APPPHYS225 - Thursday 13 November 2008

We begin class today by reviewing the representation of symmetry transformations in general, and of rotations of Cartesian 3-space in particular. These ideas will be fundamental for the material of our next few classes.

Symmetry groups and group representations

Let's begin by reviewing the algebraic notion of a *group*. A group consists of a set of elements (of arbitrary type), together with a multiplication rule, which satisfy the following properties:

1. The set must be closed under its multiplication rule. That is, if a and b are elements of the group, ab and ba must also be in the group. If $ab = ba$ for every pair of elements, the group is called 'Abelian' (or 'commutative').
2. The multiplication rule must be associative, that is, $(ab)c = a(bc)$.
3. The set must contain an identity element e, such that $ae = ea = a$ for all a.
4. Each element a must have an inverse a^{-1}, such that $a^{-1}a = aa^{-1} = e$.

A simple example of an Abelian group is the pair of numbers $\{-1, 1\}$ under normal multiplication. Another example is the set of matrices

$$\left\{ \begin{array}{l} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{array} \right\}$$

under matrix multiplication, which form a non-commutative group. Note that the pair of matrices

$$\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

form an Abelian group (under matrix multiplication) that is *isomorphic* to our first example. The essential structure of a group is its multiplication 'table,' which for the example $\{a \leftrightarrow -1, e \leftrightarrow 1\}$ is clearly

$$\begin{array}{|c|} \hline aa = e \\ \hline ae = a \\ \hline ea = a \\ \hline ee = e \\ \hline \end{array} .$$

We see that the pair of matrices has exactly the same multiplication table, under the obvious mapping

$$\left\{ a \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

In this sense it is really a 'representation' of the same group, which by mathematical convention is generally referred to as C_2. We shall use the term *linear representation* (or simply, representation) to mean an association (not necessarily one-to-one) between each element a in a group and a matrix $D(a)$ that preserves the multiplication rule:

$$D(a)D(b) = D(ab).$$

So far we have seen how C_2 can be represented by 1×1 real matrices (numbers) and 2×2 real matrices. One often speaks of these as being representations 'on' the vector spaces $\mathbf{R}^1$ and $\mathbf{R}^2$, respectively.

To emphasize the abstract nature of the concept, let's think about the *dihedral group*, denoted D_2. Its multiplication table is

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

This is clearly Abelian. With reference to the following figure [Wu-Ki Tung, *Group Theory in Physics* (World Scientific, 1985)]

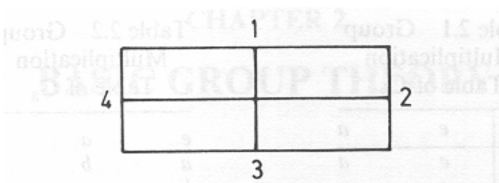

Fig. 2.1 A configuration with D_2 symmetry.

we can associate group elements with symmetries of the rectangle,

e	leave the figure unchanged,
a	reflect through the vertical axis $1 - 3$,
b	reflect through the horizontal axis $2 - 4$,
c	rotate (in the plane) about the center point by angle π.

We can easily find a representation of this group on $\mathbf{R}^2$,

$$e \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
$$a \leftrightarrow \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad c \leftrightarrow \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Sometimes we find that a group representation 'contains' smaller representations within it. For example, consider our representation of C_2 on $\mathbf{R}^2$, with

$$\left\{ a \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

We first note that a linear change of basis for the representing vector space preserves group representations, since under

$$D(a) \mapsto S^{-1}D(a)S,$$

the mapping property

$$D(a)D(b) = D(ab)$$

transforms according to

$$\begin{aligned} S^{-1}D(a)SS^{-1}D(b)S &= S^{-1}D(ab)S, \\ S^{-1}D(a)D(b)S &= S^{-1}D(ab)S, \\ D(a)D(b) &= D(ab). \end{aligned}$$

Two representations related by such a similarity transform are said to be *equivalent*. If in our example we switch to the basis

$$\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\},$$

for $\mathbf{R}^2$, our C_2 representation maps to

$$\left\{ a \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

This clearly still respects the C_2 multiplication table

$$\begin{matrix} e & a \\ a & e \end{matrix}.$$

We see that the $\{a \leftrightarrow -1, e \leftrightarrow 1\}$ representation on $\mathbf{R}^1$ actually appears as the lower-right 'corner' of the $\mathbf{R}^2$ representation, living in its own little subspace and constituting a 1×1 *subrepresentation* of the 2×2 one. What about the upper-left subspace? Here we find the *degenerate* representation $\{a \leftrightarrow 1, e \leftrightarrow 1\}$ of C_2 on $\mathbf{R}^1$, which is not a terribly useful representation but a valid one none-the-less.

In general, it is often possible to take an $n \times n$ representation and to find a basis in which all the constituent matrices assume block-diagonal form

$$D_n(a) = \begin{pmatrix} D_1(a) & 0 & 0 & \cdots \\ 0 & D_2(a) & 0 & \cdots \\ 0 & 0 & D_3(a) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $D_1(a)$ has dimensions $n_1 \times n_1$, $D_2(a)$ has dimensions $n_2 \times n_2$, etc. and $n_1 + n_2 + n_3 + \cdots = n$. Each of the sets of matrices $\{D_i(a)\}$ then forms a subrepresentation of the original one, which is therefore said to be *reducible*, and we write

$$D_n = D_1 \oplus D_2 \oplus D_3 \oplus \cdots,$$

i.e., each $D_n(a)$ is the direct sum of the $D_i(a)$. A representation that cannot be broken down in this way is called an *irreducible representation*, or *irrep* for short. Up to equivalence transformations, the decomposition of a representation into irreps is unique.

It is important to appreciate that we can use the direct sum to build larger representations out of smaller ones. For example, working again with representations of C_2, and denoting our previous representations as d_1 (on $\mathbf{R}^1$) and d_2 (on $\mathbf{R}^2$), we have

$$d_1 \oplus d_1 : \left\{ a \leftrightarrow \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

$$d_1 \oplus d_2 : \left\{ a \leftrightarrow \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\},$$

$$d_2 \oplus d_2 : \left\{ a \leftrightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right\},$$

$$d_1 \oplus d_1 \oplus d_1 : \left\{ a \leftrightarrow \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, e \leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\},$$

and so on.

Another important way to build larger representation is by taking the tensor product (also known as direct product) of smaller ones. Recall that for a pair of matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots \\ b_{21} & b_{22} & b_{23} & \cdots \\ b_{31} & b_{32} & b_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

the tensor product $A \otimes B$ is given by

$$\begin{pmatrix} a_{11}B & a_{12}B & a_{13}B & \cdots \\ a_{21}B & a_{22}B & a_{23}B & \cdots \\ a_{31}B & a_{32}B & a_{33}B & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

For 2×2 matrices, *e.g.*, we can write more concretely

$$A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}.$$

We can verify that $d_2 \otimes d_2$ for our C_2 example still constitutes a valid representation, which happens to be equivalent to $d_2 \oplus d_2$.

$$a \leftrightarrow \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad e \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$a^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = e.$$

General proofs regarding composite representations can be found in standard textbooks on group representation theory (such as Tung, referenced above).

A useful theorem for recognizing irreducible representations is given by Schur's Second Lemma, which states that [Merzbacher, p. 421] *if the matrices $D(a)$ form an irreducible representation of a group and if a matrix M commutes with all $D(a)$,*

$$[M, D(a)] = 0 \qquad \text{for every } a$$

then M is a multiple of the identity matrix. Hence, if for a given representation we can find a commuting matrix C that is not simply proportional to the identity, we know that the representation is reducible. In general (provided C is normal), there will be some change of basis (corresponding to diagonalizing C) that takes C to a matrix of the form

$$C \mapsto \begin{pmatrix} c_1 I_1 & 0 & 0 & \cdots \\ 0 & c_2 I_2 & 0 & \cdots \\ 0 & 0 & c_3 I_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where the I_i are identity matrices of dimension $n_i \times n_i$. It follows that each of the n_i –dimensional subspaces (after the change of basis) constitutes an subrepresentation.

Representations of the Rotation Group $R(3)$

It is easy to see that the set of all possible rotations in three-dimensional Euclidean space form a group. Any particular member of the group may be specified by its angle and axis of rotation, group multiplication corresponds to the successive application of the two rotations, and the identity element can be chosen as any rotation by zero angle (about any axis). For example, if we let $R_{\hat{n}}(\varphi)$ denote a clockwise rotation by angle φ about the unit vector $\hat{n}$,

$$R_{\hat{y}}(\pi) R_{\hat{x}}(\pi) = R_{\hat{z}}(\pi).$$

Closure and associativity clearly hold, and this group $R(3)$ is clearly not commutative. For example,

$$R_{\hat{x}}\left(-\frac{\pi}{2}\right) R_{\hat{x}}\left(\frac{\pi}{2}\right) = R_{\hat{y}}\left(-\frac{\pi}{2}\right) R_{\hat{y}}\left(\frac{\pi}{2}\right) = 1, \quad R_{\hat{x}}\left(-\frac{\pi}{2}\right) R_{\hat{y}}\left(-\frac{\pi}{2}\right) R_{\hat{x}}\left(\frac{\pi}{2}\right) R_{\hat{y}}\left(\frac{\pi}{2}\right) \neq 1.$$

Let us consider the association

$$(\hat{n}, \varphi) \mapsto \exp\left(\frac{-i}{\hbar} \hat{n} \cdot \vec{\mathbf{J}} \varphi\right),$$

where $(\hat{n}, \varphi) \in R(3)$ is a rotation about axis $\hat{n}$ by angle φ ($\hat{n}$ is assumed to be a unit vector) and $\vec{\mathbf{J}}$ is a vector of angular momentum operators $(\mathbf{J}_x, \mathbf{J}_y, \mathbf{J}_z)$. Any set of three operators $\{\mathbf{J}_x, \mathbf{J}_y, \mathbf{J}_z\}$ satisfying the fundamental commutation relations

$$[\mathbf{J}_x, \mathbf{J}_y] = i\hbar \mathbf{J}_z, \quad [\mathbf{J}_y, \mathbf{J}_z] = i\hbar \mathbf{J}_x, \quad [\mathbf{J}_z, \mathbf{J}_x] = i\hbar \mathbf{J}_y,$$

are legitimate angular momentum operators, with two familiar examples being $\vec{\mathbf{J}} = \vec{\mathbf{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$ (orbital angular momentum) and $\vec{\mathbf{J}} = \vec{\mathbf{S}}$ (spin-1/2). The reason for this is simply that any set of generators $\{\mathbf{J}_x, \mathbf{J}_y, \mathbf{J}_z\}$ satisfying the above commutation relations can be used in the above association between rotations and Hilbert-space operators to yield a valid representation of $R(3)$. Preservation of the appropriate multiplication table is guarenteed by the commutation relations and the operator-exponential structure of the association. The proof of this is straightforward but quite tedious – see for example Cohen-Tannoudji, Diu, and Laloe Complement B_{VI} for a patient and careful treatment.

In the case of spin-1/2, we clearly have a 2×2 representation of $R(3)$ on $\mathbf{C}^2$ (the two-dimensional complex vector space, not the two-element group!), as the generators are simply proportional to the Pauli matrices

$$\mathbf{J}_i = \frac{\hbar}{2}\boldsymbol{\sigma}_i, \quad \boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Application of the association law yields

$$(\hat{n}, \varphi) \mapsto \begin{pmatrix} \cos\frac{\varphi}{2} - in_z \sin\frac{\varphi}{2} & (-in_x - n_y)\sin\frac{\varphi}{2} \\ (-in_x + n_y)\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} + in_z \sin\frac{\varphi}{2} \end{pmatrix},$$

where $\hat{n} = (n_x, n_y, n_z)$. When the state of a spin-1/2 system is expressed in the σ_z basis, these unitary matrices are the Hilbert-space operators corresponding to rotations in the $3D$ coordinate space.

Suppose we have two spin-1/2 particles A and B. Then the joint Hilbert space $H_A \otimes H_B$ supports a tensor-product representation of the rotation group generated by the angular momentum operators

$$\mathbf{J}_x = \mathbf{J}_x^A \otimes \mathbf{1}^B + \mathbf{1}^A \otimes \mathbf{J}_x^B, \quad \mathbf{J}_y = \mathbf{J}_y^A \otimes \mathbf{1}^B + \mathbf{1}^A \otimes \mathbf{J}_y^B, \quad \mathbf{J}_z = \mathbf{J}_z^A \otimes \mathbf{1}^B + \mathbf{1}^A \otimes \mathbf{J}_z^B.$$

It is straightforward to show that this set of operators satisfies the required commutation relations.

Writing these out in matrix form, we have

$$\mathbf{J}_x \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$= \frac{\hbar}{2}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + \frac{\hbar}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

$$\mathbf{J}_y \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$= \frac{\hbar}{2}\begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} + \frac{\hbar}{2}\begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i & -i & 0 \\ i & 0 & 0 & -i \\ i & 0 & 0 & -i \\ 0 & i & i & 0 \end{pmatrix},$$

$$\mathbf{J}_z \leftrightarrow \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\otimes\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\otimes\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

$$= \frac{\hbar}{2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}+\frac{\hbar}{2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}=\frac{\hbar}{2}\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -2\end{pmatrix}.$$

This allows us to compute the Casimir operator (the C in our above discussion of Schur's Lemma)

$$\mathbf{J}^2 = \mathbf{J}_x^2 + \mathbf{J}_y^2 + \mathbf{J}_z^2$$

$$\leftrightarrow \frac{\hbar^2}{4}\left\{\begin{pmatrix}0 & 1 & 1 & 0\\ 1 & 0 & 0 & 1\\ 1 & 0 & 0 & 1\\ 0 & 1 & 1 & 0\end{pmatrix}^2+\begin{pmatrix}0 & -i & -i & 0\\ i & 0 & 0 & -i\\ i & 0 & 0 & -i\\ 0 & i & i & 0\end{pmatrix}^2+\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -2\end{pmatrix}^2\right\}$$

$$= \frac{\hbar^2}{4}\left\{\begin{pmatrix}2 & 0 & 0 & 2\\ 0 & 2 & 2 & 0\\ 0 & 2 & 2 & 0\\ 2 & 0 & 0 & 2\end{pmatrix}+\begin{pmatrix}2 & 0 & 0 & -2\\ 0 & 2 & 2 & 0\\ 0 & 2 & 2 & 0\\ -2 & 0 & 0 & 2\end{pmatrix}+\begin{pmatrix}4 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 4\end{pmatrix}\right\}=\hbar^2\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 2\end{pmatrix}.$$

We can easily read off the eigenvectors,

$$2\hbar^2 \leftrightarrow \left\{\begin{pmatrix}1\\ 0\\ 0\\ 0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 1\\ 1\\ 0\end{pmatrix}, \begin{pmatrix}0\\ 0\\ 0\\ 1\end{pmatrix}\right\}, \quad 0 \leftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 1\\ -1\\ 0\end{pmatrix},$$

so the diagonalizing transformation is

$$S = \begin{pmatrix}0 & 1 & 0 & 0\\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 0 & 1\end{pmatrix}, \quad \mathbf{J}^2 \mapsto S^{-1}(\mathbf{J}^2)S = \hbar^2\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 2\end{pmatrix}.$$

Then we have also with this transformation

$$\mathbf{J}_x \mapsto \hbar\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad \mathbf{J}_y \mapsto \hbar\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{i}{\sqrt{2}} & 0 \\ 0 & \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ 0 & 0 & \frac{i}{\sqrt{2}} & 0 \end{pmatrix}, \quad \mathbf{J}_z \mapsto \hbar\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

We now see that this is a reducible representation, with a trivial representation

$$D(\mathbf{J}_x) = D(\mathbf{J}_y) = D(\mathbf{J}_z) = 0,$$

in the upper-left corner and a three-dimensional representation in the lower-right block. Remembering that the eigenvalues of the $\mathbf{J}^2$ operator have the form

$$\mathbf{J}^2|j,m\rangle = j(j+1)\hbar^2|j,m\rangle,$$

we can confirm from the matrix representation of $\mathbf{J}^2$ in the diagonal basis that this is a spin-1 irrep. Our notation here emphasizes the fact that this basis is simultaneously an eigenbasis of $\mathbf{J}^2$ and $\mathbf{J}_z$, with

$$\mathbf{J}_z|j,m\rangle = m\hbar|j,m\rangle.$$

It follows that if we use this simultaneous eigenbasis of $\mathbf{J}^2$ and $\mathbf{J}_z$,

$$|0,0\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \leftrightarrow \frac{1}{\sqrt{2}}(|+_A\rangle \otimes |-_B\rangle - |-_A\rangle \otimes |+_B\rangle),$$

$$|1,1\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \leftrightarrow |+_A\rangle \otimes |+_B\rangle, \quad |1,-1\rangle \leftrightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \leftrightarrow |-_A\rangle \otimes |-_B\rangle,$$

$$|1,0\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \leftrightarrow \frac{1}{\sqrt{2}}(|-_A\rangle \otimes |+_B\rangle + |+_A\rangle \otimes |-_B\rangle),$$

then the representation of any rotation on our $H_A \otimes H_B$ will take the block-diagonal form

$$(\hat{n},\varphi) \mapsto \exp\left(\frac{-i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right) \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \end{pmatrix},$$

where the one in the upper-left corner is the exponential of zero and the X's in the lower-right block represent a 3×3 matrix that in general we must calculate. Most important, however, is that if we write our states $|\Psi_{AB}\rangle \in H_A \otimes H_B$ in the simultaneous eigenbasis of $\mathbf{J}^2$ and $\mathbf{J}_z$,

$$|\Psi_{AB}\rangle = c_{00}|0,0\rangle + c_{1-}|1,-1\rangle + c_{10}|1,0\rangle + c_{1+}|1,1\rangle,$$

then

$$\exp\left(\frac{-i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right)|\Psi_{AB}\rangle \sim c_{00}|0,0\rangle + c'_{1-}|1,-1\rangle + c'_{10}|1,0\rangle + c'_{1+}|1,1\rangle,$$

where the set of three primed coefficients are linearly (unitarily) related to their unprimed counterparts and

$$|c_{1-}|^2 + |c_{10}|^2 + |c_{1+}|^2 = |c'_{1-}|^2 + |c'_{10}|^2 + |c'_{1+}|^2.$$

Put another way, if we define

$$\mathbf{\Pi}_0 = |0,0\rangle\langle 0,0|, \quad \mathbf{\Pi}_1 = |1,-1\rangle\langle 1,-1| + |1,0\rangle\langle 1,0| + |1,1\rangle\langle 1,1|,$$

then

$$\exp\left(\frac{-i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right)|\Psi_{AB}\rangle = \mathbf{\Pi}_0|\Psi_{AB}\rangle + \mathbf{\Pi}_1\exp\left(\frac{-i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right)\mathbf{\Pi}_1|\Psi_{AB}\rangle,$$

$$\exp\left(\frac{-i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right)\boldsymbol{\rho}\exp\left(\frac{i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right) = \mathbf{\Pi}_0\boldsymbol{\rho}\mathbf{\Pi}_0 + \mathbf{\Pi}_1\exp\left(\frac{-i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right)\mathbf{\Pi}_1\boldsymbol{\rho}\mathbf{\Pi}_1\exp\left(\frac{i}{\hbar}\hat{n}\cdot\vec{\mathbf{J}}\varphi\right)\mathbf{\Pi}_1.$$

Hence if we start with a state that lives entirely in the $j = 1$ subspace of $H_A \otimes H_B$ there is no rotation we can apply that will cause it to acquire components in the $j = 0$ subspace, and *vice versa*.

For what follows it is important to note that the $j = 0$ state is antisymmetric under exchange of the two spin-1/2 particles, whereas the $j = 1$ states are all symmetric. Hence for any state of the form

$$|\Psi_{AB}\rangle = c_{++}|+_A \ +_B\rangle + c_{+-}(|+_A \ -_B\rangle + |-_A \ +_B\rangle) + c_{--}|-_A \ -_B\rangle,$$

we have

$$\mathbf{\Pi}_1|\Psi_{AB}\rangle = |\Psi_{AB}\rangle.$$

For factorizable pure states we see that

$$|\psi_A\rangle \otimes |\varphi_B\rangle = (a_+|+_A\rangle + a_-|-_A\rangle) \otimes (b_+|+_B\rangle + b_-|-_B\rangle)$$
$$= a_+b_+|+_A \ +_B\rangle + a_+b_-|+_A \ -_B\rangle + a_-b_+|-_A \ +_B\rangle + a_+b_+|-_A \ -_B\rangle,$$

which will be contained entirely within the $j = 1$ subspace of $H_A \otimes H_B$ if

$$a_+b_- = a_-b_+.$$

This is obviously satisfied if $a_\pm = b_\pm$, *i.e.*, if $|\psi\rangle = |\varphi\rangle$.

Encoding 'unspeakable information' in quantum states

In this section we follow the paper by N. Gisin and S. Popescu, "Spin Flips and Quantum Information for Antiparallel Spins," Phys. Rev. Lett. **83**, 432 (1999).

Suppose Alice and Bob lack a shared Cartesian reference frame. However, Alice has two spin-1/2 systems that she knows she can send to Bob without causing them to suffer any physical perturbations in transit (their Hamiltonian vanishes exactly).

Alice wants to use her two spin-1/2 systems to indicate a direction in physical space - she wants to send Bob a 'pointer' of sorts. This kind of information is sometimes called 'unspeakable' because, for example, it cannot be conveyed over the phone between parties who do not share a coordinate reference frame. So what state of the spins should Alice prepare?

If Alice's direction is represented by the unit vector $\hat{n} = (n_x, n_y, n_z)$ in her own local coordinate frame, one obvious strategy would be to prepare the state in which both spins satisfy

$$\langle \mathbf{J}_x \rangle = \frac{\hbar}{2} n_x, \quad \langle \mathbf{J}_y \rangle = \frac{\hbar}{2} n_y, \quad \langle \mathbf{J}_z \rangle = \frac{\hbar}{2} n_z.$$

We use the Dirac ket $|\hat{n}; \hat{n}\rangle$ to denote this state. In fact, it has been pointed out by Gisin and Popescu that it is provably better for Alice to prepare the state $|\hat{n}; -\hat{n}\rangle$, in which the first spin satisfies the set of expectation values above while the second satisfies

$$\langle \mathbf{J}_x \rangle = -\frac{\hbar}{2} n_x, \quad \langle \mathbf{J}_y \rangle = -\frac{\hbar}{2} n_y, \quad \langle \mathbf{J}_z \rangle = -\frac{\hbar}{2} n_z.$$

Classically it would seem that there should be no difference between using parallel or antiparallel spins.

In order to understand this intriguing fact we first need to ask what sort of measurement Bob should make in order to 'decode' the directional information carried by Alice's spins. Let us continue to use $\hat{n}$ to denote the direction in space that Alice has in mind, and let $\hat{n}_g$ be the direction that Bob guesses on the basis of a measurement he performs on the pair of spins. We imagine that Bob's measurement will have some outcome g drawn from a finite set, and that some optimal association of measurement outcomes with corresponding directional guesses $\hat{n}_g$ has been computed. Of course in order to talk about optimality we must fix a figure of merit for the accuracy of Bob's guess, such as the average fidelity (assuming uniform prior for Alice's $\hat{n}$ over the entire unit sphere)

$$F = \int d\hat{n} \sum_g P(g|\hat{n}) \frac{1 + \hat{n} \cdot \hat{n}_g}{2}.$$

Here we imagine that Bob's measurement is specified by a POVM $\{E_g\}$, so that $P(g|\hat{n}) = \langle \Psi | E_g | \Psi \rangle$, where $|\Psi\rangle$ is the state in which Alice prepares the spins. It turns out that the optimal measurement for this purpose, assuming $|\Psi\rangle$ takes the form $|\hat{n}; \hat{n}\rangle$, has four elements of the form $(j \in 1 \ldots 4)$

$$E_j = |\phi_j\rangle\langle\phi_j|, \quad |\phi_j\rangle = \frac{\sqrt{3}}{2} |\hat{n}_j; \hat{n}_j\rangle + \frac{1}{2} |\psi^-\rangle,$$

where $|\psi^-\rangle$ is the Bell singlet state and $|\hat{n}_j; \hat{n}_j\rangle$ is a spin state in which both spins point in direction $\hat{n}_j$. The unit vectors (specified in Cartesian 3-space) are

$$\hat{n}_1 = (0, 0, 1), \quad \hat{n}_2 = \left(\frac{\sqrt{8}}{3}, 0, -\frac{1}{3} \right),$$

$$\hat{n}_3 = \left(\frac{-\sqrt{2}}{3}, \sqrt{\frac{2}{3}}, -\frac{1}{3} \right), \quad \hat{n}_4 = \left(\frac{-\sqrt{2}}{3}, -\sqrt{\frac{2}{3}}, -\frac{1}{3} \right).$$

Note that these correspond to the vertices of a tetrahedron. Bob's guessing procedure is to perform the $\{E_j\}$ POVM on the spins he receives from Alice, and upon obtaining outcome j he makes the guess $\hat{n}_g = \hat{n}_j$. It is stated in the paper that with this procedure, $F = 3/4$ for the parallel-spins encoding strategy.

With anti-parallel spins, Bob's can use a guessing procedure that is basically the same but with the modified POVM

$$E_j = |\theta_j\rangle\langle\theta_j|, \quad |\theta_j\rangle = \alpha|\hat{n}_j;-\hat{n}_j\rangle - \beta\sum_{k\neq j}|\hat{n}_k;-\hat{n}_k\rangle,$$

where $\alpha = 13/\left(6\sqrt{6} - 2\sqrt{2}\right)$ and $\beta = \left(5 - 2\sqrt{3}\right)/\left(6\sqrt{6} - 2\sqrt{2}\right)$. The correspdonding value of F is $\left(5\sqrt{3} + 33\right)/\left[3\left(3\sqrt{3} - 1\right)^2\right] \approx 0.789$, which clearly beats the average fidelity in the parallel-spins protocol. The paper by Gisin and Popescu does not say whether or not this POVM is optimal.

At a qualitative level, there is in fact a simple explanation of this difference between parallel and anti-parallel spins. Note that Bob's task can be viewed as one of estimating a rotation operation that distinguishes Alices direction $\hat{n}$ from some some fiducial direction that Bob defines (for example) as his local x-axis (note that since we are only talking about a single direction and not a full set of Cartesian axes, this is technically an estimation problem in the coset $R(3)/U(1)$). The measurement that Bob performs on the spins he receives gives him information about the identity of this rotation. How much information? It cannot be any more than $\log_2(4)$ bits of information since the POVM he performs has only four outcomes, and indeed the joint quantum state of the two spins lives in a four-dimensional Hilbert space. However in the parallel-spins encoding strategy we note that in fact it cannot be more than $\log_2(3)$ bits of information, because a symmetric state such as $|\hat{n};\hat{n}\rangle$ is contained entirely within the $j = 1$ subspace of the joint Hilbert space of the two spins. Therefore, no rotation can cause it to acquire a $j = 0$ component and the state space of the two spins is then *effectively* three-dimensional. In contrast, a state such as $|\hat{n};-\hat{n}\rangle$ spans both the $j = 0$ and $j = 1$ subspaces, and can therefore 'fill out' more of the joint Hilbert space under the action of $R(3)$ rotations. Apparently it does so, by at least a little bit.

To see that $|\hat{n};-\hat{n}\rangle$ spans both subspaces, we can adopt a coordinate system in which the z-axis corresponds with $\hat{n}$. Then

$$|\hat{n};-\hat{n}\rangle \mapsto |+_A \; -_B\rangle = \frac{1}{\sqrt{2}}(|j=0,m=0\rangle + |j=1,m=0\rangle).$$

No matter what other coordinate frame you might prefer for performing your calculations, the expression for $|\hat{n};-\hat{n}\rangle$ in terms of your preferred basis states will be related to the above by some $R(3)$ rotation operator that preserves the superposition across the $j = 0$ and $j = 1$ subspaces!