i\gamma^\mu\partial_\mu - \frac{e}{\hbar c}\gamma^\mu A_\mu\right)^2 &= \gamma^\mu\gamma^\nu \left(i\partial_\mu - \frac{e}{\hbar c}A_\mu\right) \left(i\partial_\nu - \frac{e}{\hbar c}A_\nu\right) = \\
&= (g^{\mu\nu} - i\sigma^{\mu\nu}) \left(i\partial_\mu - \frac{e}{\hbar c}A_\mu\right) \left(i\partial_\nu - \frac{e}{\hbar c}A_\nu\right) = \\
&= \left(i\partial_\mu - \frac{e}{\hbar c}A_\mu\right) \left(i\partial^\mu - \frac{e}{\hbar c}A^\mu\right) - \frac{i}{2}\sigma^{\mu\nu} \left[i\partial_\mu - \frac{e}{\hbar c}A_\mu, i\partial_\nu - \frac{e}{\hbar c}A_\nu\right]
\end{aligned}$$

nas quais foi usada a relação

$$\gamma^\mu\gamma^\nu = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\} + \frac{1}{2}[\gamma^\mu, \gamma^\nu] = g^{\mu\nu} - i\sigma^{\mu\nu}$$

e também a antissimetria das matrizes $\sigma^{\mu\nu}$ para escrever

$$\sigma^{\mu\nu} = \frac{1}{2}(\sigma^{\mu\nu} - \sigma^{\nu\mu})$$

que permite exprimir o termo correspondente da equação sob a forma de um comutador. Os A_μ são de fato operadores de multiplicação, de modo que o comutador pode ser reduzido aplicando-o a uma função genérica ϕ com o resultado

$$\left[i\partial_\mu - \frac{e}{\hbar c}A_\mu, i\partial_\nu - \frac{e}{\hbar c}A_\nu\right]\phi = -i\frac{e}{\hbar c}(\partial_\mu A_\nu - \partial_\nu A_\mu)\phi = -i\frac{e}{\hbar c}F_{\mu\nu}\phi$$

onde $F_{\mu\nu}$ é o tensor do campo eletromagnético, dado explicitamente em termos das componentes de \vec{E} e \vec{B} em (11.10). A dependência desses dois campos pode ser explicitada escrevendo

$$\sigma^{\mu\nu}F_{\mu\nu} = \sigma^{jk}F_{jk} + \sigma^{0k}F_{0k} + \sigma^{k0}F_{k0} = -2\vec{\sigma} \cdot \vec{B} + 2i\vec{\alpha} \cdot \vec{E}.$$

Nessa expressão os índices j e k assumem os valores 1,2 e 3, e foi feito uso da definição (11.35) para as três componentes da matriz $\vec{\sigma}$, da anti-simetria do tensor do campo eletromagnético e ainda da relação

$$\sigma^{0k} - \sigma^{k0} = 2i\gamma^0\gamma^k = 2i\alpha_k.$$

Desse modo a equação de segunda ordem pode ser escrita como

$$\left[\left(i\partial^\nu - \frac{e}{\hbar c}A^\nu\right) \left(i\partial_\nu - \frac{e}{\hbar c}A_\nu\right) + \frac{e}{\hbar c} \left(\vec{\sigma} \cdot \vec{B} - i\vec{\alpha} \cdot \vec{E}\right) - \frac{M^2c^2}{\hbar^2}\right]\psi(x) = 0. \quad (11.43)$$

Utilizando a representação particular (11.18) das matrizes de Dirac, as matrizes $\underline{\vec{\sigma}}$ e $\vec{\alpha}$ são dadas respectivamente por (11.37) e (11.38). Nesse caso, portanto, as duas componentes superiores e as duas componentes inferiores de $\psi(x)$ são acopladas *apenas* pelo termo envolvendo $\vec{\alpha} \cdot \vec{E}$.

O termo envolvendo $\underline{\vec{\sigma}} \cdot \vec{B}$ representa claramente uma interação do spin da partícula com o campo magnético externo pois, como visto na seção anterior, o spin da partícula é representado por $\vec{S} = \hbar \underline{\vec{\sigma}}/2$. Ele indica portanto a existência de um momento magnético associado ao spin, cujo valor pode ser determinado a partir da energia de interação associada a esse termo. Uma forma de identificar tal energia de interação consiste em considerar uma solução estacionária da equação de segunda ordem, cuja dependência temporal é da forma $e^{-i\omega t}$, no limite não relativístico, caracterizado por

$$E = \hbar\omega = Mc^2 + \epsilon, \quad \text{com} \quad \epsilon \ll Mc^2$$

e de campos externos fracos, no sentido que $eA^\mu \ll Mc^2$. Nessas condições

$$\begin{aligned} \left(i\partial_0 - \frac{e}{\hbar c}A_0\right)^2 \psi(x) &= \left(\frac{\omega}{c} - \frac{e}{\hbar c}\phi\right)^2 \psi(x) = \left(\frac{Mc^2 + \epsilon}{\hbar c} - \frac{e}{\hbar c}\phi\right)^2 \psi(x) \simeq \\ &\simeq \left(\frac{M^2c^2}{\hbar^2} + \frac{2M\epsilon}{\hbar^2} - \frac{2Me}{\hbar^2}\phi\right) \psi(x) \end{aligned}$$

enquanto os termos envolvendo as componentes espaciais podem ser escritos sob a forma

$$\left(i\partial^k - \frac{e}{\hbar c}A^k\right) \left(i\partial_k - \frac{e}{\hbar c}A_k\right) \psi(x) = -\left(\frac{\vec{p}}{\hbar} - \frac{e}{\hbar c}\vec{A}\right)^2 \psi(x).$$

Levando essas expressões à equação (11.43) e multiplicando o resultado por $-\hbar^2/2M$ resulta

$$\left[\frac{\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2}{2M} + e\phi - \frac{e\hbar}{2Mc}(\underline{\vec{\sigma}} \cdot \vec{B} - i\vec{\alpha} \cdot \vec{E})\right] \psi(x) \simeq \epsilon\psi(x). \quad (11.44)$$

Nesta última equação, a energia de interação do spin com o campo magnético externo é claramente representada pelo termo

$$-\frac{e\hbar}{2Mc}\underline{\vec{\sigma}} \cdot \vec{B} \equiv -\vec{\mu} \cdot \vec{B}$$

que identifica o momento magnético $\vec{\mu}$ associado ao spin como sendo

$$\vec{\mu} = \frac{e\hbar}{2Mc}\underline{\vec{\sigma}} \equiv g\left(\frac{e}{2Mc}\right)\vec{S}.$$

O valor do fator giromagnético g associado ao spin da partícula de Dirac é portanto $g = 2$.

É importante observar que, ao contrário da equação (11.43), que é uma equação de segunda ordem obtida aplicando um operador diferencial sobre a equação de Dirac *sem qualquer aproximação*, a equação (11.44) contém aproximações nas quais foram desprezados termos envolvendo ϵ^2 , $\epsilon e\phi$ e $e^2\phi^2$ frente a termos que dependem linearmente de ϵ e de $e\phi$. Em particular, essas aproximações foram essenciais para que a equação (11.44) assumisse a forma de uma equação de autovalores. Esta, no entanto, contém um termo não hermiteano, proporcional a $i\vec{\alpha} \cdot \vec{E}$. Tendo em vista a interpretação 11.22) das matrizes $\vec{\alpha}$, este termo depende explicitamente da velocidade da partícula, ao contrário do que acontece com o termo de interação do spin com o campo magnético externo. Para ter uma idéia do seu conteúdo, convém utilizar a representação particular (11.18), na qual a equação (11.44), supondo por simplicidade $\vec{B} = 0$ com $\vec{A} = 0$ e separando as componentes “grandes” U e “pequenas” V , fica escrita como

$$\begin{pmatrix} \frac{p^2}{2M} + e\phi & \frac{ie\hbar}{2Mc} \vec{\sigma} \cdot \vec{E} \\ \frac{ie\hbar}{2Mc} \vec{\sigma} \cdot \vec{E} & \frac{p^2}{2M} + e\phi \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} = \epsilon \begin{pmatrix} U \\ V \end{pmatrix}.$$

O termo envolvendo $\vec{\sigma} \cdot \vec{E}$ acopla esses dois grupos de componentes, como mostrado explicitamente pela equação

$$\left(\frac{p^2}{2M} + e\phi \right) U + \frac{ie\hbar}{2Mc} \vec{\sigma} \cdot \vec{E} V = \epsilon U.$$

Essa equação pode ser reduzida a uma equação de duas componentes para U , apenas, utilizando a relação entre U e V que resulta da equação estacionária de primeira ordem, escrita explicitamente nessa mesma representação como

$$\begin{pmatrix} e\phi + Mc^2 & c\vec{\sigma} \cdot \vec{p} \\ c\vec{\sigma} \cdot \vec{p} & e\phi - Mc^2 \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} = E \begin{pmatrix} U \\ V \end{pmatrix}.$$

De fato, dessa equação resulta em particular que

$$V = \frac{c\vec{\sigma} \cdot \vec{p}}{E + Mc^2 - e\phi} U \simeq \frac{c\vec{\sigma} \cdot \vec{p}}{2Mc^2} U$$

onde foi feito uso de que $E = Mc^2 + \epsilon$ e de que ϵ e $e\phi$ são pequenos no limite não relativístico e de campo fraco. Com isso a equação para as componentes U fica

$$\left(\frac{p^2}{2M} + e\phi \right) U + \frac{ie\hbar}{4M^2c^2} (\vec{\sigma} \cdot \vec{E})(\vec{\sigma} \cdot \vec{p}) U = \epsilon U.$$

Utilizando a identidade $(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) \equiv \vec{A} \cdot \vec{B} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B})$ (v. problema **6.2**) o segundo termo do lado esquerdo dessa equação pode ser transformado em

$$\frac{ie\hbar}{4M^2c^2}(\vec{\sigma} \cdot \vec{E})(\vec{\sigma} \cdot \vec{p}) = \frac{ie\hbar}{4M^2c^2} \vec{E} \cdot \vec{p} - \frac{e\hbar}{4M^2c^2} \vec{\sigma} \cdot (\vec{E} \times \vec{p}).$$

O primeiro dos dois termos resultantes é um sub-produto incompleto, não hermiteano, da particular redução não relativística que foi utilizada, relacionado com o termo de Darwin (cf. eq. (6.16)). O segundo termo, por outro lado, representa uma interação spin-órbita. Ele pode ser posto numa forma mais familiar particularizando o potencial escalar externo como sendo um potencial central $\phi(r)$, que corresponde ao campo eléctrico

$$\vec{E} = -\vec{\nabla}\phi(r) = -\frac{1}{r} \frac{d\phi}{dr} \vec{r}.$$

Nesse caso, de fato, ele se reduz a

$$-\frac{e\hbar}{4M^2c^2} \vec{\sigma} \cdot (\vec{E} \times \vec{p}) \longrightarrow \frac{e\hbar}{4M^2c^2r} \left(\frac{d\phi}{dr} \right) \vec{\sigma} \cdot (\vec{r} \times \vec{p}) = \frac{1}{2M^2c^2r} \left(e \frac{d\phi}{dr} \right) \vec{S} \cdot \vec{L}$$

onde $\vec{S} = \hbar\vec{\sigma}/2$. Essa forma do termo de interação spin-órbita reproduz a equação (6.12), e inclui portanto os efeitos da precessão de Thomas. Ele corresponde, de fato, à interação, com o campo magnético gerado pelo movimento do eléctron no campo eléctrico central, de um momento magnético

$$\vec{\mu}_{LS} = g_{LS} \left(\frac{e}{2Mc} \right) \vec{S}$$

com $g_{LS} = 1$, e não 2, como no caso da interação com o campo magnético externo.

11.5 Estados ligados de um eléctron de Dirac em um campo coulombiano externo, atrativo.

O modelo mais simples que pode ser feito no contexto da equação de Dirac para os estados ligados de um átomo hidrogenóide (átomo de um eléctron cujo núcleo tem carga positiva Ze) consiste tratar o hamiltoniano de Dirac num campo externo, representado pelo potencial escalar $\phi(r) = Ze/r$ e com o potencial vetor nulo, $\vec{A} \equiv 0$. Desse modo a equação de autovalores a ser estudada é

$$\left[\frac{\hbar c}{i} \vec{\alpha} \cdot \vec{\nabla} + \beta m_e c^2 + v(r) \right] \psi_E(\vec{r}) = E \psi_E(\vec{r})$$

com $v(r) \equiv -e\phi(r) = -Ze^2/r$ e $E < m_e c^2$. Essa equação representa na realidade um conjunto de quatro equações de derivadas parciais acopladas para as quatro componentes das funções de onda $\psi_E(\vec{r})$. O operador diferencial que ela contém é simplesmente o hamiltoniano de Dirac escrito na forma usada na equação (11.19), explicitando a representação de Schrödinger do operador momento, $\vec{p} = \hbar \vec{\nabla}/i$ e introduzindo o potencial escalar externo $\phi(r)$ que interage com o elétron de carga $-e$ e massa m_e .

11.5.1 Simetrias e funções angulares.

De modo semelhante ao usado no caso da equação não relativística, é possível usar constantes do movimento do hamiltoniano para reduzir essas equações a um sistema de equações ordinárias acopladas, envolvendo a coordenada radial apenas. As constantes do movimento relevantes são, neste caso, o momento angular total (note que o potencial $V(r)$ é invariante sob rotações)

$$\vec{J} = \vec{r} \times \vec{p} + \frac{\hbar}{2} \vec{\sigma}$$

e a paridade (o potencial $V(r)$ sendo também invariante sob uma inversão espacial)

$$\mathcal{P} = \gamma^0 R_P = \beta R_P.$$

De fato, usando a representação particular (11.18) para as matrizes de Dirac, a equação de autovalores adquire a forma

$$\left[\frac{\hbar c}{i} \begin{pmatrix} 0 & \vec{\sigma} \cdot \vec{\nabla} \\ \vec{\sigma} \cdot \vec{\nabla} & 0 \end{pmatrix} + \begin{pmatrix} m_e c^2 & 0 \\ 0 & -m_e c^2 \end{pmatrix} + \begin{pmatrix} v(r) & 0 \\ 0 & v(r) \end{pmatrix} \right] \begin{pmatrix} U \\ V \end{pmatrix} = E \begin{pmatrix} U \\ V \end{pmatrix} \quad (11.45)$$

onde U e V contém respectivamente as duas componentes superiores e inferiores de $\psi_E(\vec{r})$. As autofunções com momento angular e paridade bem definidos

$$\psi_{Ejm}(\vec{r}) \equiv \begin{pmatrix} U_{jm}(\vec{r}) \\ V_{jm}(\vec{r}) \end{pmatrix}$$

devem portanto satisfazer

$$J^2 \psi_{Ejm}(\vec{r}) = \hbar^2 j(j+1) \psi_{Ejm}(\vec{r}), \quad J_3 \psi_{Ejm}(\vec{r}) = \hbar m \psi_{Ejm}(\vec{r})$$

e também

$$\mathcal{P} \psi_{Ejm}(\vec{r}) = \beta \psi_{Ejm}(-\vec{r}) = \pm \psi_{Ejm}(\vec{r})$$

ou seja, em termos de $U_{jm}(\vec{r})$ e $V_{jm}(\vec{r})$,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} U_{jm}(-\vec{r}) \\ V_{jm}(-\vec{r}) \end{pmatrix} = \begin{pmatrix} U_{jm}(-\vec{r}) \\ -V_{jm}(-\vec{r}) \end{pmatrix} = \pm \begin{pmatrix} U_{jm}(\vec{r}) \\ V_{jm}(\vec{r}) \end{pmatrix}$$

o que significa que $U_{jm}(\vec{r})$ e $V_{jm}(\vec{r})$ devem ter momento angular bem definido e paridades opostas.

A parte angular dessas funções de onda pode ser implementada em termos das funções angulares de duas componentes

$$\mathcal{Y}_{jm}^{\pm}(\hat{r}, s) \equiv \sum_{m_s} C_{m-m_s}^{j \pm \frac{1}{2}} \frac{1}{2} \frac{j}{m_s} Y_{j \pm \frac{1}{2}}^{m-m_s}(\hat{r}) \chi_{m_s}(s)$$

onde $s = \pm 1/2$ é uma variável de spin, $\chi_{m_s}(s) = \delta_{m_s, s}$ são funções usuais de spin $1/2$ e os coeficientes C são coeficientes de Clebsch-Gordan. Note que j é *semi-inteiro*. Essas funções angulares de duas componentes podem ser escritas explicitamente como

$$\mathcal{Y}_{jm}^{\pm}(\hat{r}) = \begin{pmatrix} C_{m-\frac{1}{2}}^{j \pm \frac{1}{2}} \frac{1}{2} \frac{j}{m} Y_{j \pm \frac{1}{2}}^{m-\frac{1}{2}}(\hat{r}) \\ C_{m+\frac{1}{2}}^{j \pm \frac{1}{2}} \frac{1}{2} \frac{j}{m} Y_{j \pm \frac{1}{2}}^{m+\frac{1}{2}}(\hat{r}) \end{pmatrix}$$

onde as funções de spin foram representadas por

$$\chi_{\frac{1}{2}}(s) \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{e} \quad \chi_{-\frac{1}{2}}(s) \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Os $\mathcal{Y}_{jm}^{\pm}(\hat{r}, s)$ são claramente autofunções simultâneas de J^2 , J_3 , L^2 e S^2 , com autovalores dados respectivamente por $\hbar^2 j(j+1)$, $\hbar m$, $\hbar^2(j \pm 1/2)(j \pm 1/2 + 1)$ e $\hbar^2 3/4$. As componentes do spin \vec{S} são representadas da forma usual por matrizes de Pauli (de dimensão 2). A paridade de $\mathcal{Y}_{jm}^{\pm}(\hat{r}, s)$ é dada então por $(-1)^{j \pm 1/2}$, de modo que para um dado valor de j os dois tipos (\pm) de função angular têm paridades opostas. Existem portanto dois possíveis *ansätze* para as funções de onda com quatro componentes e com momento angular total bem definido. Eles são

$$\text{Tipo 1 : } \psi_{jm}^{(1)} = \begin{pmatrix} \mathcal{Y}_{jm}^{-}(\hat{r}, s) g_1(r) \\ \mathcal{Y}_{jm}^{+}(\hat{r}, s) i f_1(r) \end{pmatrix}; \quad \text{Tipo 2 : } \psi_{jm}^{(2)} = \begin{pmatrix} \mathcal{Y}_{jm}^{+}(\hat{r}, s) g_2(r) \\ \mathcal{Y}_{jm}^{-}(\hat{r}, s) i f_2(r) \end{pmatrix}, \quad (11.46)$$

onde foram introduzidas amplitudes radiais $g_k(r)$ e $f_k(r)$, $k = 1, 2$. Os fatores i que acompanham as amplitudes f_k foram introduzidos para conveniência posterior. Note que as funções radiais

não dependem da variável de spin s , pois essa dependência destruiria a propriedade de que U_{jm} e V_{jm} sejam autofunções do quadrado do momento angular total J^2 . É claro que os dois tipos de função de onda de *quatro* componentes *não* têm um valor bem definido do momento angular orbital, pois ambos contém componentes tanto com $l = j + 1/2$ como componentes com $l = j - 1/2$. Na medida em que as componentes V_{jm} sejam pequenas comparadas com as componentes U_{jm} , as soluções de tipo 1 e 2 terão *predominantemente* $l = j - 1/2$ e $l = j + 1/2$ respectivamente.

11.5.2 Equações radiais acopladas.

O desempenho desses *ansätze* pode ser avaliado substituindo a solução de tipo 1 (por exemplo) na equação (11.45). Disso resultam as equações acopladas

$$\begin{pmatrix} m_e c^2 + v(r) - E & c\vec{\sigma} \cdot \vec{p} \\ c\vec{\sigma} \cdot \vec{p} & -m_e c^2 + v(r) - E \end{pmatrix} \begin{pmatrix} \mathcal{Y}_{jm}^-(\hat{r}, s) g_1(r) \\ \mathcal{Y}_{jm}^+(\hat{r}, s) i f_1(r) \end{pmatrix} = 0$$

ou, explicitamente, e usando unidades tais que $\hbar = c = m_e = 1$ para simplificar a notação,

$$\begin{cases} [1 + v(r) - E] \mathcal{Y}_{jm}^-(\hat{r}, s) g_1(r) + (\vec{\sigma} \cdot \vec{p}) \mathcal{Y}_{jm}^+(\hat{r}, s) i f_1(r) = 0 \\ (\vec{\sigma} \cdot \vec{p}) \mathcal{Y}_{jm}^-(\hat{r}, s) g_1(r) - [1 - v(r) + E] \mathcal{Y}_{jm}^+(\hat{r}, s) i f_1(r) = 0. \end{cases} \quad (11.47)$$

Essas equações envolvem, por um lado, operadores de multiplicação proporcionais à matriz identidade de dimensão 2, que aparecem entre colchetes, e, por outro lado, o operador diferencial $\vec{\sigma} \cdot \vec{p} = (1/i)\vec{\sigma} \cdot \vec{\nabla}$, que envolve as matrizes de Pauli $\vec{\sigma}$. Este último pode ser re-escrito fazendo uso da identidade

$$1_2 \equiv \frac{(\vec{\sigma} \cdot \vec{r})(\vec{\sigma} \cdot \vec{r})}{r^2}, \quad (11.48)$$

onde 1_2 é a matriz identidade de dimensão 2, como

$$\vec{\sigma} \cdot \vec{p} = \frac{(\vec{\sigma} \cdot \vec{r})(\vec{\sigma} \cdot \vec{r})}{r^2} (\vec{\sigma} \cdot \vec{p}) = \frac{(\vec{\sigma} \cdot \vec{r})}{r^2} [\vec{r} \cdot \vec{p} + i\vec{\sigma} \cdot (\vec{r} \times \vec{p})] = (\vec{\sigma} \cdot \hat{r}) \left(\frac{1}{i} \frac{\partial}{\partial r} + \frac{i}{r} \vec{\sigma} \cdot \vec{L} \right).$$

A derivada com relação à coordenada radial age apenas sobre as funções radiais $g_1(r)$ e $f_1(r)$, enquanto que $\vec{\sigma} \cdot \vec{L}$ age apenas sobre as funções angulares $\mathcal{Y}_{jm}^\pm(\hat{r}, s)$. Esta última ação pode ser explicitada lembrando que

$$\vec{\sigma} \cdot \vec{L} = 2\vec{S} \cdot \vec{L} = J^2 - L^2 - \frac{3}{4} \quad \text{com} \quad \vec{J} = \vec{L} + \vec{S}$$

o que dá imediatamente

$$(\vec{\sigma} \cdot \vec{L})\mathcal{Y}_{jm}^{\pm}(\hat{r}, s) = \left[\mp \left(j + \frac{1}{2} \right) - 1 \right] \mathcal{Y}_{jm}^{\pm}(\hat{r}, s).$$

Resta portanto estudar a ação sôbre as funções angulares do operador $\vec{\sigma} \cdot \hat{r}$. Este operador é um operador invariante sob rotações (isto é, um tensor de Racah de ordem zero) e portanto conserva o número quântico j , associado ao momento angular total. Por outro lado, como $\vec{\sigma}$ é uma das componentes do momento angular, a combinação $\vec{\sigma} \cdot \hat{r}$ é *ímpar* sob uma inversão espacial, de modo que sua ação sôbre as funções angulares deve necessariamente ser da forma

$$(\vec{\sigma} \cdot \hat{r})\mathcal{Y}_{jm}^{\pm}(\hat{r}, s) = \xi_{\pm} \mathcal{Y}_{jm}^{\mp}(\hat{r}, s).$$

Por outro lado, em vista da identidade (11.48),

$$\mathcal{Y}_{jm}^{\pm}(\hat{r}, s) = (\vec{\sigma} \cdot \hat{r})^2 \mathcal{Y}_{jm}^{\pm}(\hat{r}, s) = (\vec{\sigma} \cdot \hat{r}) \xi_{\pm} \mathcal{Y}_{jm}^{\mp}(\hat{r}, s) = \xi_{\mp} \xi_{\pm} \mathcal{Y}_{jm}^{\pm}(\hat{r}, s)$$

o que mostra que $\xi_{\pm} = 1/\xi_{\mp}$. O valor desses objetos pode na realidade ser calculado explicitamente através do elemento de matriz

$$\langle \mathcal{Y}_{jm}^{-}(\hat{r}, s) | \vec{\sigma} \cdot \hat{r} | \mathcal{Y}_{jm}^{+}(\hat{r}, s) \rangle = \xi_{+},$$

o que pode ser feito elementarmente, por exemplo, usando a relação (9.19) com $K = Q = 0$ para escrever o produto escalar $\vec{\sigma} \cdot \hat{r}$ em termos das componentes esféricas $\sigma_0 = \sigma_3$, $\sigma_{\pm} = \mp(\sigma_1 \pm i\sigma_2)/\sqrt{2}$ de $\vec{\sigma}$ e das componentes esféricas de \hat{r} , estas últimas expressas em termos em termos de harmônicas esféricas de ordem 1. O resultado desse cálculo é $\xi_{+} = \xi_{-} = -1$, de modo que as equações acopladas (11.47) para as soluções de tipo 1 ficam reduzidas a

$$\begin{cases} [1 + v(r) - E]g_1(r) - \left[\frac{\partial}{\partial r} + \frac{1 + (j + \frac{1}{2})}{r} \right] f_1(r) = 0 \\ -[1 - v(r) + E]f_1(r) + \left[\frac{\partial}{\partial r} + \frac{1 - (j + \frac{1}{2})}{r} \right] g_1(r) = 0. \end{cases}$$

Uma simplificação útil adicional desse sistema de equações pode ser obtida introduzindo as funções

$$F_1(r) \equiv r f_1(r), \quad G_1(r) \equiv r g_1(r)$$

em termos das quais, com uma re-ordenação dos termos e das equações, ele pode ser escrito como

$$\text{Tipo 1} \quad \left\{ \begin{array}{l} \left[\frac{\partial}{\partial r} + \frac{j + \frac{1}{2}}{r} \right] F_1(r) - [1 + v(r) - E] G_1(r) = 0 \\ \left[\frac{\partial}{\partial r} - \frac{j + \frac{1}{2}}{r} \right] G_1(r) - [1 - v(r) + E] F_1(r) = 0. \end{array} \right.$$

Um procedimento inteiramente análogo leva ao sistema de equações que devem ser satisfeitas pelas soluções de tipo 2

$$\text{Tipo 2} \quad \left\{ \begin{array}{l} \left[\frac{\partial}{\partial r} - \frac{j + \frac{1}{2}}{r} \right] F_2(r) - [1 + v(r) - E] G_2(r) = 0 \\ \left[\frac{\partial}{\partial r} + \frac{j + \frac{1}{2}}{r} \right] G_2(r) - [1 - v(r) + E] F_2(r) = 0 \end{array} \right.$$

onde foram usadas as definições

$$F_2(r) \equiv r f_2(r), \quad \text{e} \quad G_2(r) \equiv r g_2(r).$$

As funções $F_i(r)$ e $G_i(r)$, $i = 1, 2$, podem ser vistas como dependentes do parâmetro $\lambda \equiv j + 1/2$. Escrevendo-as então como $F_i^{(\lambda)}(r)$ e $G_i^{(\lambda)}(r)$, a forma dos dois sistemas de equações mostra claramente que

$$F_1^{(\lambda)} = F_2^{(-\lambda)} \quad \text{e} \quad G_1^{(\lambda)} = G_2^{(-\lambda)}.$$

Com essa observação, os dois sistemas de equações podem ser reduzidos ao sistema único

$$\text{Tipos 1 e 2} \quad \left\{ \begin{array}{l} \left[\frac{\partial}{\partial r} + \frac{\lambda}{r} \right] F_i^{(\lambda)}(r) - [1 + v(r) - E] G_i^{(\lambda)}(r) = 0 \\ \left[\frac{\partial}{\partial r} - \frac{\lambda}{r} \right] G_i^{(\lambda)}(r) - [1 - v(r) + E] F_i^{(\lambda)}(r) = 0 \end{array} \right. \quad (11.49)$$

com os valores do parâmetro λ dados por

$$\text{Tipo 1: } \lambda = j + \frac{1}{2} \quad \text{e} \quad \text{Tipo 2: } \lambda = -\left(j + \frac{1}{2}\right). \quad (11.50)$$

11.5.3 Soluções para o espectro discreto.

Os estados ligados deste modelo para o átomo hidrogenóide correspondem, como já dito, a um conjunto discreto de autovalores $E_n < m_e c^2$ (ou seja, $E_n < 1$ nas unidades em que $\hbar = c = m_e = 1$) e são descritos por suas respectivas autofunções, que devem ser funções de quadrado integrável, de quatro componentes. Isto significa, em particular, que elas devem ser soluções regulares na origem e tendendo a zero para $r \rightarrow \infty$ das equações acopladas radiais (11.49), para os diferentes possíveis valores do parâmetro λ , segundo (11.50). O potencial $v(r)$ será daqui em diante escrito explicitamente como

$$v(r) = -\frac{Ze^2}{r}.$$

As equações acopladas (11.49) podem na realidade ser reduzidas a equações diferenciais conhecidas através de transformações apropriadas. Substituindo as funções $F_i^{(\lambda)}$ e $G_i^{(\lambda)}$ por novas funções $K_i^{(\lambda)}$ e $I_i^{(\lambda)}$ definidas pelas relações (note que as soluções buscadas devem ter $E < 1$)

$$K_i^{(\lambda)} = \frac{G_i^{(\lambda)}}{\sqrt{1+E}} + \frac{F_i^{(\lambda)}}{\sqrt{1-E}}, \quad I_i^{(\lambda)} = \frac{G_i^{(\lambda)}}{\sqrt{1+E}} - \frac{F_i^{(\lambda)}}{\sqrt{1-E}}$$

e introduzindo a notação $\mathfrak{p} = \sqrt{1-E^2}$ elas se escrevem como

$$\begin{cases} \frac{dK_i^{(\lambda)}}{dr} + \left(\mathfrak{p} - \frac{Ze^2}{\mathfrak{p}} \frac{1}{r} \right) K_i^{(\lambda)} = \left(\lambda + \frac{Ze^2}{\mathfrak{p}} \right) \frac{1}{r} I_i^{(\lambda)} \\ \frac{dI_i^{(\lambda)}}{dr} - \left(\mathfrak{p} - \frac{Ze^2}{\mathfrak{p}} \frac{1}{r} \right) I_i^{(\lambda)} = \left(\lambda - \frac{Ze^2}{\mathfrak{p}} \right) \frac{1}{r} K_i^{(\lambda)}. \end{cases}$$

A eliminação de $I_i^{(\lambda)}$ entre essas duas equações leva à equação de segunda ordem para $K_i^{(\lambda)}$

$$r \frac{d^2 K_i^{(\lambda)}}{dr^2} + \frac{dK_i^{(\lambda)}}{dr} + \left(2Ze^2 E + \mathfrak{p} - \mathfrak{p}^2 r - \frac{\gamma^2}{r} \right) K_i^{(\lambda)} = 0$$

onde ainda foi definido o parâmetro $\gamma \equiv \sqrt{\lambda^2 - Z^2 e^4}$. Finalmente, exprimindo a variável r em termos de $x \equiv 2\mathfrak{p}r$ e efetuando a transformação adicional

$$K_i^{(\lambda)} = e^{-\frac{x}{2}} x^\gamma M_i^{(\lambda)}(x) \quad (11.51)$$

a equação resultante para $M_i^{(\lambda)}(x)$ é

$$x \frac{d^2 M_i^{(\lambda)}(x)}{dx^2} + (2\gamma + 1 - x) \frac{dM_i^{(\lambda)}(x)}{dx} - \left(\gamma - \frac{Ze^2 E}{p} \right) M_i^{(\lambda)}(x) = 0$$

cuja solução é a *função hipergeométrica confluyente*²

$$M_i^{(\lambda)}(x) = F \left(\gamma - \frac{Ze^2 E}{p} \middle| 2\gamma + 1 \middle| x \right).$$

As propriedades da função hipergeométrica confluyente que são de relevância mais imediata para o espectro discreto do átomo hidrogenóide são as que garantem a normalizabilidade das funções de onda (11.46). Para que isso ocorra, é preciso exigir condições suficientes de regularidade de $K_i^{(\lambda)}$, e portanto também de $M_i^{(\lambda)}(x)$, através da relação (11.51). Dessa relação segue, em primeiro lugar, que a regularidade de $K_i^{(\lambda)}$ na origem impõe a condição de que $\gamma = \sqrt{\lambda^2 - Z^2 e^4} > 0$. Com relação a essa condição, vale a pena observar que

$$\lambda^2 = \left(j + \frac{1}{2} \right)^2 \geq 1 \quad \text{enquanto} \quad Z^2 e^4 = \left(\frac{Z}{137} \right)^2$$

de modo que o radicando é sempre positivo para $Z < 137$. A exigência de positividade para γ significa que a raiz deve ser tomada com o sinal positivo, apenas. Em segundo lugar, o anulamento suficiente de $K_i^{(\lambda)}$ no limite $x \rightarrow \infty$ exige que o eventual crescimento com x da função hipergeométrica confluyente seja de fato dominado pela exponencial decrescente, o que de fato ocorre *apenas* nos casos em que a função hipergeométrica $F(a|b|x)$ se reduz a um polinômio. Para que isso aconteça é preciso que o parâmetro a seja um *inteiro não positivo*. Tendo em vista a expressão resultante para esse parâmetro em $M_i^{(\lambda)}(x)$, essa condição dá

$$\gamma - \frac{Ze^2 E}{p} = -n', \quad n' \geq 0, \text{ inteiro.}$$

Para cada um dos valores $n' > 0$ existem duas soluções normalizáveis das equações radiais, correspondentes respectivamente a $\lambda = \pm(j + 1/2)$, mas no caso particular em que $n' = 0$ apenas as soluções com $\lambda = +(j + 1/2)$ (isto é, as soluções de tipo 1) sobrevivem. Lembrando

²Uma notação algo curiosa mas comum para essa função é $F(a|b|x)$, onde a e b são os parâmetros da equação

$$x \frac{d^2 F}{dx^2} + (b - x) \frac{dF}{dx} - aF = 0.$$

V. e.g. P. Morse e H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill Book Co., N.Y. 1953, pag. 551.

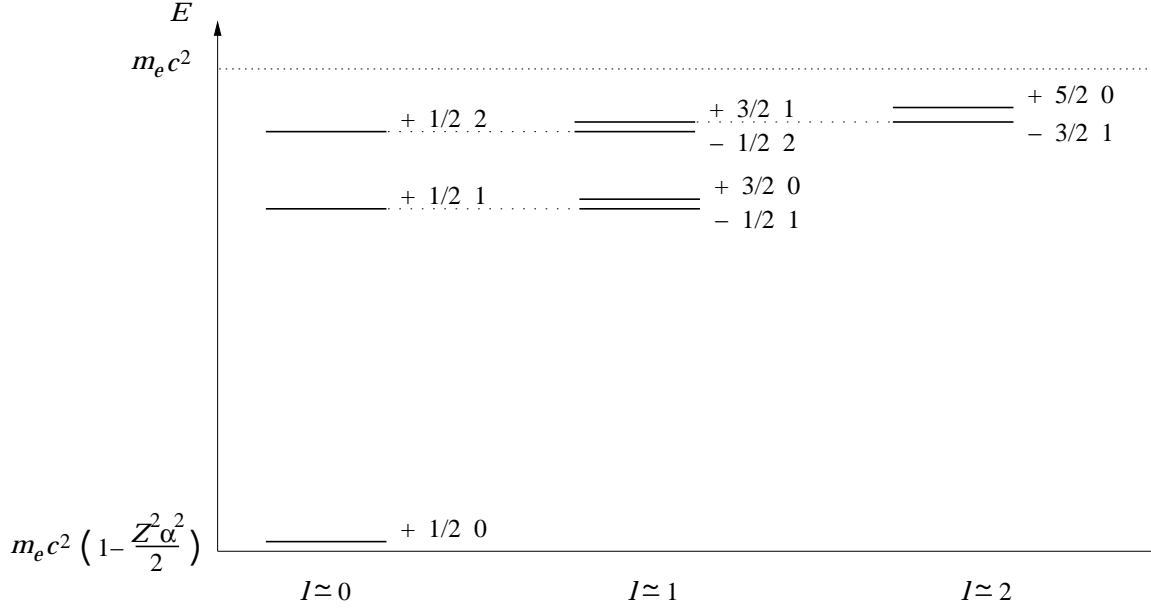


Figura 11.1: Esquema semi-quantitativo do estado fundamental e dos primeiros níveis excitados do átomo hidrogenóide de Dirac (cf. figs. 5.3 e 6.3). Cada um dos níveis mostrados é rotulado pelo sinal de $\lambda = \pm(j + 1/2)$ e pelos valores de j e n' (por exemplo, no estado fundamental o sinal de λ é $+$, $j = 1/2$ e $n' = 0$). As degenerescências são indicadas por linhas pontilhadas, e abaixo de cada grupo de estados está indicado o valor dominante do momento angular orbital l .

que $p = \sqrt{1 - E^2}$ é possível resolver esta última condição para E em termos de n' e dos demais parâmetros, do que resulta

$$E = \frac{1}{\sqrt{1 + \frac{Z^2 e^4}{(\gamma + n')^2}}} = \frac{m_e c^2}{\sqrt{1 + \left(\frac{Z\alpha}{\gamma + n'}\right)^2}} \quad (11.52)$$

onde, na última forma da expressão, foram re-introduzidos os fatores dimensionais apropriados, com

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137} \quad \text{e} \quad \gamma = \sqrt{\left(j + \frac{1}{2}\right)^2 + Z^2 \alpha^2}.$$

Os diferentes valores de E dados pela expressão (11.52) são os autovalores correspondentes ao espectro discreto do átomo hidrogenóide de Dirac. Uma primeira observação com relação a

esse espectro é a de que não existem estados ligados de energia negativa. O estado de menor energia é o estado de tipo 1 correspondente aos valores $n' = 0$, $j = 1/2$. Os demais estados de tipo 1 com $n' = 0$ (e valores de $j > 1/2$) correspondem aos estados de menor energia para cada um dos respectivos valores do momento angular j . Outros estados para cada um dos valores possíveis de j existem, com $n' \geq 1$, e aparecem em pares degenerados, contendo uma solução de tipo 1 e uma de tipo 2. Os primeiros níveis do espectro de energia resultante estão representados na figura 11.1. Sempre que $m_e c^2 - E \ll m_e c^2$ as duas componentes superiores das funções de onda de Dirac serão muito mais importantes que as duas componentes inferiores, como pode ser verificado a partir das equações acopladas (11.45) ou (11.47). Isso significa que, embora o momento angular orbital não seja um bom número quântico para as funções de onda de quatro componentes como um todo, os valores do momento angular orbital $l = j - 1/2$ e $l = j + 1/2$ serão dominantes respectivamente para as soluções de tipo 1 (que correspondem a $\lambda > 0$) e de tipo 2 (que correspondem a $\lambda < 0$). Esses valores dominantes de l também estão indicados na figura 11.1.

Finalmente, é fácil estabelecer contacto com o espectro obtido da teoria não relativística fazendo uma expansão do espectro relativístico (11.52) em potências de $Z^2 \alpha^2$. Levando a expansão até a ordem $Z^4 \alpha^4$ o resultado que se obtém pode ser escrito como

$$E = m_e c^2 \left[1 - \frac{Z^2 \alpha^2}{2n^2} - \frac{Z^4 \alpha^4}{2n^4} \left(\frac{n}{j + \frac{1}{2}} - \frac{3}{4} \right) + \dots \right]$$

onde o “número quântico principal” n aparece como uma abreviatura para $n = n' + j + 1/2$. O primeiro termo dessa expansão é simplesmente a energia de repouso do electron livre. O termo proporcional a α^2 reproduz o espectro de Bohr, enquanto o termo proporcional a α^4 contém a estrutura fina do espectro. Em particular, a separação obtida nessa ordem entre os níveis $\{+ 3/2 \ 0\}$ e $\{- 1/2 \ 1\}$ é

$$-m_e c^2 \frac{Z^4 \alpha^4}{2 \times 2^4} \left(\frac{2}{\frac{3}{2} + \frac{1}{2}} - \frac{2}{\frac{1}{2} + \frac{1}{2}} \right) = m_e c^2 \frac{Z^2 \alpha^4}{2 \times 16}$$

que, lembrando que a constante de Rydberg é dada por $R = m_e c^2 \alpha^2 / 2$, coincide com o resultado perturbativo obtido na página 308 (v. fig. 6.3), após a inclusão da correção relativística à energia cinética e do termo de Darwin.

No espectro hidrogenóide de Dirac, representado na figura 11.1, os estados $\{+ 1/2 \ 1\}$ e $\{- 1/2 \ 1\}$ são exatamente degenerados. Como já mencionado na conclusão do capítulo 5, esse resultado está em desacordo com o resultado experimental obtido em 1947 por Lamb e Retherford, que mostra o segundo desses dois estados como tendo uma energia algo *menor* que

o primeiro. O primeiro tratamento teórico dessa discrepância é devido a Bethe³, que obteve uma estimativa quantitativamente correta do efeito através de um tratamento *não relativístico* em que as propriedades do elétron são modificadas perturbativamente devido a processos virtuais de emissão e re-absorção de fótons. Tratamentos posteriores incluíram a possibilidade de conversão dos fótons virtuais em pares elétron-pósitron, tendo desempenhado um papel importante na reformulação da teoria quântica relativística em termos de uma teoria quântica de campos.

Problemas.

11.1 - Lembrando que qualquer matriz 2×2 pode ser escrita como combinação linear das tres matrizes de Pauli $\vec{\sigma}$ e da matriz identidade 2×2 , mostre que não há nenhuma matriz não nula 2×2 que *anti-comute* com *as tres* matrizes de Pauli.

11.2 - Demonstre a afirmação feita na página 602 com relação à equação (11.29), isto é, que o anulamento de A , juntamente com as propriedades das matrizes de Dirac, implica no anulamento dos coeficientes a , b_μ , $c_{\mu\nu}$, d_μ e e . Demonstre ainda, usando um procedimento análogo, a independência linear do conjunto de quatro matrizes formado pela matriz identidade de dimensão 2 e pelas três matrizes de Pauli.

11.3 - Sejam $e^{i\vec{k}\cdot\vec{r}-iEt/\hbar}u_n(\vec{k}, E)$, $n = 1, 2, 3, 4$ as quatro soluções da equação de Dirac livre. Usando a representação (11.18) para as matrizes de Dirac, calcule

$$u_n^\dagger(\vec{k}, E)\vec{\alpha}u_n(\vec{k}, E) \quad \text{e} \quad u_n^\dagger(\vec{k}, E)\beta u_n(\vec{k}, E)$$

para $n = 1, 2, 3, 4$. **Sugestão:** Calcule $(\not{p}^\dagger + M)\vec{\alpha}(\not{p} + M)$ e tome como resposta os elementos de matriz diagonais. Interprete o resultado em termos das equações (11.22) e (11.23).

11.4 - Escreva uma solução da equação de Dirac livre, que corresponde a uma onda plana se propagando na direção z , em termos de ondas esféricas.

11.5 - Mostre que uma transformação de Lorentz preserva o sinal da energia de uma solução da equação de Dirac.

11.6 - Resolva a equação de Dirac livre pelo seguinte método:

³H. Bethe, *Phys. Rev.* **72**, 339 (1947). Este trabalho está reproduzido na coletânea *Quantum Electrodynamics*, J. Schwinger (ed.), Dover, N. Y., 1958, pag. 139.

a) note que, para $\vec{p} = 0$, as soluções podem ser escolhidas como sendo

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

b) Obtenha soluções para um valor finito de \vec{p} fazendo a transformação de Lorentz apropriada.

Referências gerais.

Os textos gerais sobre mecânica quântica incluem, usualmente em sua parte final, capítulos sobre as equações relativísticas de Klein-Gordon e Dirac. Um tratamento que utiliza as mesmas convenções para a cinemática relativística que as adotadas aqui se encontra em

★ **A. Messiah**, *op. cit.*, Capítulo XX.

Uma introdução às equações relativísticas de Klein-Gordon e Dirac se encontra também nos capítulos iniciais de textos sobre a teoria quântica de campos. Um tal texto que, além disso, usa as convenções hoje “normais” para a cinemática relativística, é

★ **C. Itzykson e J.-B. Zuber**, *Quantum Field Theory*, McGraw-Hill, Inc., 1980, Capítulo 2, especialmente seções 2-1 a 2-4.

Há, além disso, textos “intermediários” da década de 1960, que tratam das equações relativísticas em considerável detalhe. Um texto desse tipo, que usa no entanto a chamada métrica pseudo-euclídeana para o espaço-tempo, é

★ **J. J. Sakurai**, *Advanced Quantum Mechanics*, Addison-Wesley Publishing Co., Inc., 1967 (Tenth Printing, 1987), Cap. 3.

Outro, que utiliza as convenções hoje tidas como normais, é

★ **J. D. Bjorken e S. D. Drell**, *Relativistic quantum mechanics*, McGraw-Hill, Inc., 1964.