Contents

Contents i

List of Figures ii

List of Tables iii

7 Time Reversal 1

7.1 The Poincaré Group . . . 1

7.1.1 Space inversion and time-reversal . . . 1

7.1.2 Representations of the Poincaré Lie algebra . . . 2

7.1.3 Whither time-reversal? . . . 3

7.2 Antilinearity : The Solution to All Our Problems . . . 4

7.2.1 Properties of antilinear operators . . . 4

7.2.2 Position and momentum eigenstates . . . 6

7.2.3 Change of basis for time-reversal . . . 7

7.2.4 Time reversal with spin . . . 8

7.2.5 Kramers degeneracy . . . 9

7.2.6 External fields . . . 9

7.3 Time Reversal and Point Group Symmetries . . . 10

7.3.1 Complex conjugate representations . . . 11

7.3.2 Generalized time-reversal . . . 12

7.4 Consequences of Time-Reversal . . . 13

7.4.1 Selection rules and time-reversal 13
7.4.2 Onsager reciprocity 14
7.5 Color groups 16
7.5.1 Magnetic Bravais lattices and magnetic space groups 19
7.5.2 Corepresentations of color groups 23
7.6 Appendix : The Foldy-Wouthuysen Transformation 23
7.6.1 Derivation of the Spin-Orbit Interaction 24

List of Figures

7.1 Colored squares and their magnetic point group symmetries 17
7.2 The two magnetic point groups deriving from $C_{2v} = mm2$ 18
7.3 The five new black and white Bravais lattices in two dimensions 19

List of Tables

7.1 Representations and degeneracies for generalized time reversal $\check{\Theta} = \hat{r}\,\check{T}$ 13
7.2 The 58 colored three-dimensional magnetic point groups 21
7.3 True Facts about magnetic lattices, point groups, and space groups 22

Chapter 7

Time Reversal

7.1 The Poincaré Group

7.1.1 Space inversion and time-reversal

Recall that the Poincaré group $\mathsf{P}(1,n)$ in n space dimensions is the set of matrices

$$R(L,\boldsymbol{b}) = \left[\begin{array}{ccc|c} L_{00} & \cdots & L_{0n} & b_0 \\ \vdots & \ddots & \vdots & \vdots \\ L_{n0} & \cdots & L_{nn} & b_n \\ \hline 0 & \cdots & 0 & 1 \end{array}\right] \quad , \tag{7.1}$$

where $L \in \mathsf{O}(1,n)$ is a Lorentz transformation, meaning $L^{\mathsf{T}} g L = g$ with $g = \mathsf{diag}(1,-1,\ldots,-1)$ is a diagonal matrix of rank $(n+1)$, and b is an $(n+1)$-component column vector[1]. Note g is of rank $n+2$, and its action on a vector ξ whose transpose is $\xi^{\mathsf{T}} = (x_0\,,\,x_1\,,\,\ldots\,,\,x_n\,,\,1)$ is given by

$$R(L,\boldsymbol{b})\,\xi = \left[\begin{array}{ccc|c} L_{00} & \cdots & L_{0n} & b_0 \\ \vdots & \ddots & \vdots & \vdots \\ L_{n0} & \cdots & L_{nn} & b_n \\ \hline 0 & \cdots & 0 & 1 \end{array}\right] \begin{pmatrix} x_0 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \begin{pmatrix} x'_0 \\ \vdots \\ x'_n \\ 1 \end{pmatrix} \equiv \xi' \quad , \tag{7.2}$$

where $x'_\mu = L_{\mu\nu} x_\nu + b_\mu$. The space inversion and time-reversal operators, I and T, respectively, expressed as elements of $\mathsf{O}(1,n)$, are then

$$I = \left[\begin{array}{c|c} 1 & 0_{1\times n} \\ \hline 0_{n\times 1} & -1_{n\times n} \end{array}\right] \qquad , \qquad T = \left[\begin{array}{c|c} -1 & 0_{1\times n} \\ \hline 0_{n\times 1} & 1_{n\times n} \end{array}\right] \quad , \tag{7.3}$$

[1] In chapter 1, we called this group $\mathsf{P}(n,1)$, which is equivalent to $\mathsf{P}(1,n)$. In both cases, the metric tensor $g_{\mu\nu}$ is diagonal and the temporal entry g_{00} is of opposite sign to the spatial entries $g_{11} = g_{22} = \cdots = g_{nn}$.

and are both of rank $(n+1)$. Their corresponding rank matrices $\mathcal{I}$ and $\mathcal{T}$, which are elements of $\mathsf{P}(1,n)$ and therefore of rank $(n+2)$, are then given by

$$\mathcal{I} = \left[\begin{array}{c|c} I_{(n+1)\times(n+1)} & 0_{(n+1)\times 1} \\ \hline 0_{1\times(n+1)} & 1 \end{array}\right] \quad , \quad \mathcal{T} = \left[\begin{array}{c|c} T_{(n+1)\times(n+1)} & 0_{(n+1)\times 1} \\ \hline 0_{1\times(n+1)} & 1 \end{array}\right] \quad . \tag{7.4}$$

Note that $\mathcal{I}^{-1} = \mathcal{I}$ and $\mathcal{T}^{-1} = \mathcal{T}$, and furthermore that

$$\begin{aligned} \mathcal{I}\, R(L,b)\, \mathcal{I}^{-1} &= R(ILI^{-1}, Ib) \\ \mathcal{T}\, R(L,b)\, \mathcal{T}^{-1} &= R(TLT^{-1}, Tb) \quad . \end{aligned} \tag{7.5}$$

The product $\mathcal{I}\mathcal{T} = \mathcal{T}\mathcal{I}$ is

$$\mathcal{I}\mathcal{T} = \left[\begin{array}{c|c} -1_{(n+1)\times(n+1)} & 0_{(n+1)\times 1} \\ \hline 0_{1\times(n+1)} & 1 \end{array}\right] \quad , \tag{7.6}$$

which commutes with all pure Lorentz transformations, but fails to commute with space-time translations, since

$$\mathcal{I}\mathcal{T}\, R(E,b)\, (\mathcal{I}\mathcal{T})^{-1} = R(E,-b) \quad . \tag{7.7}$$

7.1.2 Representations of the Poincaré Lie algebra

We now restrict our attention to the case $n = 3$, where the Poincaré group consists of 5×5 matrices. The generators of the Poincaré Lie algebra $\mathsf{p}(1,3)$ are classified as being translations, rotations, or boosts. The lowest order variation of $R(L,b)$ about the identity $R(E,0)$ is

$$\delta R = \left[\begin{array}{c|ccc|c} 0 & \delta\omega_{01} & \delta\omega_{02} & \delta\omega_{03} & \delta b_0 \\ \hline \delta\omega_{01} & 0 & \delta\omega_{12} & -\delta\omega_{31} & \delta b_1 \\ \delta\omega_{02} & -\delta\omega_{12} & 0 & \delta\omega_{23} & \delta b_2 \\ \delta\omega_{03} & \delta\omega_{31} & -\delta\omega_{23} & 0 & \delta b_3 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right] \equiv -iP^\mu\, \delta b_\mu - iJ^{\mu\nu}\, \delta\omega_{\mu\nu} \quad . \tag{7.8}$$

More precisely,

$$\begin{aligned} R(E, \delta b) &= \exp\bigl(-iP^\mu\, \delta b_\mu\bigr) \\ R(E+\delta\omega, 0) &= \exp\bigl(-iJ^{\mu\nu}\, \delta\omega_{\mu\nu}\bigr) \quad , \end{aligned} \tag{7.9}$$

where $A^\mu B_\mu = g_{\mu\nu} A^\mu B^\nu$, and where $\delta\omega^{\mathsf{T}} = -g\, \delta\omega\, g$ and $J^{\mathsf{T}} = -gJg$. We stress that for each $\mu \in \{0,1,2,3\}$ the generator P^μ is a 5×5 matrix, as is each of the six independent elements of $J^{\mu\nu}$. The latter is further split into its rotation and boost components by writing

$$J_i = \tfrac{1}{2}\epsilon_{ijk}\, J^{jk} \quad , \quad K_i = J_{i0} \quad , \tag{7.10}$$

with $i \in \{1,2,3\}$. $\boldsymbol{J}$ are the generators of rotations, and $\boldsymbol{K}$ the generators of boosts. Thus, we have a total of ten generators of the Lie algebra $\mathsf{p}(1,3)$: P^0, $\boldsymbol{P}$, $\boldsymbol{J}$, and $\boldsymbol{K}$, the 5×5 matrices of which can be read off from Eqn. 7.8. Under space inversion and time-reversal,

$$\mathcal{I}\,P^0\,\mathcal{I}^{-1} = +P_0 \quad , \quad \mathcal{I}\,\boldsymbol{P}\,\mathcal{I}^{-1} = -\boldsymbol{P} \quad , \quad \mathcal{I}\,\boldsymbol{J}\,\mathcal{I}^{-1} = +\boldsymbol{J} \quad , \quad \mathcal{I}\,\boldsymbol{K}\,\mathcal{I}^{-1} = -\boldsymbol{K} \tag{7.11}$$

$$\mathcal{T}\,P^0\,\mathcal{T}^{-1} = -P_0 \quad , \quad \mathcal{T}\,\boldsymbol{P}\,\mathcal{T}^{-1} = +\boldsymbol{P} \quad , \quad \mathcal{T}\,\boldsymbol{J}\,\mathcal{T}^{-1} = +\boldsymbol{J} \quad , \quad \mathcal{T}\,\boldsymbol{K}\,\mathcal{T}^{-1} = -\boldsymbol{K} \quad . \tag{7.12}$$

7.1.3 Whither time-reversal?

The problem is that P^0 changes sign under $\mathcal{T}$, but $\boldsymbol{P}$ does not. In classical mechanics, the action of time-reversal is

$$\boldsymbol{r}^T = \boldsymbol{r} \quad , \quad \boldsymbol{p}^T = -\boldsymbol{p} \quad , \quad \boldsymbol{L}^T = -\boldsymbol{L} \quad , \quad \boldsymbol{E}^T = \boldsymbol{E} \quad , \quad \boldsymbol{B}^T = -\boldsymbol{B} \quad . \tag{7.13}$$

Thus, if $H(\boldsymbol{E},\boldsymbol{B})$ is the Hamiltonian for a charged particle in the presence of electric and magnetic fields, $\big[H(\boldsymbol{E},\boldsymbol{B})\big]^T = H(\boldsymbol{E},-\boldsymbol{B})$. Unlike space inversion, time-reversal in classical mechanics is *not* a canonical transformation, since it does not preserve the Poisson bracket $\big\{x^\mu\,,\,p^\nu\big\}_{\text{PB}} = \delta^{\mu\nu}$. This is our first clue that there is something special about time-reversal and that attempting to implement it in quantum mechanics via a unitary transformation is doomed to fail.

Indeed, if we use Eqn. 7.12 to define Hermitian generators of $\mathsf{p}(1,3)$, we run into problems quantizing because the generator of time translations, P^0, which is the Hamiltonian, is apparently odd under time-reversal, while the momentum $\boldsymbol{P}$, which is the generator of space translations, is even under time-reversal. This poses severe problems for the classical-quantum correspondence.

Indeed, suppose we define a time-reversal operator $\hat{T}$ whose action on wavefunctions $\psi(\boldsymbol{x},t)$ is

$$\psi'(\boldsymbol{x},t) = \hat{T}\psi(\boldsymbol{x},t) = \lambda\,\psi(\boldsymbol{x},-t) \quad , \tag{7.14}$$

where $\lambda \in \mathbb{C}$. Does $\psi'(\boldsymbol{x},t)$ satisfy the Schrödinger equation?

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\,\psi'(\boldsymbol{x},t) = i\hbar\lambda\,\frac{\partial}{\partial t}\,\psi(\boldsymbol{x},-t) &= -i\hbar\lambda\frac{\partial}{\partial(-t)}\,\psi(\boldsymbol{x},-t) \\ &= -\lambda\,\hat{H}\,\psi(\boldsymbol{x},-t) = -\hat{H}\,\psi'(\boldsymbol{x},t) \quad . \end{aligned} \tag{7.15}$$

No, it does not. In hindsight, this was obvious from the start. The Schrödinger equation is first order in time, hence it is not invariant under $t \to -t$. So at this point we are left with three possibilities:

(i) Quantum physics, unlike Newtonian physics, is not invariant under time-reversal. *[horrible!]*

(ii) $\hat{T}\hat{H}\hat{T}^{-1} = -\hat{H}$ and the correspondence principle fails. *[horrible!]*

(iii) $i\hbar\frac{\partial}{\partial t}$ does not change sign under time-reversal. *[hmmm...]*

7.2 Antilinearity : The Solution to All Our Problems

I don't know about you, but I'm putting my money on option (iii). How could we make that work out? Well, suppose that the action of time-reversal is not given by Eqn. 7.14, but rather by

$$\psi'(\boldsymbol{x},t) = \lambda\check{K}\psi(\boldsymbol{x},-t) = \lambda\,\psi^*(\boldsymbol{x},-t) \quad , \tag{7.16}$$

where $\check{K}$ is the scalar *complex conjugation operator*[2] which complex conjugates every scalar to its right. Now we have

$$i\hbar\frac{\partial}{\partial t}\,\psi'(\boldsymbol{x},t) = i\hbar\lambda\,\frac{\partial}{\partial t}\,\psi^*(\boldsymbol{x},-t) = \lambda\left[i\hbar\frac{\partial}{\partial(-t)}\,\psi(\boldsymbol{x},-t)\right]^* = \lambda\,\hat{H}^*\psi^*(\boldsymbol{x},-t) = \hat{H}^*\psi'(\boldsymbol{x},t) \quad , \tag{7.17}$$

and so long as $\hat{H} = \hat{H}^*$, the Schrödinger equation remains invariant under time-reversal. Now that was so fun, let's do it again:

$$\psi''(\boldsymbol{x},t) = \lambda\check{K}\,\lambda\check{K}\,\psi\big(\boldsymbol{x},-(-t)\big) = |\lambda|^2\,\psi(\boldsymbol{x},t) \quad , \tag{7.18}$$

and if time-reversal applied twice preserves the state $\psi(\boldsymbol{x},t)$ up to a phase, we conclude that phase $|\lambda|^2$ must be unity, *i.e.* $\check{T}^2 = 1$ and $\lambda = e^{i\theta}$ is a unimodular complex number. This result is applicable to scalar wavefunctions ψ. When there is a spinor component due to intrinsic angular momentum, then the most general form for $\check{T}$ is $\check{T} = \hat{U}\check{K}$, where $\hat{U}$ is a unitary operator. In this case the unitary $\hat{U}$ may act on the spinor coordinates of Ψ, but since two applications of $\check{T}$ must result in the same state, *i.e.* must preserve the *ray* in Hilbert space, we conclude $\check{T}^2\,\Psi = e^{i\alpha}\,\Psi$, *i.e.* $\check{T}^2 = e^{i\alpha}$ is at most a constant phase. But then

$$\begin{aligned}\check{T}^3\,\Psi &= \check{T}\,(\check{T}^2\,\Psi) = \check{T}\,e^{i\alpha}\,\Psi = e^{-i\alpha}\,\check{T}\,\Psi \\ &= (\check{T}^2)\,\check{T}\,\Psi = e^{i\alpha}\,\check{T}\,\Psi \quad ,\end{aligned} \tag{7.19}$$

and we conclude $e^{2i\alpha} = 1$, hence $\alpha = 0$ or $\alpha = \pi$. As we shall see below, the case $\alpha = 0$ applies when the intrinsic angular momentum is $j \in \mathbb{Z}$, while $\alpha = \pi$ applies when $j \in \mathbb{Z} + \frac{1}{2}$.

An operator $\check{A}$ for which

$$\check{A}\Big[\alpha\,\big|\,\phi\,\big\rangle + \beta\,\big|\,\psi\,\big\rangle\Big] = \alpha^*\check{A}\,\big|\,\phi\,\big\rangle + \beta^*\check{A}\,\big|\,\psi\,\big\rangle \tag{7.20}$$

is called *antilinear*. Thus, time-reversal operator for spinless particles, $\check{T} = e^{i\theta}\check{K}$, is antilinear. Note that we use the inverted hat symbol (ˇ) to denote an antilinear operator.

7.2.1 Properties of antilinear operators

The following are True Facts about antilinear operators:

- An antilinear operator does not commute with complex numbers. Rather

$$\check{A}\,c = c^*\check{A} \qquad \Rightarrow \qquad \check{A}\,c\,\check{A}^{-1} = c^* \quad . \tag{7.21}$$

[2] We shall see below in §7.2.4 how to define time-reversal for particles with $S = \frac{1}{2}$.

- Rather than $\langle\,\phi\,|\,\hat{A}\,\psi\,\rangle = \langle\,\hat{A}^\dagger\,\phi\,|\,\psi\,\rangle$ as for linear $\hat{A}$, for antilinear $\check{A}$ we have

$$\langle\,\phi\,\big|\,\check{A}\,\psi\,\rangle = \langle\,\check{A}^\dagger\phi\,\big|\,\psi\,\rangle^* \quad . \tag{7.22}$$

  Indeed, the familiar Dirac notation $\langle\,\phi\,|\,\hat{A}\,|\,\psi\,\rangle = \langle\,\phi\,|\,\hat{A}\,\psi\,\rangle = \langle\,\hat{A}^\dagger\phi\,|\,\psi\,\rangle$ is misleading and should be eschewed in the case of antilinear operators, for which we may write

$$\Big(\langle\,\phi\,|\check{A}\Big) = \langle\,\check{A}^\dagger\,\phi\,| \quad , \quad \Big(\check{A}\,\big|\,\psi\,\rangle\Big) = |\,\check{A}\,\psi\,\rangle \tag{7.23}$$

  and

$$\Big(\langle\,\phi\,|\check{A}\Big)|\,\psi\,\rangle = \Big[\langle\,\phi\,|\Big(\check{A}\,\big|\,\psi\,\rangle\Big)\Big]^* \tag{7.24}$$

  Thus, $\langle\,\check{A}^\dagger\,\phi\,|\,\psi\,\rangle = \langle\,\phi|\check{A}\,\psi\,\rangle^*$. It is very dangerous and often wrong to remove the parentheses in the above relations!

- Though this follows from the first bullet, it is worth emphasizing:

$$(c\,\check{A})^{-1} = \check{A}^{-1}\,c^{-1} = c^{*-1}\,\check{A}^{-1} \quad . \tag{7.25}$$

- The time-reversal operator is both unitary and antilinear, *i.e.* it is *antiunitary*. Because it is unitary, $\check{T}^\dagger\check{T} = \hat{E}$. Thus entails

$$\big\langle\,\check{T}\phi\,\big|\,\check{T}\psi\,\big\rangle = \big\langle\,\check{T}^\dagger\check{T}\,\phi\,\big|\,\psi\,\big\rangle^* = \big\langle\,\phi\,\big|\,\psi\,\big\rangle^* = \big\langle\,\psi\,\big|\,\phi\,\big\rangle \quad . \tag{7.26}$$

  Thus,

$$\big|\big\langle\,\check{T}\phi\,\big|\,\check{T}\psi\,\big\rangle\big|^2 = \big|\big\langle\,\phi\,\big|\,\psi\,\big\rangle\big|^2 \tag{7.27}$$

  for all $|\,\phi\,\rangle$ and $|\,\psi\,\rangle$. So time-reversal preserves probabilities.

- Let $\hat{U}$ and $\hat{V}$ be unitary, and let $\check{A}$ and $\check{B}$ be antiunitary. Then $\hat{U}\hat{V}$ and $\check{A}\check{B}$ are both unitary, while $\hat{U}\check{B}$ and $\hat{V}\check{A}$ are both antiunitary. These follow directly from Eqn. 7.20.

- Any symmetry operation which preserves probabilities can be represented as an operator acting on the Hilbert space of states that is either both linear and unitary, or antilinear and antiunitary.

The proof of the last bullet point is elementary[3]. Let $\widetilde{Q}$ be an operator which preserves probabilities. Then it must preserve the norm, *i.e.*

$$\big\langle\,\widetilde{Q}\,\Psi\,\big|\,\widetilde{Q}\,\Psi\,\big\rangle = \big\langle\,\Psi\,\big|\,\Psi\,\big\rangle \quad . \tag{7.28}$$

for all $|\,\Psi\,\rangle$. Now let $|\,\Psi\,\rangle = c\,|\,\phi\,\rangle + |\,\psi\,\rangle$. Then

$$\begin{aligned}
\big\langle\,\widetilde{Q}\,(c\,\phi+\psi)\,\big|\,\widetilde{Q}\,(c\,\phi+\psi)\,\big\rangle &= \big\langle\,c\,\phi+\psi\,\big|\,c\phi+\psi\,\big\rangle \\
&= |c|^2\,\big\langle\,\phi\,\big|\,\phi\,\big\rangle + \big\langle\,\psi\,\big|\,\psi\,\big\rangle + 2\,\mathsf{Re}\Big[c\,\big\langle\,\psi\,\big|\,\phi\,\big\rangle\Big] \\
&= \big\langle\,\tilde{c}\,\widetilde{Q}\,\phi + \widetilde{Q}\,\psi\,\big|\,\tilde{c}\,\widetilde{Q}\,\phi + \widetilde{Q}\,\psi\,\big\rangle \\
&= |\tilde{c}|^2\,\big\langle\,\phi\,\big|\,\phi\,\big\rangle + \big\langle\,\psi\,\big|\,\psi\,\big\rangle + 2\,\mathsf{Re}\Big[\tilde{c}\,\big\langle\,\psi\,\big|\,\phi\,\big\rangle\Big] \quad ,
\end{aligned} \tag{7.29}$$

[3]Wigner said it, so it must be true. See, *e.g.*, Appendix A of S. Weinberg, *The Quantum Theory of Fields* (vol. 1) or Theorem 10.4.2 of Lax. Here we follow Lax's proof.

where $\tilde{c} = c$ if $\widetilde{Q}$ is linear and $\tilde{c} = c^*$ if $\widetilde{Q}$ is antilinear. Note $|\tilde{c}|^2 = |c|^2$ in either case. Then setting $c = 1$ and $c = i$ gives

$$\begin{aligned} c = 1 \qquad &\Rightarrow \qquad \mathsf{Re}\,\big\langle\, \widetilde{Q}\,\psi \,\big|\, \widetilde{Q}\,\phi\,\big\rangle = \mathsf{Re}\,\big\langle\, \psi \,\big|\, \phi\,\big\rangle \\ c = i \qquad &\Rightarrow \qquad \mathsf{Im}\,\big\langle\, \widetilde{Q}\,\psi \,\big|\, \widetilde{Q}\,\phi\,\big\rangle = \pm\,\mathsf{Im}\,\big\langle\, \psi \,\big|\, \phi\,\big\rangle \quad , \end{aligned} \tag{7.30}$$

where the top sign holds for $\widetilde{Q}$ linear and the bottom sign for $\widetilde{Q}$ antilinear. For $\widetilde{Q}$ linear, we have $\langle\, \widetilde{Q}\,\psi \,|\, \widetilde{Q}\,\phi\,\rangle = \langle\, \psi \,|\, \phi\,\rangle$, which establishes that $\widetilde{Q}$ is unitary. For $\widetilde{Q}$ antilinear, $\langle\, \widetilde{Q}\,\psi \,|\, \widetilde{Q}\,\phi\,\rangle = \langle\, \phi \,|\, \psi\,\rangle$, which establishes that $\widetilde{Q}$ is antiunitary.

As a result of the complex conjugation, we now have an updated and more suitable version of Eqn. 7.12,

$$\hat{I}\,\hat{P}^0\,\hat{I}^{-1} = +P_0 \quad , \quad \hat{I}\,\hat{\boldsymbol{P}}\,\hat{I}^{-1} = -\hat{\boldsymbol{P}} \quad , \quad \hat{I}\,\hat{\boldsymbol{J}}\,\hat{I}^{-1} = +\hat{\boldsymbol{J}} \quad , \quad \hat{I}\,\hat{\boldsymbol{K}}\,\hat{I}^{-1} = -\hat{\boldsymbol{K}} \tag{7.31}$$

$$\check{T}\,\hat{P}^0\,\check{T}^{-1} = +P_0 \quad , \quad \check{T}\,\hat{\boldsymbol{P}}\,\check{T}^{-1} = -\hat{\boldsymbol{P}} \quad , \quad \check{T}\,\hat{\boldsymbol{J}}\,\check{T}^{-1} = -\hat{\boldsymbol{J}} \quad , \quad \check{T}\,\hat{\boldsymbol{K}}\,\check{T}^{-1} = +\hat{\boldsymbol{K}} \quad . \tag{7.32}$$

All is well!

7.2.2 Position and momentum eigenstates

We may now compute the action of $\check{T}$ on operators, but how does it act on basis states? We first consider the case in which the particles are spinless. With respect to time-reversal, one can define an orthonormal basis $|\,\psi_\mu\,\rangle$ which is defined to be real, *i.e.* for which $|\,\check{T}\,\psi_\mu\,\rangle = |\,\psi_\mu\,\rangle$. To see why this is so, consider an arbitrary basis vector $|\,\phi_\mu\,\rangle$ and form $|\,\widetilde{\psi}_\mu\,\rangle \equiv |\,\phi_\mu\,\rangle + |\,\check{T}\,\phi_\mu\,\rangle$. For spinless particles, $\check{T}^2 = 1$, hence $|\,\check{T}\,\widetilde{\psi}_\mu\,\rangle = |\,\widetilde{\psi}_\mu\,\rangle$ and $|\,\widetilde{\psi}_\mu\,\rangle$ is an eigenstate of $\check{T}$ with eigenvalue $+1$. Now consider any vector $|\,\phi_\nu\,\rangle$ satisfying $\langle\,\widetilde{\psi}_\mu\,|\,\phi_\nu\,\rangle = 0$ and form $|\,\widetilde{\psi}_\nu\,\rangle \equiv |\,\phi_\nu\,\rangle + |\,\check{T}\,\phi_\nu\,\rangle$. Then

$$\big\langle\, \widetilde{\psi}_\mu \,\big|\, \widetilde{\psi}_\nu\,\big\rangle = \big\langle\, \widetilde{\psi}_\mu \,\big|\, \check{T}\,\phi_\nu\,\big\rangle = \big\langle\, \check{T}\,\widetilde{\psi}_\mu \,\big|\, \check{T}^2\,\phi_\nu\,\big\rangle^* = \big\langle\, \widetilde{\psi}_\mu \,\big|\, \phi_\nu\,\big\rangle^* = 0 \tag{7.33}$$

because $\check{T}^2 = +1$ and $\langle\,\check{T}\,\widetilde{\psi}_\mu\,| = \langle\,\widetilde{\psi}_\mu\,|$. Followed to its conclusion (for a finite-dimensional Hilbert space), this procedure results in a complete set of mutually orthogonal vectors, which can further be normalized so as to be orthonormal, *viz.* $\langle\,\psi_\mu\,|\,\psi_\nu\,\rangle = \delta_{\mu\nu}$ with $|\,\check{T}\,\psi_\mu\,\rangle = |\,\psi_\mu\,\rangle$ for all μ. Furthermore, if $\check{T}\hat{H}\check{T} = \hat{H}$, then its matrix elements in this basis are

$$H_{\mu\nu} = \big\langle\, \psi_\mu \,\big|\, \hat{H}\,\psi_\nu\,\big\rangle = \big\langle\, \check{T}\,\psi_\mu \,\big|\, \check{T}\,\overbrace{\check{T}^{-1}\hat{H}\,\check{T}}^{\hat{H}}\psi_\nu\,\big\rangle = \big\langle\, \psi_\mu \,\big|\, \hat{H}\,\big|\,\psi_\nu\,\big\rangle^* = H^*_{\mu\nu} \quad , \tag{7.34}$$

and therefore all the matrix elements $H_{\mu\nu}$ are real.

Typically this is taken to be the case for position eigenstates, $|\,\boldsymbol{r}\,\rangle$, *i.e.*

$$\big|\,\boldsymbol{r}\,\big\rangle = \big|\,\check{T}\boldsymbol{r}\,\big\rangle = \big|\,\check{T}^\dagger\boldsymbol{r}\,\big\rangle \quad , \tag{7.35}$$

and therefore

$$\begin{aligned} \big\langle\, \boldsymbol{r} \,\big|\, \check{T}\Big(\alpha\,\big|\,\phi\,\big\rangle + \beta\,\big|\,\psi\,\big\rangle\Big) &= \Big[\alpha\,\big\langle\,\check{T}^\dagger\boldsymbol{r}\,\big|\,\phi\,\big\rangle + \beta\,\big\langle\,\check{T}^\dagger\boldsymbol{r}\,\big|\,\psi\,\big\rangle\Big]^* \\ &= \alpha^*\big\langle\,\boldsymbol{r}\,\big|\,\phi\,\big\rangle^* + \beta^*\big\langle\,\boldsymbol{r}\,\big|\,\psi\,\big\rangle^* = \alpha^*\phi^*(\boldsymbol{r}) + \beta^*\psi^*(\boldsymbol{r}) \quad . \end{aligned} \tag{7.36}$$

Furthermore, if $|\,\psi\,\rangle = \int d^dr\,|\,\boldsymbol{r}\,\rangle\langle\,\boldsymbol{r}\,|\,\psi\,\rangle$, then

$$|\,\check{T}\psi\,\rangle = \check{T}\int d^dr\,\langle\,\boldsymbol{r}\,|\,\psi\,\rangle\,\check{T}^{-1}\,|\,\check{T}\boldsymbol{r}\,\rangle = \int d^dr\,|\,\boldsymbol{r}\,\rangle\,\langle\,\boldsymbol{r}\,|\,\psi\,\rangle^* \quad . \tag{7.37}$$

We also have $\check{T}\hat{H}\check{T}^{-1} = \hat{H}^*$, so if $\hat{H} = \hat{H}^*$, then $\hat{H}\,|\,\psi\,\rangle = E\,|\,\psi\,\rangle$ entails

$$\hat{H}\,\big|\,\check{T}\psi\,\big\rangle = \check{T}\hat{H}\check{T}^{-1}\,\big|\,\check{T}\psi\,\big\rangle = \check{T}\hat{H}\,\big|\,\psi\,\big\rangle = E\,\big|\,\hat{T}\psi\,\big\rangle \quad , \tag{7.38}$$

and since

$$\big\langle\,\boldsymbol{r}\,\big|\,\check{T}\psi\,\big\rangle = \big\langle\,\check{T}^\dagger\boldsymbol{r}\,\big|\,\psi\,\big\rangle^* = \big\langle\,\boldsymbol{r}\,\big|\,\psi\,\big\rangle^* = \psi^*(\boldsymbol{r}) \quad , \tag{7.39}$$

if the eigenstate $|\,\psi\,\rangle$ is nondegenerate, $\psi(\boldsymbol{r})$ can be chosen to be real. This is very important, so let's say it again, this time with feeling:

- ⋆ *For spinless particles, if $\check{T}\hat{H} = \hat{H}\check{T}$, the non-degenerate eigenstates of $\hat{H}$ are real, possibly multiplied by a constant phase.*

For momentum eigenstates, we have

$$e^{i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar} = \big\langle\,\boldsymbol{r}\,\big|\,\boldsymbol{p}\,\big\rangle = \big\langle\,\check{T}\boldsymbol{p}\,\big|\,\check{T}\boldsymbol{r}\,\big\rangle = \big\langle\,\boldsymbol{r}\,\big|\,\check{T}\boldsymbol{p}\,\big\rangle^* \quad , \tag{7.40}$$

and we conclude $|\,\check{T}\boldsymbol{p}\,\rangle = |\,\check{T}^\dagger\boldsymbol{p}\,\rangle = |\,-\boldsymbol{p}\,\rangle$. This can also be deduced from the operator transformation properties,

$$\check{T}\,\hat{\boldsymbol{r}}\,\check{T}^{-1} = +\hat{\boldsymbol{r}} \qquad , \qquad \check{T}\,\hat{\boldsymbol{p}}\,\check{T}^{-1} = -\hat{\boldsymbol{p}} \quad , \tag{7.41}$$

where the latter follows from the action of complex conjugation on $\hat{\boldsymbol{p}} = -i\hbar\boldsymbol{\nabla}$.

If $\hat{Q}$ is any operator with a definite signature under spinless time-reversal, *i.e.* if

$$\check{K}\,\hat{Q}\,\check{K}^{-1} = \eta_{\hat{Q}}\,\hat{Q} \tag{7.42}$$

with $\eta_{\hat{Q}} = \pm 1$, then if $\hat{H} = \hat{H}^*$ is time-reversal invariant,

$$\check{K}\,\hat{Q}(t)\,\check{K}^{-1} = \check{K}\,e^{i\hat{H}t/\hbar}\,\hat{Q}\,e^{-i\hat{H}t/\hbar}\,\check{K}^{-1} = e^{-i\hat{H}t/\hbar}\,\check{K}\,\hat{Q}\,\check{K}^{-1}\,e^{i\hat{H}t/\hbar} = \eta_{\hat{Q}}\,\hat{Q}(-t) \quad . \tag{7.43}$$

7.2.3 Change of basis for time-reversal

Recall $\check{T} = \lambda\check{K}$ with $|\lambda| = 1$. We are free to choose $\lambda = 1$, in which case $\check{T} = \check{K}$ is the complex conjugation operator. Thus far we have defined $\check{K}$ with respect to a particular $\check{T}$-invariant basis, *i.e.* the position basis[4] We could choose a different basis, $\big\{|\,n\,\rangle\big\}$, and define the action of a new time-reversal operator $\check{K}'$ as

$$\check{K}'\sum_n \psi_n\,\big|\,n\,\big\rangle = \sum_n \psi_n^*\,\big|\,n\,\big\rangle \quad . \tag{7.44}$$

[4]The "time-reversal basis" is one for which the basis states have time-reversal eigenvalue $+1$.

Since $\check{T} = \check{K}'$ and $\check{K}$ are both antilinear, they must be related by a unitary operator, *i.e.* $\check{T} = \hat{U}\check{K}$. Then

$$\check{T}^2 = \hat{U}\check{K}\hat{U}\check{K} = \hat{U}\hat{U}^* \equiv e^{i\alpha} \quad , \tag{7.45}$$

since a second time-reversal operation must restore the original state up to a phase. Thus, $e^{i\alpha}\,\hat{U}^\dagger = \hat{U}^*$, and taking the transpose we obtain $e^{i\alpha}\,\hat{U}^* = \hat{U}^\dagger$, whence $e^{2i\alpha} = 1$, which says that $\alpha = 0$ or $\alpha = \pi$. For spinless particles, as we have seen, $\check{T}^2 = +1$. In the next section, we consider the case of $S = \frac{1}{2}$.

7.2.4 Time reversal with spin

In order that the spin-orbit term in the electron Hamiltonian[5],

$$\hat{H}_{\rm SO} = \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V\times\boldsymbol{p} \quad , \tag{7.46}$$

remain invariant under time-reversal, we must have $\check{T}\boldsymbol{\sigma}\check{T}^{-1} = -\boldsymbol{\sigma}$. With $\check{T} = \hat{U}\check{K}$,

$$\begin{aligned}\check{T}\boldsymbol{\sigma}\check{T}^{-1} &= \hat{U}\check{K}\boldsymbol{\sigma}\check{K}^{-1}\hat{U}^{-1} \\ &= \hat{U}\big(\sigma^x\,\hat{\boldsymbol{x}} - \sigma^y\,\hat{\boldsymbol{y}} + \sigma^z\,\hat{\boldsymbol{z}}\big)\,\hat{U}^\dagger = -\boldsymbol{\sigma} \quad .\end{aligned} \tag{7.47}$$

Thus, $\hat{U}\sigma^x\hat{U}^\dagger = -\sigma^x$, $\hat{U}\sigma^y\hat{U}^\dagger = +\sigma^y$, $\hat{U}\sigma^z\hat{U}^\dagger = -\sigma^z$, and $\hat{U} = e^{i\beta}\sigma^y$ is a solution, where $e^{i\beta}$ is an arbitrary phase, which we may take to be i, so that $U = i\sigma^y \in \mathsf{SU}(2)$ has unit determinant. We now have

$$\check{T}^2 = (i\sigma^y)\,\check{K}\,(i\sigma^y)\,\check{K} = (i\sigma^y)^2 = -1 \quad . \tag{7.48}$$

For N spins each with $S = \frac{1}{2}$,

$$\check{T} = (i\sigma_1^y)\cdots(i\sigma_N^y)\,\check{K} \tag{7.49}$$

and $\check{T}^2 = (-1)^N$. Note that $\prod_{n=1}^N (i\sigma_n^y) = \exp(i\pi S^y/\hbar)$ corresponds to a rotation by π of all the spins about the y-axis in internal space.

Now consider the case of general angular momentum $\hat{\boldsymbol{J}}$, with $\hat{\boldsymbol{J}}^2 = \hbar^2 j(j+1)$. We have $\check{T}\hat{\boldsymbol{J}}\check{T}^{-1} = -\hat{\boldsymbol{J}}$, which entails $\check{T}\hat{\boldsymbol{J}}^2\check{T}^{-1} = \hat{\boldsymbol{J}}^2$, hence $\check{T}$ preserves the j quantum number. We also have

$$\hat{J}^z\,\check{T}\,\big|\,j,m\,\big\rangle = \check{T}\,(\check{T}^{-1}\hat{J}^z\,\check{T})\,\big|\,j,m\,\big\rangle = -\hbar m\,\check{T}\,\big|\,j,m\,\big\rangle \quad , \tag{7.50}$$

from which we conclude $\check{T}\,|\,j,m\,\rangle = C_{j,m}\,|\,j,-m\,\rangle$, where $C_{j,m}$ is a complex scalar. We furthermore have

$$\begin{aligned}\hat{J}^\pm\,\check{T}\,\big|\,j,m\,\big\rangle &= -\check{T}\,\hat{J}^\mp\,\big|\,j,m\,\big\rangle = -A_\mp(j,m)\,\check{T}\,\big|\,j,m\mp1\,\big\rangle = -A_\mp(j,m)\,C_{j,m\mp1}\,\big|\,j,-m\pm1\,\big\rangle \\ &= C_{j,m}\,\hat{J}^\pm\,\big|\,j,-m\,\big\rangle = C_{j,m}\,A_\pm(j,-m)\,\big|\,j,-m\pm1\,\big\rangle \quad ,\end{aligned} \tag{7.51}$$

where $A_\pm(j,m) = A_\mp(j,-m) = \hbar\sqrt{j(j+1)-m(m\mp1)}$. Thus, we conclude $C_{j,m\mp1} = -C_{j,m}$, which we may choose to solve with the assignment $C_{j,m} = (-1)^{j+m}$. One more time, with great feeling:

$$\check{T}\,\big|\,j,m\,\big\rangle = (-1)^{j+m}\,\big|\,j,-m\,\big\rangle \quad . \tag{7.52}$$

[5]Our notation is somewhat inconsistent as we generally do not place hats on $\boldsymbol{r}$, $\boldsymbol{p}$, and $\boldsymbol{\sigma}$. That these entities function as operators on Hilbert space is taken for granted.

Thus, $\check{T} = \hat{U}\check{K}$ with $\check{K}\,\big|\,j,m\,\big\rangle = \big|\,j,m\,\big\rangle$ and

$$\hat{U} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & -1 & 0 \\ \vdots & \iddots & & \vdots & \vdots \\ (-1)^{2j} & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{7.53}$$

so that

$$\check{T}^2 = \hat{U}\,\check{K}\,\hat{U}\,\check{K} = \hat{U}^2 = (-1)^{2j} \quad . \tag{7.54}$$

Therefore, in the case of a single j-quantum, we have $\check{T}^2 = +1$ or $j \in \mathbb{Z}$, and $\check{T}^2 = -1$ for $j \in \mathbb{Z} + \frac{1}{2}$. For the general case of N j-quanta, $\check{T}^2 = (-1)^{2jN}$.

7.2.5 Kramers degeneracy

When $\big[\check{T}, \hat{H}\big] = 0$ and $\check{T}^2 = -1$, all states in the eigenspectrum of $\hat{H}$ are evenfold degenerate. We prove this by showing that $|\,\psi\,\rangle$ and $|\,\check{T}\psi\,\rangle$ are degenerate and orthogonal. The proof of degeneracy is provided in Eqn. 7.38. As to orthogonality,

$$\big\langle\,\psi\,\big|\,\check{T}\psi\,\big\rangle = \big\langle\,\check{T}\psi\,\big|\,\check{T}^2\psi\,\big\rangle^* = -\big\langle\,\check{T}\psi\,\big|\,\psi\,\big\rangle^* = -\big\langle\,\psi\,\big|\,\check{T}\psi\,\big\rangle \quad , \tag{7.55}$$

and therefore $\langle\,\psi\,|\,\check{T}\psi\,\rangle = 0$. Thus, the dimension of Hilbert space must be even. For N spin-$\frac{1}{2}$ objects, this requires N odd, *i.e.* the total spin $S_{\rm tot}$ is a half odd integer.

In the absence of $\hat{H}_{\rm SO}$, we are free to define $\check{T} = \check{K}$ even though electrons have $S = \frac{1}{2}$. The reasoning is the same as that which permitted us to use the ordinary point group and not the double group in such circumstances. In the context of time-reversal, $\big[\check{K}, \hat{H}\big] = 0$ in the absence of $\hat{H}_{\rm SO}$, so we are free to classify states by their properties with respect to $\check{K}$ alone.

7.2.6 External fields

In the presence of external fields, the one electron Hamiltonian is given by

$$\begin{aligned} \hat{H} = \frac{1}{2m}\big(\boldsymbol{p} + \tfrac{e}{c}\boldsymbol{A}\big)^2 + V(\boldsymbol{r}) + \frac{e\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\boldsymbol{B} \\ + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V\times\big(\boldsymbol{p} + \tfrac{e}{c}\boldsymbol{A}\big) + \frac{\hbar^2}{8m^2c^2}\,\nabla^2 V + \mathcal{O}(m^{-3}) \quad , \end{aligned} \tag{7.56}$$

where $V(\boldsymbol{r}) = -e\,\phi(\boldsymbol{r})$, $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$, and $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$. Beyond the kinetic and potential energy terms in this expression, we have, respectively, the Zeeman and spin-orbit terms, both of which involve the electron's spin, and the Darwin term, which in the presence of a potential $V(\boldsymbol{r}) = -Ze^2/r$ is proportional to $\nabla^2(1/r) = -4\pi\,\delta(\boldsymbol{r})$. In general $\phi(\boldsymbol{r})$ is generated by an external charge density $\rho(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$ by an external current density $\boldsymbol{j}(\boldsymbol{r})$. Where do all these terms come from, by the way? From the Dirac equation, of course:

$$i\hbar\,\frac{\partial\Psi}{\partial t} = \begin{pmatrix} mc^2 + V & c\,\boldsymbol{\sigma}\cdot\big(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\big) \\ c\,\boldsymbol{\sigma}\cdot\big(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\big) & -mc^2 + V \end{pmatrix}\Psi \quad . \tag{7.57}$$

The wavefunction Ψ is a four-component Dirac spinor; each of the entries in the above Hamiltonian matrix is a 2×2 subblock. Since mc^2 is the largest energy scale in town, the coupling between the upper two "positive energy" components and the lower two "negative energy" components is relatively weak, and can be eliminated order by order in $(mc^2)^{-1}$ by successive canonical transformations of the type discussed in §6.7. This procedure is known as the *Foldy-Wouthuysen transformation* and is described in standard texts of yore such as Bjorken and Drell (see the Appendix §7.6 for a derivation). The Dirac equation is of course wrong[6], and the real theory of electrons interacting with photons is given by quantum electrodynamics. Thus the g-factor multiplying $(e/2mc)\,\boldsymbol{B}\cdot\boldsymbol{S}$, where $\boldsymbol{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$, is $g = 2$, which is the "tree level" value. Radiative corrections, calculable within QED, lead to $g = 2 + \frac{\alpha}{\pi} + \mathcal{O}(\alpha^2)$, where $\alpha = e^2/\hbar c \approx 1/137$ is the fine structure constant. But I digress.

The Hamiltonian $\hat{H}^T = \check{T}\hat{H}\check{T}^{-1}$ is invariant under time-reversal provided

$$\rho^T(\boldsymbol{r},t) = \rho(\boldsymbol{r},-t) \quad , \quad \boldsymbol{j}^T(\boldsymbol{r},t) = -\boldsymbol{j}(\boldsymbol{r},-t) \tag{7.58}$$

$$\phi^T(\boldsymbol{r},t) = \phi(\boldsymbol{r},-t) \quad , \quad \boldsymbol{A}^T(\boldsymbol{r},t) = -\boldsymbol{A}(\boldsymbol{r},-t) \tag{7.59}$$

$$\boldsymbol{E}^T(\boldsymbol{r},t) = \boldsymbol{E}(\boldsymbol{r},-t) \quad , \quad \boldsymbol{B}^T(\boldsymbol{r},t) = -\boldsymbol{B}(\boldsymbol{r},-t) \quad , \tag{7.60}$$

where the conditions on ϕ and $\boldsymbol{A}$ are of course true up to a gauge transformation. We then have $\hat{H}^T(\boldsymbol{E},\boldsymbol{B}) = \hat{H}(\boldsymbol{E}^T,\boldsymbol{B}^T)$ and $\psi^T(t\,;\,\boldsymbol{E},\boldsymbol{B}) = \hat{U}\check{K}\psi(-t\,;\,\boldsymbol{E}^T,\boldsymbol{B}^T)$.

7.3 Time Reversal and Point Group Symmetries

All point group operations $\hat{g} \equiv \hat{U}(g)$ commute with time-reversal: $\big[\check{T},\hat{g}\big] = 0$. The reason is that proper point group operations are rotations, hence $\hat{g} = \exp\big(-i\xi\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{J}}/\hbar\big)$, and $\check{T}\,i\hat{\boldsymbol{J}}\,\check{T}^{-1} = i\hat{\boldsymbol{J}}$. The improper operations include spatial inversion $\hat{I}$, which also commutes with $\check{T}$.

Consider the case of a particle of total spin j subjected to point group operations. We have

(i) $\check{T}^2 = +1$, $j = 0$: The time-reversal operator is then simply complex conjugation, *i.e.* $\check{T} = \check{K}$. Consider any complex scalar basis function $\psi(\boldsymbol{r})$. Then

$$\check{T}\,\hat{g}\,\psi(\boldsymbol{r}) = \hat{g}\,\check{T}\,\psi(\boldsymbol{r}) = \psi^*(\boldsymbol{r}g) \quad . \tag{7.61}$$

(ii) $\check{T}^2 = -1$, $j = \frac{1}{2}$: The time-reversal operator is $\hat{T} = i\sigma^y\check{K}$, and we write the wavefunction as a two-component column vector, with $\psi^{\mathsf{T}}(\boldsymbol{r}) = \big(\psi_\uparrow(\boldsymbol{r})\,,\,\psi_\downarrow(\boldsymbol{r})\big)$. Then it can be shown that

$$\check{T}\,\hat{g}\,\psi(\boldsymbol{r}) = \hat{g}\,\check{T}\,\psi(\boldsymbol{r}) = i\sigma^y\left[D^{1/2}(g)\right]^*\begin{pmatrix}\psi^*_\uparrow(\boldsymbol{r}g)\\ \psi^*_\downarrow(\boldsymbol{r}g)\end{pmatrix} \quad . \tag{7.62}$$

The essential step in establishing the above result is to show $\sigma^y D^{1/2}(g)\,\sigma^y = \big[D^{1/2}(g)\big]^*$, which is left as an exercise for the student.

[6] With respect to QED, the Dirac equation is correct "at tree level".

(iii) For general j, one has $\psi^{\mathsf{T}}(\boldsymbol{r}) = \big(\psi_m(\boldsymbol{r})\ ,\ \ldots\ ,\ \psi_{-m}(\boldsymbol{r})\big)$, and

$$\check{T}\,\hat{g}\,\psi(\boldsymbol{r}) = \hat{g}\,\check{T}\,\psi(\boldsymbol{r}) = \psi^{T}(\boldsymbol{r}g) \tag{7.63}$$

whose components are given by

$$\psi^{T}_{m}(\boldsymbol{r}g) = (-1)^{j-m}\left[D^{j}_{-m,m'}(g)\right]^{*}\psi^{*}_{m'}(\boldsymbol{r}g) \quad . \tag{7.64}$$

This is readily derived using the definition and properties of rotation matrices, discussed in §4.3.2.

7.3.1 Complex conjugate representations

Suppose $|\,\psi^{\Gamma}_{\mu}\,\rangle$ for $\mu \in \{1,\ldots,d_{\Gamma}\}$ are basis vectors for an invariant subspace $\mathcal{V}^{\Gamma}$ transforming according to a representation Γ of the point group G. Then combining point group and time-reversal operations yields

$$\hat{g}\,\big|\,\check{T}\,\psi^{\Gamma}_{\nu}\,\big\rangle = \check{T}\,\hat{g}\,\big|\,\psi^{\Gamma}_{\nu}\,\big\rangle = \check{T}\,\Big[\,\big|\,\psi^{\Gamma}_{\mu}\,\big\rangle\,D^{\Gamma}_{\mu\nu}(g)\Big] = \big|\,\check{T}\,\psi^{\Gamma}_{\mu}\,\big\rangle\,D^{\Gamma^{*}}_{\mu\nu}(g) \quad . \tag{7.65}$$

This tells us that the basis vectors $|\,\check{T}\psi^{\Gamma}_{\mu}\,\rangle$ transform as the complex conjugate representation Γ^{*}. In §2.5, we discussed the significance of the Frobenius-Schur indicator,

$$\varepsilon_{\Gamma} = \frac{1}{N_G}\sum_{g\in G}\chi^{\Gamma}(g^2) \quad , \tag{7.66}$$

which takes the values $\varepsilon_{\Gamma} \in \{-1,0,+1\}$, in determining whether a given IRREP can be made real, *i.e.* whether it is equivalent to one whose representation matrices are all real. We found

(i) $\varepsilon_{\Gamma} = +1$: The representation matrices $D^{\Gamma}(G)$ may be brought to real form by a similarity transformation $S\,D^{\Gamma}(G)\,S^{-1}$. All characters $\chi^{\Gamma}(g)$ are real. In such cases, Γ is said to be *real*.

(ii) $\varepsilon_{\Gamma} = -1$: $D^{\Gamma}(G)$ and $D^{\Gamma^{*}}$ are equivalent, meaning they are related by a similarity transformation, but they cannot be brought to real form. All characters $\chi^{\Gamma}(g)$ are real. In such cases, Γ is said to be *pseudoreal*.

(iii) $\varepsilon_{\Gamma} = 0$: $D^{\Gamma}(G)$ and $D^{\Gamma^{*}}$ are inequivalent, and $\chi^{\Gamma}(g) \notin \mathbb{R}$ for some group elements g. In such cases, Γ is said to be *complex*.

The single crystallographic point groups have no pseudoreal IRREPs. The following crystallographic groups have complex IRREPs: C_3 , C_{3h} , C_4 , C_{4h} , C_6 , C_{6h} , T, T_h. The spin IRREPs of the double point groups all have dimensions $d_{\Gamma} = 1$, 2, or 4. They are real only for the case $d_{\Gamma} = 1$. For $d_{\Gamma} = 2$ and 4, they are pseudoreal. For $\mathsf{SO}(3)$, we use the invariant measure to compute the Frobenius-Schur indicator. The invariant measure is given by

$$d\mu(\xi,\hat{\boldsymbol{n}}) = (1-\cos\xi)\,\frac{d\xi}{2\pi}\,\frac{d\hat{\boldsymbol{n}}}{4\pi} \quad , \tag{7.67}$$

hence

$$\begin{aligned}\varepsilon_j &= \frac{1}{2\pi}\int\limits_0^{2\pi} d\xi\,(1-\cos\xi)\,\chi^{(j)}(2\xi) \\ &= \frac{1}{2\pi}\int\limits_0^{2\pi} d\xi\,(1-\cos\xi)\,\frac{\sin(2j+1)\xi}{\sin\xi} = \begin{cases} +1 & \text{if } j\in\mathbb{Z} \\ -1 & \text{if } j\in\mathbb{Z}+\frac{1}{2} \end{cases} \quad .\end{aligned} \tag{7.68}$$

Does the presence of time-reversal symmetry lead to additional degeneracies in the eigenspectrum? We state the following results without proof:[7]

(i) If $\varepsilon_\Gamma\check{T}^2 = +1$, then there are no additional degeneracies. This is the case for real IRREPs when $\check{T}=+1$, and for pseudoreal IRREPs when $\check{T}^2=-1$. In the latter case, one can redefine the states so that $|\,\psi_\mu^{\Gamma^*}\,\rangle = |\,\check{T}\psi_\mu^\Gamma\,\rangle$.

(ii) If $\varepsilon_\Gamma\check{T}^2 = -1$, then there is a doubling, and $|\,\check{T}\psi_\mu^\Gamma\,\rangle$ is orthogonal to $|\,\psi_\nu^\Gamma\,\rangle$ for all μ and ν. This is the case for real IRREPs when $\check{T}^2=-1$ and for pseudoreal IRREPs when $\check{T}^2=+1$. Doubling means that a given representation appears twice, with degenerate energies.

(iii) If $\varepsilon_\Gamma\check{T}^2 = 0$, the IRREPs Γ and Γ^* are inequivalent and degenerate. Such degenerate IRREPs are called *paired*.

7.3.2 Generalized time-reversal

In cases where $\big[\check{T},\hat{H}\big]\neq 0$, but $\big[\hat{r}\,\check{T},\hat{H}\big]=0$ for some point group operation $r\in G$, the following result is useful:

$$\hat{g}\,|\,\check{\Theta}\,\psi_\nu^\Gamma\,\rangle = \Big(\hat{r}\,|\,\check{\Theta}\,\psi_\mu^\Gamma\,\rangle\Big)\,D^{\Gamma^*}_{\mu\nu}(h^{-1}gh) \quad , \tag{7.69}$$

where $\check{\Theta}=\hat{r}\,\check{T}$ is an antiunitary operator which effectively stands in for time-reversal $\check{T}$. This state of affairs persists, for example, when a magnetic field $\boldsymbol{H}$ is present, which by itself breaks time-reversal symmetry, but there is a point group operation $\hat{r}$, such as a twofold axis perpendicular to $\boldsymbol{H}$ or a mirror plane containing $\boldsymbol{H}$[8], both of which reverse $\boldsymbol{H}$. Thus $\check{\Theta}$ preserves $\boldsymbol{H}$ and is a symmetry, assuming time-reversal is otherwise unbroken. Above we considered the time-reverse representation $\Gamma^T=\Gamma^*$. We denote the generalized time-reverse representation by Γ^Θ. The details are worked out in §10.6 of Lax, and we present the results in Tab. 7.1. We define

$$\widetilde{\varepsilon}_\Gamma = \frac{1}{N_G}\sum_{g\in G}\chi^\Gamma\big((rg)^2\big) \tag{7.70}$$

as well as

$$D^{\Gamma^\Theta}(g) = \big[D^\Gamma(r^{-1}gr)\big]^* \quad , \qquad \chi^{\Gamma^\Theta}(g) = \mathsf{Tr}\,D^{\Gamma^\Theta}(g) = \big[\chi^\Gamma(r^{-1}gr)\big]^* \quad . \tag{7.71}$$

[7] For a proof, see Lax §10.7.

[8] Since $\boldsymbol{H}$ is a pseudovector, it is reversed by a mirror containing $\boldsymbol{H}$ and preserved by a mirror orthogonal to $\boldsymbol{H}$.

		equivalence of representations				degeneracies	
type	$\widetilde{\varepsilon}_\Gamma$	$\chi^{\Gamma^\Theta} = \chi^\Gamma$?	$D^{\Gamma^\Theta} = S^{-1} D^\Gamma S$?	S symmetry	$D^{\Gamma^\Theta} = D^\Gamma$?	$\check{T}^2 = +1$	$\check{T}^2 = -1$
1	$+1$	yes	yes	$S = D^\Gamma(r^2)\, S^{\mathsf{T}}$	if $S^{\mathsf{T}} = S$	none	doubling
2	-1	yes	yes	$S = -D^\Gamma(r^2)\, S^{\mathsf{T}}$	if $S^{\mathsf{T}} = S$	doubling	none
3	0	no	no	none	no	pairing	pairing

Table 7.1: Representations and degeneracies for generalized time reversal $\check{\Theta} = \hat{r}\,\check{T}$.

The issue is whether a given IRREP Γ is equivalent to its generalized time-reverse Γ^Θ, as well as whether the generalized time-reversal symmetry entails any extra degeneracies in the spectrum of $\hat{H}$. Equivalence of IRREPs means that there exists a fixed matrix S with $D^{\Gamma^\Theta}(g) = S^{-1} D^\Gamma(g)\, S$ for all $g \in G$. Since we presume our representations to be unitary, S is also unitary. The analysis hinges on the value of the generalized Frobenius-Schur indicator, $\widetilde{\varepsilon}_\Gamma$, defined above in eqn. 7.70. When $r = E$, the results in Tab. 7.1 recapitulate those already stated. For the cases $\widetilde{\varepsilon}_\Gamma = \pm 1$ (types 1 and 2), the representation matrices may be made to be identical, *i.e.* $D^{\Gamma^\Theta}(g) = D^\Gamma(g)$, provided $S = S^{\mathsf{T}}$, *i.e.* if $D^\Gamma(r^2) = \widetilde{\varepsilon}_\Gamma\, D^\Gamma(E)$.

7.4 Consequences of Time-Reversal

7.4.1 Selection rules and time-reversal

Selection rules in quantum mechanics refer to symmetries which result in the vanishing of certain matrix elements of the form $V_{\alpha\beta}^{\Gamma_a\Gamma_b} = \langle\, \Psi_\alpha^{\Gamma_a} \,|\, \hat{V} \,|\, \Psi_\beta^{\Gamma_b} \,\rangle$. From the point of view of the Wigner-Eckart theorem, we may always decompose the potential $\hat{V}$ into components which transform according to IRREPs of our symmetry group, *viz.*

$$\hat{V} = \sum_{\Gamma_c,\gamma} v_{\Gamma_c,\gamma}\, \hat{Q}_\gamma^{\Gamma_c} \quad , \tag{7.72}$$

in which the various matrix elements are subject to the considerations underlying the Wigner-Eckart theorem of §3.2.5. Here we are interested in the consequences of time-reversal symmetry.

Consider a matrix element

$$\big\langle\, \Psi \,\big|\, \hat{V} \,\big|\, \Psi' \,\big\rangle = \big\langle\, \hat{V}^\dagger\, \Psi \,\big|\, \Psi' \,\big\rangle = \big\langle\, \check{T}\Psi' \,\big|\, \check{T}\hat{V}^\dagger\check{T}^{-1}\, \check{T}\Psi \,\big\rangle = \eta_V \big\langle\, \check{T}\Psi' \,\big|\, \hat{V}\, \check{T}\Psi \,\big\rangle \quad , \tag{7.73}$$

where we assume $\check{T}\, \hat{V}^\dagger\, \check{T}^{-1} = \eta_V\, \hat{V}$. Typically $\hat{V}$ will be Hermitian, *i.e.* $\hat{V} = \hat{V}^\dagger$. Suppose further that $\Psi' = \check{T}\,\Psi$ is the time-reverse mate of Ψ. If $\check{T}^2 = \varepsilon_T$ then we have

$$\big\langle\, \Psi \,\big|\, \hat{V} \,\big|\, \Psi' \,\big\rangle = \eta_V\, \varepsilon_T \big\langle\, \Psi \,\big|\, \hat{V} \,\big|\, \Psi' \,\big\rangle \quad , \tag{7.74}$$

which means $\langle\, \Psi \,|\, \hat{V} \,|\, \Psi' \,\rangle = 0$ if $\eta_V\, \varepsilon_T = -1$. This is of course to be expected.

Now consider the thermal average of some Hermitian operator, $\langle \hat{V}(t) \rangle = \mathsf{Tr}\left(\hat{\varrho}\,\hat{V}(t)\right) = \sum_n P_n \langle\, n \,|\, \hat{V}(t) \,|\, n \,\rangle$, where $\hat{\varrho}$ is the equilibrium density matrix $Z^{-1}\exp(-\beta\hat{H})$ and $P_n = Z^{-1}\exp(-\beta E_n)$. For $\check{T}\hat{H}\check{T}^{-1} = \hat{H}$, we can equally well take the trace over the states $|\,\check{T}n\,\rangle$, which are each energetically degenerate with the corresponding $|\,n\,\rangle$. Then

$$\langle\, \hat{V}(t)\,\rangle = \sum_n P_n \left\langle\, \check{T}n \,\middle|\, \hat{V}(t)\,\check{T}n \,\right\rangle = \sum_n P_n\,\eta_V \left\langle\, n \,\middle|\, \hat{V}^\dagger(-t) \,\middle|\, n \,\right\rangle = \langle\, \hat{V}(-t)\,\rangle \quad , \tag{7.75}$$

where we have used

$$\left\langle\, \check{T}n \,\middle|\, \hat{V}(t)\,\check{T}n \,\right\rangle = \left\langle\, \check{T}n \,\middle|\, \check{T}\,\overbrace{\check{T}^{-1}\,\hat{V}(t)\,\check{T}}^{\eta_V \hat{V}(-t)}\, n \,\right\rangle = \eta_V \left\langle\, \hat{V}(-t)\,n \,\middle|\, n \,\right\rangle = \left\langle\, n \,\middle|\, \hat{V}^\dagger(-t) \,\middle|\, n \,\right\rangle \quad . \tag{7.76}$$

More generally, time-reversal symmetry has the following consequences,

$$\langle\, \hat{A}(t)\,\hat{B}(0)\,\rangle = \eta_A\,\eta_B\,\langle\, \hat{B}(0)\,\hat{A}(-t)\,\rangle \quad , \tag{7.77}$$

which leads us to the following discussion.

7.4.2 Onsager reciprocity

Now consider a general quantum mechanical system with a Hamiltonian $\hat{H}_0$ subjected to a time-dependent perturbation, $\hat{H}_1(t)$, where

$$\hat{H}_1(t) = -\sum_i \hat{Q}_i\,\phi_i(t)\;. \tag{7.78}$$

Here, the $\{\hat{Q}_i\}$ are a set of Hermitian operators, and the $\{\phi_i(t)\}$ are fields or potentials. Some examples:

$$\hat{H}_1(t) = \begin{cases} -\hat{\boldsymbol{M}}\cdot\boldsymbol{B}(t) & \text{magnetic moment – magnetic field} \\ \int d^3r\;\hat{\varrho}(\boldsymbol{r})\,\phi(\boldsymbol{r},t) & \text{density – scalar potential} \\ -\frac{1}{c}\int d^3r\;\hat{\boldsymbol{j}}(\boldsymbol{r})\cdot\boldsymbol{A}(\boldsymbol{r},t) & \text{electromagnetic current – vector potential} \end{cases}$$

We now ask, what is $\langle \hat{Q}_i(t)\rangle$? We assume that the lowest order response is linear, *i.e.*

$$\langle \hat{Q}_i(t)\rangle = \int\limits_{-\infty}^{\infty} dt'\;\chi_{ij}(t-t')\,\phi_j(t') + \mathcal{O}(\phi_k\,\phi_l)\;. \tag{7.79}$$

Note that we assume that the $\mathcal{O}(\phi^0)$ term vanishes, which can be assured with a judicious choice of the $\{Q_i\}$[9]. We also assume that the responses are all causal, *i.e.* $\chi_{ij}(t-t') = 0$ for $t < t'$. To compute $\chi_{ij}(t-t')$, we will use first order perturbation theory to obtain $\langle \hat{Q}_i(t)\rangle$ and then functionally differentiate with respect to $\phi_j(t')$:

$$\chi_{ij}(t-t') = \frac{\delta\left\langle \hat{Q}_i(t)\right\rangle}{\delta\phi_j(t')}\;. \tag{7.80}$$

[9] If not, define $\delta\hat{Q}_i \equiv \hat{Q}_i - \langle \hat{Q}_i\rangle_0$ and consider $\langle \delta\hat{Q}_i(t)\rangle$.

The first step is to establish the result,

$$\big|\,\Psi(t)\,\big\rangle = \mathcal{T}\,\exp\left\{-\frac{i}{\hbar}\int\limits_{t_0}^{t} dt'\,\Big[\hat{H}_0 + \hat{H}_1(t')\Big]\right\}\big|\,\Psi(t_0)\,\big\rangle\,, \tag{7.81}$$

where $\mathcal{T}$ is the *time ordering operator*, which places earlier times to the right. This is easily derived starting with the Schrödinger equation,

$$i\hbar\frac{d}{dt}\big|\,\Psi(t)\,\big\rangle = \hat{H}(t)\big|\,\Psi(t)\,\big\rangle\,, \tag{7.82}$$

where $\hat{H}(t) = \hat{H}_0 + \hat{H}_1(t)$. Integrating this equation from t to $t + dt$ gives

$$\begin{aligned}\big|\,\Psi(t+dt)\,\big\rangle &= \left(1 - \frac{i}{\hbar}\hat{H}(t)\,dt\right)\big|\,\Psi(t)\,\big\rangle \\ \big|\,\Psi(t_0 + N\,dt)\,\big\rangle &= \left(1 - \frac{i}{\hbar}\hat{H}\big(t_0 + (N-1)\,dt\big)\right)\cdots\left(1 - \frac{i}{\hbar}\hat{H}(t_0)\right)\big|\,\Psi(t_0)\,\big\rangle\,,\end{aligned} \tag{7.83}$$

hence

$$\big|\,\Psi(t_2)\,\big\rangle = U(t_2,t_1)\,\big|\,\Psi(t_1)\,\big\rangle \tag{7.84}$$

where

$$\hat{U}(t_2,t_1) = \mathcal{T}\exp\left\{-\frac{i}{\hbar}\int\limits_{t_1}^{t_2} dt\,\hat{H}(t)\right\} \tag{7.85}$$

is a unitary linear operator, known as the *time evolution operator* between times t_1 and t_2. It satisfies a composition law, $U(t_3,t_1) = U(t_3,t_2)\,U(t_2,t_1)$.

If $t_1 < t < t_2$, then differentiating $\hat{U}(t_2,t_1)$ with respect to $\phi_i(t)$ yields

$$\frac{\delta\hat{U}(t_2,t_1)}{\delta\phi_j(t)} = \frac{i}{\hbar}\,\hat{U}(t_2,t)\,\hat{Q}_j\,\hat{U}(t,t_1)\,, \tag{7.86}$$

since $\partial\hat{H}(t)/\partial\phi_j(t) = -\hat{Q}_j$. We may therefore write (assuming $t_0 < t, t'$)

$$\begin{aligned}\left.\frac{\delta\,|\,\Psi(t)\,\rangle}{\delta\phi_j(t')}\right|_{\{\phi_i=0\}} &= \frac{i}{\hbar}\,e^{-i\hat{H}_0(t-t')/\hbar}\,\hat{Q}_j\,e^{-i\hat{H}_0(t'-t_0)/\hbar}\,\big|\,\Psi(t_0)\,\big\rangle\,\Theta(t-t') \\ &= \frac{i}{\hbar}\,e^{-i\hat{H}_0 t/\hbar}\,\hat{Q}_j(t')\,e^{+i\hat{H}_0\,t_0/\hbar}\,\big|\,\Psi(t_0)\,\big\rangle\,\Theta(t-t')\,,\end{aligned} \tag{7.87}$$

where

$$\hat{Q}_j(t) \equiv e^{i\hat{H}_0 t/\hbar}\,\hat{Q}_j\,e^{-i\hat{H}_0 t/\hbar} \tag{7.88}$$

is the operator $\hat{Q}_j$ in the time-dependent *interaction representation*. Finally, we have

$$\begin{aligned}\chi_{ij}(t-t') &= \frac{\delta}{\delta\phi_j(t')}\big\langle\,\Psi(t)\,\big|\,\hat{Q}_i\,\big|\,\Psi(t)\,\big\rangle = \frac{\delta\,\langle\,\Psi(t)\,|}{\delta\phi_j(t')}\,\hat{Q}_i\,\big|\,\Psi(t)\,\big\rangle + \big\langle\,\Psi(t)\,\big|\,\hat{Q}_i\,\frac{\delta\,|\,\Psi(t)\,\rangle}{\delta\phi_j(t')} \\ &= \Big\{-\frac{i}{\hbar}\big\langle\,\Psi(t_0)\,\big|\,e^{-i\hat{H}_0 t_0/\hbar}\,\hat{Q}_j(t')\,e^{+i\hat{H}_0 t/\hbar}\,\hat{Q}_i\,\big|\,\Psi(t)\,\big\rangle \\ &\qquad\qquad + \frac{i}{\hbar}\big\langle\,\Psi(t)\,\big|\,\hat{Q}_i\,e^{-i\hat{H}_0 t/\hbar}\,\hat{Q}_j(t')\,e^{+i\hat{H}_0 t_0/\hbar}\,\big|\,\Psi(t_0)\,\big\rangle\Big\}\,\Theta(t-t') \\ &= \frac{i}{\hbar}\big\langle\big[\hat{Q}_i(t),\hat{Q}_j(t')\big]\big\rangle\,\Theta(t-t')\,,\end{aligned} \tag{7.89}$$

were averages are with respect to the wavefunction $|\,\Psi\,\rangle \equiv \exp(-i\hat{H}_0\,t_0/\hbar)\,|\,\Psi(t_0)\,\rangle$, with $t_0 \to -\infty$. This is sometimes known as the *retarded* response function. This result is valid at finite temperature if we take the bracket $\langle\cdots\rangle$ to denote thermal averaging, *viz.*

$$\begin{aligned}\chi_{ij}(t-t') &= \frac{i}{\hbar}\sum_m P_m \,\langle\, m \,|\, \big[\hat{Q}_i(t), \hat{Q}_j(t')\big] \,|\, m\,\rangle\,\Theta(t-t') \\ &= \frac{i}{\hbar}\sum_{m,n} P_m \Big\{ \langle\, m \,|\, \hat{Q}_i(t) \,|\, n\,\rangle\langle\, n \,|\, \hat{Q}_j(t') \,|\, m\,\rangle - \langle\, m \,|\, \hat{Q}_j(t') \,|\, n\,\rangle\langle\, n \,|\, \hat{Q}_i(t) \,|\, m\,\rangle \Big\}\,\Theta(t-t') \quad .\end{aligned} \tag{7.90}$$

Now the sums over states $|\,m\,\rangle$ and $|\,n\,\rangle$ can equally well be performed over their time-reverses $|\,\check{T}m\,\rangle$ and $|\,\check{T}n\,\rangle$. Assuming $\mathcal{H}_0$ is time-reversal invariant, the energy spectrum is identical. Note then that

$$\begin{aligned}\langle\, \check{T}m \,|\, \hat{Q}_i(t)\,\check{T}n\,\rangle\langle\, \check{T}n \,|\, \hat{Q}_j(t')\,\check{T}m\,\rangle &= \langle\, \check{T}^{-1}\hat{Q}_i(t)\check{T}\,n \,|\, m\,\rangle\langle\, \check{T}^{-1}\hat{Q}_j(t')\check{T}\,m \,|\, n\,\rangle \\ &= \eta_i\,\eta_j\,\langle\, \hat{Q}_i(-t)\,n \,|\, m\,\rangle\langle\, \hat{Q}_j(-t')\,m \,|\, n\,\rangle \\ &= \eta_i\,\eta_j\,\langle\, m \,|\, Q_j^\dagger(-t') \,|\, n\,\rangle\langle\, n \,|\, \hat{Q}_i^\dagger(-t) \,|\, m\,\rangle \quad ,\end{aligned} \tag{7.91}$$

where

$$\check{T}^{-1}\,\hat{Q}_i(t)\,\check{T} = \check{T}^{-1}\,e^{i\hat{H}_0 t/\hbar}\,\hat{Q}_i\,e^{-i\hat{H}_0 t/\hbar}\,\check{T} = e^{-i\hat{H}_0 t/\hbar}\,\check{T}^{-1}\,\hat{Q}_i\,\check{T}\,e^{+i\hat{H}_0 t/\hbar} = \eta_i\,\hat{Q}_i(-t) \quad . \tag{7.92}$$

Here, $\eta_i = \pm 1$ is the signature of the operator $\hat{Q}_i$ under time-reversal. Appealing now to hermiticity of $\hat{Q}_i$, we have

$$\chi_{ij}(t-t') = \frac{i}{\hbar}\,\eta_i\,\eta_j\,\langle\big[\hat{Q}_j(-t'), \hat{Q}_i(-t)\big]\rangle\,\Theta(t-t') = \eta_i\,\eta_j\,\chi_{ji}(t-t') \quad , \tag{7.93}$$

where we have also appealed to time translation invariance. The above relation is a consequence of time-reversal invariance, and is known as *Onsager reciprocity*. In the frequency domain, the linear response of a system to a finite frequency perturbations $\phi_j(\omega)$ is given by $\langle\hat{Q}_i(\omega)\rangle = \chi_{ij}(\omega)\,\phi_j(\omega)$.

We have assumed here that $\check{T}\hat{H}_0\check{T}^{-1} = \hat{H}_0$. Suppose though that $\hat{H}_0 = \hat{H}_0(\boldsymbol{H})$ where $\boldsymbol{H}$ is a magnetic field, which reverses under time-reversal. Then the derivation goes through as before, with the important *caveat* that $\hat{H}_0(\boldsymbol{H})$ must be replaced by $\hat{H}_0(-\boldsymbol{H})$ after time-reversal. The statement of Onsager reciprocity is then

$$\chi_{ij}(t-t') = \eta_i\,\eta_j\,\chi^T_{ji}(t-t') \quad , \tag{7.94}$$

where $\chi_{ij}(t-t')$ is the response function for a system with Hamiltonian $\hat{H}_0(\boldsymbol{H})$, and $\chi^T_{ji}(t-t')$ is the response function for a system with Hamiltonian $\hat{H}_0(-\boldsymbol{H})$.

7.5 Color groups

Cesium chloride is a cubic structure in which the Cs and Cl ions lie on interpenetrating simple cubic lattices. Its space group is $Pm\overline{3}m$, and consists of operations $\{\, g \,|\, \boldsymbol{R}\,\}$ where $g \in O_h$ and $\boldsymbol{R}$ is a simple cubic direct lattice vector. We could also describe the symmetries of CsCl by including another space

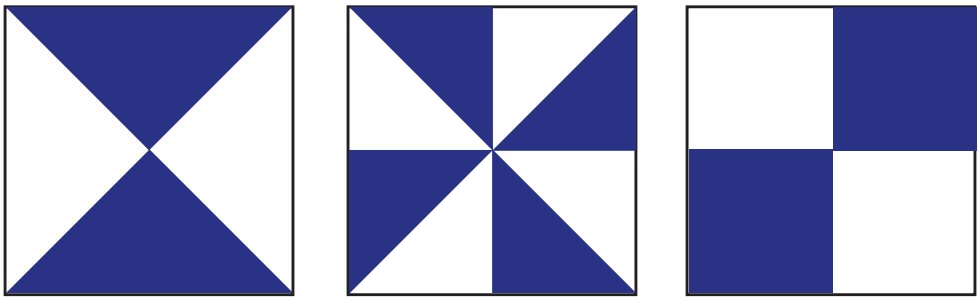

Figure 7.1: Colored squares and their magnetic point group symmetries.

group generator $\{\,\theta\,|\,\boldsymbol{\tau}\,\}$, where $\boldsymbol{\tau} = \frac{1}{2}(\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3)$ translates from a corner to a center of any cubic cell, and θ is an alchemy operation which converts Cs to Cl and vice versa. Thus $\theta^2 = E$.

Consider the symmetries of the three colored squares in Fig. 7.1. The symmetry of an uncolored square is C_{4v}, which has eight elements: $\{E, C_2, 2C_4, 2\sigma_v, 2\sigma_d\}$. In the figure, the square on the left is symmetric under the subgroup $\{E, C_2, 2\sigma_v\}$, but any of the operations $\{2C_4, 2\sigma_d\}$ is a symmetry only if they are accompanied by an operation θ which exchanges blue and white; again $\theta^2 = E$. Thus, it is symmetric under the operations of the *magnetic point group*[10]

$$4'm'm = \{E, C_2, 2\sigma_v, 2\theta C_4, 2\theta\sigma_d\} \quad . \tag{7.95}$$

The square in the center is symmetric under the subgroup $\{E, C_2, 2C_4\}$ and under $\{2\theta\sigma_v, 2\theta\sigma_d\}$, which altogether constitute the magnetic point group $4m'm'$:

$$4m'm' = \{E, C_2, 2C_4, 2\theta\sigma_v, 2\theta\sigma_d\} \quad . \tag{7.96}$$

Finally, the square on the right is symmetric under the subgroup $\{E, C_2, 2\sigma_d\}$ and under $\{2\theta C_4, 2\theta\sigma_v\}$, which altogether constitute the magnetic point group $4mm'$:

$$4'mm' = \{E, C_2, 2\sigma_v, 2\theta C_4, 2\theta\sigma_d\} \quad . \tag{7.97}$$

Since $4'mm'$ and $4m'm$ differ only by swapping the mirrors, they are equivalent, as can be seen by redefining their respective unit cells after a 45° rotation. We shall comment on the significance of the primes presently (astute students should be able to infer their meaning!).

Let $\mathcal{P}$ be an ordinary point group which we wish to extend to a magnetic point group $\mathcal{P}_{\rm M}$. It is easy to see that there are only the following three possibilities:

(i) No group operations involve color changes and $\mathcal{P}_{\rm M} = \mathcal{P}$. Such magnetic point groups are *uncolored*. Uncolored point groups describe nonmagnetic structures, or ferromagnets where all the local moments are of the same polarization ("color").

[10] See Mirman [1999], Joshua [1991], V. Kopský, *Symmetry* **7**, 135 (2015) and R. Lifshitz, *Magnetic Point Groups and Space Groups* in *Encyclopedia of Condensed Matter* **3**, 219 (Elsevier, 2005).

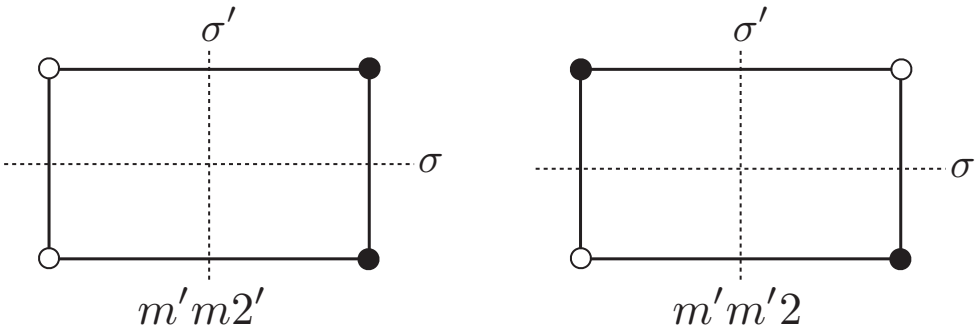

Figure 7.2: The two magnetic point groups deriving from $C_{2v} = mm2$. A black dot indicates spin polarization $m = +j$ and a white dot indicates a spin polarization $m = -j$, with $j \in \mathbb{Z}$.

(ii) To every element $g \in \mathcal{P}$ corresponds an element $\theta\, g \in \mathcal{P}_{\rm M}$. Thus $\mathcal{P}_{\rm M} = \mathcal{P} \times 1'$ where $1' = \{E, \theta\}$ is a $\mathbb{Z}_2$ clone. Such magnetic point groups are called *grey*. Note that $\theta \in \mathcal{P}_{\rm M}$ for all grey groups, which cannot be a symmetry element for any site group $\mathcal{P}_{\rm M}(\boldsymbol{r})$, because it changes the color on each site. However, coupled with lattice translations, the θ operation does appear in space group elements. For example, consider the one-dimensional antiferromagnet consisting of identical atoms whose local moments are arranged as $\big|\, \cdots \uparrow\downarrow\uparrow\downarrow\uparrow\downarrow \cdots \,\big\rangle$. Then if τ is half the wavelength of the spin pattern, *i.e.* the distance between consecutive $\uparrow$ and $\downarrow$ sites, $\big\{\, \theta \,\big|\, \tau \,\big\} \in \mathcal{S}_{\rm M}$ is an element of the magnetic space group and then by definition $\theta \in \mathcal{P}_{\rm M}$ is in the magnetic point group. The situation is roughly analogous to the status of the inversion operation I in diamond. The maximally symmetric site group for diamond is T_d, which does not contain I. But diamond is nonsymmorphic and $\big\{\, I \,\big|\, \boldsymbol{\tau}_I \,\big\} \in \mathcal{S}$ is in diamond's space group $Fd\overline{3}m$. Clearly this state of affairs requires translational symmetries and thus grey groups do not occur in finite systems such as molecules. Grey groups have twice the number of elements as their corresponding ordinary point groups.

(iii) Suppose $\mathcal{P}$ has a normal subgroup $\mathcal{B}$ of index two, which means that $\mathcal{P} = \mathcal{B} \cup (\mathcal{P} - \mathcal{B})$ and furthermore that $\mathcal{P} - \mathcal{B} = u\mathcal{B}$ for any $u \notin \mathcal{B}$. Thus the order $N_{\mathcal{P}}$ is even, and both $\mathcal{B}$ and $\mathcal{P} - \mathcal{B}$ contain $\frac{1}{2}N_{\mathcal{P}}$ elements. Now form the group $\mathcal{P}_{\rm M} = \mathcal{B} \cup \theta(\mathcal{P} - \mathcal{B})$, whose order is also $N_{\mathcal{P}}$ This is the familiar coset construction via Lagrange's theorem which we discussed in the dim and distant past (see §1.3.1). How do we find which point groups have normal subgroups of order two? Check the character tables for their one-dimensional representations other than the trivial IRREP. Such one-dimensional IRREPs will necessarily have $\chi^{\Gamma}(g) = -1$ for half the group elements, and the classes for which $\chi^{\Gamma}(\mathcal{C}) = +1$ contain the elements of an index two normal subgroup $\mathcal{B}$. Such groups $\mathcal{P}_{\rm M}$ are called *black and white groups*.

For example, C_{4v} has three nontrivial one-dimensional IRREPs: A_2, B_1, and B_2. In A_2, the classes with $\chi^{A_2}(\mathcal{C}) = 1$ are $\{E, C_2, 2C_4\}$, whose elements form an index two normal subgroup $\boldsymbol{B}$, whence the construction of $\mathcal{P}_{\rm M} = 4m'm'$ in Eqn. 7.96. Choosing the B_1 IRREP, we find $\mathcal{B} = \{E, C_2, 2\sigma_v\}$, whence the construction of $\mathcal{P}_{\rm M} = 4'm'm$ in Eqn. 7.95. Choosing B_2 just swaps the mirrors and yields a group equivalent to $4'm'm$. Another example is shown in Fig. 7.2, which depicts the two magnetic point groups deriving from $C_{2v} = mm2$.

The notation for magnetic space groups is to place a prime on elements of the Hermann-Maugin symbol which are paired with the θ operation in order to produce a symmetry. In other words, the primed

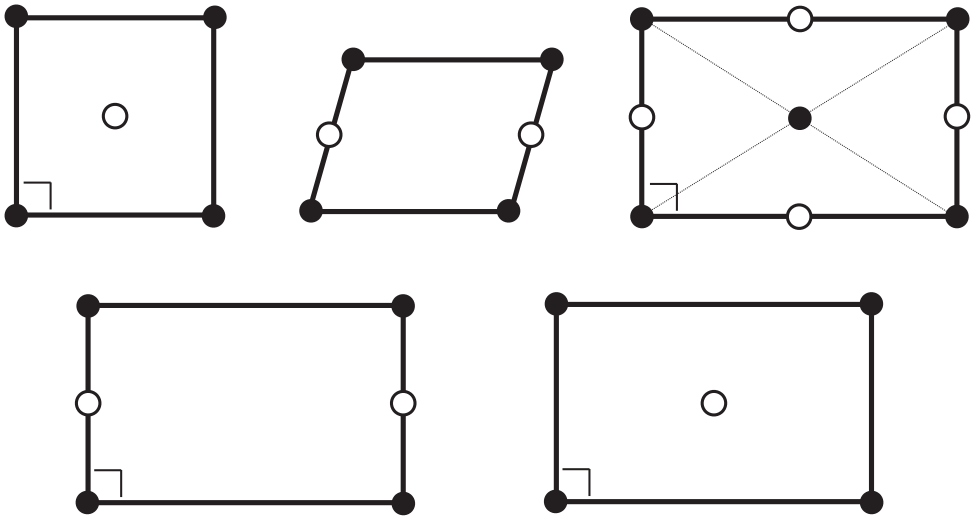

Figure 7.3: The five new black and white Bravais lattices in two dimensions. Clockwise from upper left, they derive from the square (1), oblique (1), centered rectangular (1), and rectangular (2) uncolored lattices. When the colors both fade to grey, each of these figures becomes a Bravais lattice (45° rotated square, oblique, rectangular, centered rectangular, and rectangular, respectively).

elements are those *not* in the subgroup $\mathcal{B}$.

7.5.1 Magnetic Bravais lattices and magnetic space groups

In constructing magnetic space groups, we must include translations. For ordinary space groups $\mathcal{S}$, the elements $\{ E \,|\, \boldsymbol{R} \}$ represent translations by a Bravais lattice vector $\boldsymbol{R}$. But for magnetic space groups $\mathcal{S}_{\rm M}$ we can have operations such as $\{ \theta \,|\, \boldsymbol{R} \}$, where $\boldsymbol{R}$ is a translation vector in the direct *magnetic Bravais lattice*. A magnetic Bravais lattice (or *colored lattice*) is one in which there are two sublattices, one black and one white. An important restriction is that a colored lattice must turn into a regular lattice if the colors fade, *i.e.* if the distinction between the black and white sites is removed. In two dimensions, there are five uncolored lattices (oblique, hexagonal, rectangular, centered rectangular, and square), and it turns out there are five black and white (BW) lattices, for a total of ten. In three dimensions, there are 14 uncolored lattices and 22 BW ones for a total of 36. Why are there almost double the number of BW as compared to uncolored lattices in $d = 3$? The reason is that the BW lattices have an additional primitive lattice vector $\boldsymbol{\tau}_{\rm BW}$, *i.e.* that which connects the B and W sublattices. There can be more than one possible such $\boldsymbol{\tau}_{\rm BW}$, however. The situation is roughly analogous to the different centering possibilities for cubic, tetragonal, orthorhombic, and monoclinic Bravais lattices. In Fig. 7.3 we show the additional five BW lattices which arise in two dimensions. Recall in $d = 2$ there are five lattices: oblique, hexagonal, rectangular, centered rectangular, and square. From the all-black oblique lattice, we can imagine adding white sites in the center of each plaquette or in the middle of one set of parallel edges. However these options are equivalent, as one can simply redefine the original direct lattice vectors such that one of them extends diagonally across the cell, *i.e.* replace $\boldsymbol{a}_2$ by $\boldsymbol{a}_2' = \boldsymbol{a}_1 + \boldsymbol{a}_2$. The center of the original oblique plaquette now lies at the midpoint of the side $\boldsymbol{a}_2'$. When faded, this BW lattice becomes an

oblique uncolored lattice. So we can only conjure up one additional BW lattice from the black oblique. The hexagonal case adds nothing new to what we have just derived, since the addition of a sublattice breaks the hexagonal symmetry anyway. With rectangular, we may place white sites in the center of each cell, or at the midpoint of a parallel pair of sides. When faded, the former yields a centered rectangular lattice and the latter another rectangular lattice. Attempting to place white sites at the midpoints of all sides of the rectangle does not produce a Bravais lattice when faded and is therefore not a permitted extension of the rectangular case, but it is a permitted extension of the centered rectangular case, and when faded produces a rectangular lattice. Finally, placing white sites at the centers of a pair of parallel sides of a square is equivalent to the same extension of the rectangular cell, hence yields nothing new. But placing a white site at the center of each square is permitted and leads to a θC_4 symmetry not present in any of the rectangular extensions. So there are a total of five additional BW lattices in two dimensoins. In three dimensions, to the 14 all black Bravais lattices, we get unique BW extensions from triclinic (1), monoclinic (5), orthorhombic (8), tetragonal (4), trigonal (0), hexagonal (2), and cubic (2), for a total of 22 BW extensions, and 36 colored Bravais lattices in all.

When it comes to point groups, we follow the recipe in §7.5 above. In any dimension, there is a grey (G) point group $\mathcal{P}' = \mathcal{P} \times \{E, \theta\}$ for each uncolored (U) point group $\mathcal{P}$. To create a black and white (BW) point group $\mathcal{P}_{\rm M}$, we must identify uncolored point groups $\mathcal{P}$ with normal subgroups $\mathcal{B}$ of index 2, and then construct $\mathcal{P}_{\rm M} = \mathcal{B} \cup \theta(\mathcal{P} - \mathcal{B})$. Since a given $\mathcal{P}$ may have several such maximal proper subgroups, the number of BW point groups is greater than then number of U or G point groups. In three dimensions, for example, there are 32 U and 32 G point groups, and 58 BW point groups. The latter are listed in Tab. 7.2. A summary of the numbers of U, G, and BW point groups in $d = 2$ and $d = 3$ dimensions is given in Tab. 7.3.

Next, we come to space groups. These may be build upon either uncolored or black and white Bravais lattices. For example, in $d = 3$ we have learned that there are a total of 230 uncolored space groups. For each space group $\mathcal{S}$ we can add the color changing element θ as a generator of the point group to create a grey space group $\mathcal{S}'$ with twice the number of elements of $\mathcal{S}$. There is a one-to-one correspondence between uncolored and grey groups, hence there are 230 grey space groups as well. When it comes to building the black and white space groups, if the underlying Bravais lattice is uncolored, the recipe is the same as for the point groups. That is, we start with a space group $\mathcal{S}$ generated by $\{ E \,|\, \boldsymbol{R} \}$ and $\{ g \,|\, \boldsymbol{\tau}_g \}$, where $g \in \mathcal{P} - \mathcal{B}$, and $\{ \theta\, h \,|\, \boldsymbol{\tau}_h \}$, where $h \in \mathcal{B}$. Thus results in 674 new black and white space groups. But we are not quite done! We could have started with a black and white Bravais lattice, in which case the black and white space group generators are $\{ E \,|\, \boldsymbol{R} \}$, $\{ g \,|\, \boldsymbol{\tau}_g \}$ where $g \in \mathcal{P} - \mathcal{B}$, and $\{ \theta\, h \,|\, \boldsymbol{\tau}_h + \boldsymbol{\tau}_{\rm BW} \}$ where $h \in \mathcal{B}$. This adds another 517 lattices, for a total of 1191 BW space groups. Adding this to the 230 uncolored and 230 grey space groups, we arrive at a total of 1,651 colored three-dimensional space groups, the properties of which are tabulated in a riveting 11,976 page text by D. B. Litvin, entitled *Magnetic Group Tables, Part 2* (International Union of Crystallography, 2013).

system	$\mathcal{P}^{(\mathrm{Sch})}$	$\mathcal{P}^{\mathrm{HM}}$	$N_{\mathcal{P}}$	colored point groups $\mathcal{P}_{\mathrm{M}}$
triclinic	C_1	1	1	none
	C_i	$\overline{1}$	2	$\overline{1}'$
monoclinic	C_2	2	2	$2'$
	C_s	m	2	m'
	C_{2h}	$2/m$	4	$2/m'$, $2'/m$, $2'/m'$
orthorhombic	D_2	222	4	$2'2'2$
	C_{2v}	$mm2$	4	$m'm2$, $m'm'2$
	D_{2h}	mmm	8	$m'mm$, $m'm'm$, $m'm'm'$
tetragonal	C_4	4	4	$4'$
	S_4	$\overline{4}$	4	$\overline{4}'$
	C_{4h}	$4/m$	8	$4/m'$, $4'/m$, $4'/m'$
	D_4	422	8	$4'22'$, $4'2'2'$
	C_{4v}	$4mm$	8	$4'm'm$, $4m'm'$
	D_{2d}	$\overline{4}2m$	8	$\overline{4}'2'm$, $\overline{4}'2m'$, $\overline{4}'2'm'$
	D_{4h}	$4/mmm$	16	$4/m'mm$, $4/mm'm'$, $4/m'm'm'$, $4'/mm'm$, $4'/m'm'm$
trigonal	C_3	3	3	none
	S_6	$\overline{3}$	3	$\overline{3}'$
	D_3	32	6	$32'$
	C_{3v}	$3m$	6	$3m'$
	D_{3d}	$\overline{3}m$	12	$\overline{3}'m$, $\overline{3}m'$, $\overline{3}'m'$
hexagonal	C_6	6	6	$6'$
	C_{3h}	$\overline{6}$	6	$\overline{6}'$
	C_{6h}	$6/m$	12	$6'/m$, $6/m'$, $6'/m'$
	D_6	622	12	$6'2'2$, $62'2'$
	C_{6v}	$6mm$	12	$6'm'm$, $6m'm'$
	D_{3h}	$\overline{6}m2$	12	$\overline{6}'m'2$, $\overline{6}'m2'$, $\overline{6}m'2'$
	D_{6h}	$6/mmm$	24	$6/m'mm$, $6/mm'm'$, $6/m'm'm'$, $6'/mm'm$, $6'/m'm'm$
cubic	T	23	12	none
	T_h	$m\overline{3}$	24	$m'\overline{3}$
	O	432	24	$4'32'$
	T_d	$\overline{4}3m$	24	$\overline{4}'3m'$
	O_h	$m\overline{3}m$	48	$m'\overline{3}m$, $m\overline{3}m'$, $m'\overline{3}m'$

Table 7.2: The 58 colored three-dimensional magnetic point groups.

	Bravais lattices			point groups				space groups			
	U	BW	total	U	G	BW	total	U	G	BW	total
$d = 2$	5	5	10	10	10	11	31	17	17	46	80
$d = 3$	14	22	36	32	32	58	122	230	230	1191	1651

Table 7.3: True Facts about magnetic lattices, point groups, and space groups. Notation: U (uncolored), G (grey), and BW (black and white).

7.5.2 Corepresentations of color groups

7.6 Appendix : The Foldy-Wouthuysen Transformation

The Dirac Hamiltonian is

$$H = mc^2\,\gamma^0 + c\gamma^0\,\boldsymbol{\gamma}\cdot\boldsymbol{\pi} + V \quad , \tag{7.98}$$

where $\boldsymbol{\pi} = \boldsymbol{p} + \frac{e}{c}\boldsymbol{A}$ is the dynamical momentum and where the γ^μ are the Dirac matrices,

$$\gamma^0 = \begin{pmatrix} 1_{2\times2} & 0_{2\times2} \\ 0_{2\times2} & -1_{2\times2} \end{pmatrix} \qquad , \qquad \boldsymbol{\gamma} = \begin{pmatrix} 0_{2\times2} & \boldsymbol{\sigma}_{2\times2} \\ -\boldsymbol{\sigma}_{2\times2} & 0_{2\times2} \end{pmatrix} \quad . \tag{7.99}$$

Here $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The Dirac equation is

$$i\hbar\frac{d\Psi}{dt} = H\Psi \quad , \tag{7.100}$$

where Ψ is a four-component Dirac spinor.

The idea behind the FW transformation is to unitarily transform to a different Hilbert space basis such that the coupling in H between the upper and lower components of the Dirac spinor vanishes. This may be done systematically as an expansion in inverse powers of the electron mass m. We begin by defining $K \equiv c\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\pi} + V$ so that $H = mc^2\,\gamma^0 + K$. Note that K is of order m^0. We then write

$$\widetilde{H} = e^{iS}\,H\,e^{-iS} = H + i\big[S, H\big] + \frac{(i)^2}{2!}\,\big[S,[S,H]\big] + \ldots \quad , \tag{7.101}$$

where S itself is written as a power series in $(mc^2)^{-1}$:

$$S = \frac{S_0}{mc^2} + \frac{S_1}{(mc^2)^2} + \ldots \ . \tag{7.102}$$

The job now is to write $\widetilde{H}$ as a power series in m^{-1}. The first few terms are easy to find:

$$\widetilde{H} = mc^2\,\gamma^0 + K + i\big[S_0,\gamma^0\big] + \frac{1}{mc^2}\bigg(i\big[S_0,K\big] + i\big[S_1,\gamma^0\big] - \tfrac{1}{2}\big[S_0,[S_0,\gamma^0]\big]\bigg) + \ldots \tag{7.103}$$

We choose the operators S_n so as to cancel, at each order in m^{-1}, the off-diagonal terms in $\widetilde{H}$ that couple the upper two components of Ψ to the lower two components of Ψ. To order m^0, we then demand

$$c\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\pi} + i\big[S_0,\gamma^0\big] = 0\,. \tag{7.104}$$

Note that we do not demand that $i\big[S_0,\gamma^0\big]$ completely cancel K – indeed it is impossible to find such an S_0, and one way to see this is to take the trace. The trace of any commutator must vanish, but $\mathrm{Tr}\,K = 4V$, which is in general nonzero. But this is of no concern to us, since we only need cancel the (traceless) *off-diagonal* part of K, which is to say $c\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\pi}$.

To solve for S_0, one can write it in terms of its four 2×2 subblocks, compute the commutator with γ^0, and then impose eqn. 7.104. One then finds $S_0 = -\frac{i}{2} c \boldsymbol{\gamma} \cdot \boldsymbol{\pi}$, the derivation of which is left as an exercise.

At the next level, we have to deal with the term in the round brackets in eqn. 7.103. Since we know S_0, we can compute the first and the third terms therein. In general, this will leave us with an off-diagonal term coupling upper and lower components of Ψ. We then choose S_1 so as to cancel this term. This calculation already is tedious, and we haven't even gotten to the spin-orbit interaction term yet, since it is of order m^{-2}.

7.6.1 Derivation of the Spin-Orbit Interaction

Here's a simpler way to proceed to order m^{-2}. Let a, b be block indices and i, j be indices within each block. Thus, the component Ψ_{ai} is the $i^{\rm th}$ component of the $a^{\rm th}$ block; $\Psi_{a=1,i=2}$ is the lower component of the upper block, *i.e.* the second component of the four-vector Ψ.

Write the Hamiltonian as

$$H = mc^2 \, \tau^z + c \boldsymbol{\sigma} \cdot \boldsymbol{\pi} \, \tau^x + V(\boldsymbol{r}) \quad , \tag{7.105}$$

where τ^μ are Pauli matrices with indices a, b and σ^ν are Pauli matrices with indices i, j. The σ and τ matrices commute because they act on different indices.

A very important result regarding Pauli matrices:

$$e^{i\theta \, \hat{\boldsymbol{n}} \cdot \boldsymbol{\tau}/2} \, \tau^\alpha \, e^{-i\theta \, \hat{\boldsymbol{n}} \cdot \boldsymbol{\tau}/2} = n^\alpha n^\beta \, \tau^\beta + \cos\theta \, (\delta^{\alpha\beta} - n^\alpha n^\beta) \, \tau^\beta + \sin\theta \, \epsilon^{\alpha\beta\gamma} \, n^\beta \, \tau^\gamma \quad . \tag{7.106}$$

STUDENT EXERCISE: Verify and interpret the above result.

Using this result, we can write

$$A \, \tau^z + B \, \tau^x = \sqrt{A^2 + B^2} \cdot e^{-i \tan^{-1}(B/A) \, \tau^y/2} \, \tau^z \, e^{i \tan^{-1}(B/A) \, \tau^y/2} \, , \tag{7.107}$$

and, for our specific purposes,

$$mc^2 \, \tau^z + \boldsymbol{\sigma} \cdot \boldsymbol{\pi} \, \tau^x = \sqrt{(mc^2)^2 + (c \boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2} \cdot U \, \tau^z \, U^\dagger \quad , \tag{7.108}$$

where $U = \exp\Big(- i \tan^{-1}(\frac{\boldsymbol{\sigma} \cdot \boldsymbol{\pi}}{mc}) \, \tau^y/2 \Big)$. The fact that $\boldsymbol{\sigma} \cdot \boldsymbol{\pi}$ is an operator is no obstacle here, since it commutes with the τ^μ matrices. We can give meaning to expressions like $\tan^{-1}(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}/mc)$ in terms of their Taylor series expansions.

We therefore have the result,

$$U^\dagger \, H \, U = \sqrt{(mc^2)^2 + (c \boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2} \, \tau^z + U^\dagger \, V(\boldsymbol{r}) \, U \quad . \tag{7.109}$$

The first term is diagonal in the block indices. Expanding the square root, we have

$$\begin{aligned} mc^2 \sqrt{1 + \Big(\frac{\boldsymbol{\sigma} \cdot \boldsymbol{\pi}}{mc}\Big)^2} &= mc^2 + \frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2}{2m} + \mathcal{O}(m^{-3}) \\ &= mc^2 + \frac{\boldsymbol{\pi}^2}{2m} + \frac{e\hbar}{2mc} \, \boldsymbol{B} \cdot \boldsymbol{\sigma} + \mathcal{O}(m^{-3}) \quad , \end{aligned} \tag{7.110}$$

since

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 &= \sigma^\mu\sigma^\nu\,\pi^\mu\pi^\nu = (\delta^{\mu\nu} + i\epsilon^{\mu\nu\lambda}\sigma^\lambda)\,\pi^\mu\pi^\nu \\ &= \boldsymbol{\pi}^2 + \tfrac{i}{2}\epsilon^{\mu\nu\lambda}\left[p^\mu + \tfrac{e}{c}A^\mu, p^\nu + \tfrac{e}{c}A^\nu\right] \\ &= \boldsymbol{\pi}^2 + \frac{e\hbar}{c}\,\boldsymbol{B}\cdot\boldsymbol{\sigma} \quad .\end{aligned} \tag{7.111}$$

We next need to compute $U^\dagger\,V(\boldsymbol{r})\,U$ to order m^{-2}. To do this, first note that

$$U = 1 - \frac{i}{2}\,\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{mc}\,\tau^y - \frac{1}{8}\Big(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{mc}\Big)^2 + \ldots \quad , \tag{7.112}$$

Thus,

$$U^\dagger\,V\,U = V + \frac{i}{2mc}\big[\boldsymbol{\sigma}\cdot\boldsymbol{\pi}, V\big]\,\tau^y - \frac{1}{8m^2c^2}\big[\boldsymbol{\sigma}\cdot\boldsymbol{\pi}, [\boldsymbol{\sigma}\cdot\boldsymbol{\pi}, V]\big] + \ldots \quad . \tag{7.113}$$

Upon reflection, one realizes that, to this order, it suffices to take the first term in the Taylor expansion of $\tan^{-1}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}/mc)$ in the expression for U, in which case one can then invoke eqn. 7.101 to obtain the above result. The second term on the RHS of eqn. 7.113 is simply $\frac{\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\,\tau^y$. The third term is

$$\begin{aligned}\frac{i\hbar}{8m^2c^2}\big[\sigma^\mu\pi^\mu, \sigma^\nu\partial^\nu V\big] &= \frac{i\hbar}{8m^2c^2}\Big\{\sigma^\mu\,\big[\pi^\mu, \sigma^\nu\partial^\nu V\big] + \big[\sigma^\mu, \sigma^\nu\partial^\nu V\big]\,\pi^\mu\Big\} \\ &= \frac{i\hbar}{8m^2c^2}\Big\{\frac{\hbar}{i}\partial^\mu\partial^\nu V\,\sigma^\mu\sigma^\nu + 2i\epsilon^{\mu\nu\lambda}\sigma^\lambda\partial^\nu V\,\pi^\mu\Big\} \\ &= \frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\times\boldsymbol{\pi} \quad .\end{aligned} \tag{7.114}$$

Therefore,

$$\begin{aligned}U^\dagger\,H\,U = \left(mc^2 + \frac{\boldsymbol{\pi}^2}{2m} + \frac{e\hbar}{2mc}\,\boldsymbol{B}\cdot\boldsymbol{\sigma}\right)\tau^z + V + \frac{\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\,\tau^y \\ + \frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\times\boldsymbol{\pi} + \mathcal{O}(m^{-3}) \quad .\end{aligned} \tag{7.115}$$

This is not block-diagonal, owing to the last term on the RHS of the top line. We can eliminate this term by effecting yet another unitary transformation. However, this will result in a contribution to the energy of order m^{-3}, so we can neglect it. To substantiate this last claim, drop all the block-diagonal terms except for the leading order one, $mc^2\,\tau^z$, and consider the Hamiltonian

$$K = mc^2\,\tau^z + \frac{\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\,\tau^y\,. \tag{7.116}$$

We now know how to bring this to block-diagonal form. The result is

$$\widetilde{K} = mc^2\sqrt{1 + \left(\frac{\hbar\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V}{2m^2c^3}\right)^2}\,\tau^z = \left(mc^2 + \frac{\hbar^2(\boldsymbol{\nabla}\,V)^2}{8m^3c^4} + \ldots\right)\tau^z \quad , \tag{7.117}$$

and the correction is of order m^{-3}, as promised.

We now assume all the negative energy ($\tau^z = -1$) states are filled. The Hamiltonian for the electrons, valid to $\mathcal{O}(m^{-3})$, is then

$$\widetilde{H} = mc^2 + V + \frac{\boldsymbol{\pi}^2}{2m} + \frac{e\hbar}{2mc}\,\boldsymbol{B}\cdot\boldsymbol{\sigma} + \frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\times\boldsymbol{\pi} + \mathcal{O}(m^{-3}) \quad . \tag{7.118}$$