2.5 The Schmidt decomposition and purifications

Density operators and the partial trace are just the beginning of a wide array of tools useful for the study of composite quantum systems, which are at the heart of quantum computation and quantum information. Two additional tools of great value are the *Schmidt decomposition* and *purifications*. In this section we present both these tools, and try to give the flavor of their power.

Theorem 2.7: (**Schmidt decomposition**) Suppose $|\psi\rangle$ is a pure state of a composite system, AB. Then there exist orthonormal states $|i_A\rangle$ for system A, and orthonormal states $|i_B\rangle$ of system B such that

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle, \tag{2.202}$$

where λ_i are non-negative real numbers satisfying $\sum_i \lambda_i^2 = 1$ known as *Schmidt co-efficients*.

This result is very useful. As a taste of its power, consider the following consequence: let $|\psi\rangle$ be a pure state of a composite system, AB. Then by the Schmidt decomposition $\rho^A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$ and $\rho^B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$, so the eigenvalues of ρ^A and ρ^B are identical, namely λ_i^2 for both density operators. Many important properties of quantum systems are completely determined by the eigenvalues of the reduced density operator of the system, so for a pure state of a composite system such properties will be the same for both systems. As an example, consider the state of two qubits, $(|00\rangle + |01\rangle + |11\rangle)/\sqrt{3}$. This has no obvious symmetry property, yet if you calculate $\mathrm{tr}\left((\rho^A)^2\right)$ and $\mathrm{tr}\left((\rho^B)^2\right)$ you will discover that they have the same value, $7/9$ in each case. This is but one small consequence of the Schmidt decomposition.

Proof

We give the proof for the case where systems A and B have state spaces of the same dimension, and leave the general case to Exercise 2.76. Let $|j\rangle$ and $|k\rangle$ be any fixed orthonormal bases for systems A and B, respectively. Then $|\psi\rangle$ can be written

$$|\psi\rangle = \sum_{jk} a_{jk} |j\rangle|k\rangle, \tag{2.203}$$

for some matrix a of complex numbers a_{jk}. By the singular value decomposition, $a = udv$, where d is a diagonal matrix with non-negative elements, and u and v are unitary matrices. Thus

$$|\psi\rangle = \sum_{ijk} u_{ji} d_{ii} v_{ik} |j\rangle|k\rangle. \tag{2.204}$$

Defining $|i_A\rangle \equiv \sum_j u_{ji}|j\rangle$, $|i_B\rangle \equiv \sum_k v_{ik}|k\rangle$, and $\lambda_i \equiv d_{ii}$, we see that this gives

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle. \tag{2.205}$$

It is easy to check that $|i_A\rangle$ forms an orthonormal set, from the unitarity of u and the orthonormality of $|j\rangle$, and similarly that the $|i_B\rangle$ form an orthonormal set. □

Exercise 2.76: Extend the proof of the Schmidt decomposition to the case where A and B may have state spaces of different dimensionality.

Exercise 2.77: Suppose ABC is a three component quantum system. Show by example that there are quantum states $|\psi\rangle$ of such systems which can not be written in the form

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle, \tag{2.206}$$

where λ_i are real numbers, and $|i_A\rangle, |i_B\rangle, |i_C\rangle$ are orthonormal bases of the respective systems.

The bases $|i_A\rangle$ and $|i_B\rangle$ are called the *Schmidt bases* for A and B, respectively, and the number of non-zero values λ_i is called the *Schmidt number* for the state $|\psi\rangle$. The Schmidt number is an important property of a composite quantum system, which in some sense quantifies the 'amount' of entanglement between systems A and B. To get some idea of why this is the case, consider the following obvious but important property: the Schmidt number is preserved under unitary transformations on system A or system B alone. To see this, notice that if $\sum_i \lambda_i |i_A\rangle|i_B\rangle$ is the Schmidt decomposition for $|\psi\rangle$ then $\sum_i \lambda_i (U|i_A\rangle)|i_B\rangle$ is the Schmidt decomposition for $U|\psi\rangle$, where U is a unitary operator acting on system A alone. Algebraic invariance properties of this type make the Schmidt number a very useful tool.

Exercise 2.78: Prove that a state $|\psi\rangle$ of a composite system AB is a product state if and only if it has Schmidt number 1. Prove that $|\psi\rangle$ is a product state if and only if ρ^A (and thus ρ^B) are pure states.

A second, related technique for quantum computation and quantum information is *purification*. Suppose we are given a state ρ^A of a quantum system A. It is possible to introduce another system, which we denote R, and *define* a *pure state* $|AR\rangle$ for the joint system AR such that $\rho^A = \text{tr}_R(|AR\rangle\langle AR|)$. That is, the pure state $|AR\rangle$ reduces to ρ^A when we look at system A alone. This is a purely mathematical procedure, known as *purification*, which allows us to associate pure states with mixed states. For this reason we call system R a *reference* system: it is a fictitious system, without a direct physical significance.

To prove that purification can be done for *any* state, we explain how to construct a system R and purification $|AR\rangle$ for ρ^A. Suppose ρ^A has orthonormal decomposition $\rho^A = \sum_i p_i |i^A\rangle\langle i^A|$. To purify ρ^A we introduce a system R which has the same state space as system A, with orthonormal basis states $|i^R\rangle$, and define a pure state for the combined system

$$|AR\rangle \equiv \sum_i \sqrt{p_i}|i^A\rangle|i^R\rangle. \tag{2.207}$$

We now calculate the reduced density operator for system A corresponding to the state $|AR\rangle$:

$$\text{tr}_R(|AR\rangle\langle AR|) = \sum_{ij} \sqrt{p_i p_j}|i^A\rangle\langle j^A| \,\text{tr}(|i^R\rangle\langle j^R|) \tag{2.208}$$

$$= \sum_{ij} \sqrt{p_i p_j}|i^A\rangle\langle j^A| \,\delta_{ij} \tag{2.209}$$

$$= \sum_i p_i |i^A\rangle\langle i^A| \tag{2.210}$$

$$= \rho^A. \tag{2.211}$$

Thus $|AR\rangle$ is a purification of ρ^A.

Notice the close relationship of the Schmidt decomposition to purification: the procedure used to purify a mixed state of system A is to define a pure state whose Schmidt basis for system A is just the basis in which the mixed state is diagonal, with the Schmidt coefficients being the square root of the eigenvalues of the density operator being purified.

In this section we've explained two tools for studying composite quantum systems, the Schmidt decomposition and purifications. These tools will be indispensable to the study of quantum computation and quantum information, especially quantum information, which is the subject of Part III of this book.

Exercise 2.79: Consider a composite system consisting of two qubits. Find the Schmidt decompositions of the states

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}; \quad \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2}; \text{ and } \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}. \tag{2.212}$$

Exercise 2.80: Suppose $|\psi\rangle$ and $|\varphi\rangle$ are two pure states of a composite quantum system with components A and B, with identical Schmidt coefficients. Show that there are unitary transformations U on system A and V on system B such that $|\psi\rangle = (U \otimes V)|\varphi\rangle$.

Exercise 2.81: (Freedom in purifications) Let $|AR_1\rangle$ and $|AR_2\rangle$ be two purifications of a state ρ^A to a composite system AR. Prove that there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I_A \otimes U_R)|AR_2\rangle$.

Exercise 2.82: Suppose $\{p_i, |\psi_i\rangle\}$ is an ensemble of states generating a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ for a quantum system A. Introduce a system R with orthonormal basis $|i\rangle$.

(1) Show that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification of ρ.
(2) Suppose we measure R in the basis $|i\rangle$, obtaining outcome i. With what probability do we obtain the result i, and what is the corresponding state of system A?
(3) Let $|AR\rangle$ be *any* purification of ρ to the system AR. Show that there exists an orthonormal basis $|i\rangle$ in which R can be measured such that the corresponding post-measurement state for system A is $|\psi_i\rangle$ with probability p_i.

2.6 EPR and the Bell inequality

Anybody who is not shocked by quantum theory has not understood it.
– Niels Bohr

I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it. The rest of this walk was devoted to a discussion of what a physicist should mean by the term 'to exist'.
– Abraham Pais

...quantum phenomena do not occur in a Hilbert space, they occur in a laboratory.
– Asher Peres

...what is proved by impossibility proofs is lack of imagination.
– John Bell

This chapter has focused on introducing the tools and mathematics of quantum mechanics. As these techniques are applied in the following chapters of this book, an important recurring theme is the unusual, *non-classical* properties of quantum mechanics. But what exactly is the difference between quantum mechanics and the classical world? Understanding this difference is vital in learning how to perform information processing tasks that are difficult or impossible with classical physics. This section concludes the chapter with a discussion of the Bell inequality, a compelling example of an essential difference between quantum and classical physics.

When we speak of an object such as a person or a book, we assume that the physical properties of that object have an existence independent of observation. That is, measurements merely act to *reveal* such physical properties. For example, a tennis ball has as one of its physical properties its *position*, which we typically measure using light scattered from the surface of the ball. As quantum mechanics was being developed in the 1920s and 1930s a strange point of view arose that differs markedly from the classical view. As described earlier in the chapter, according to quantum mechanics, an unobserved particle does not possess physical properties that exist independent of observation. Rather, such physical properties arise as a consequence of measurements performed upon the system. For example, according to quantum mechanics a qubit does not possess definite properties of 'spin in the z direction, σ_z', and 'spin in the x direction, σ_x', each of which can be revealed by performing the appropriate measurement. Rather, quantum mechanics gives a set of rules which specify, given the state vector, the probabilities for the possible measurement outcomes when the observable σ_z is measured, or when the observable σ_x is measured.

Many physicists rejected this new view of Nature. The most prominent objector was Albert Einstein. In the famous 'EPR paper', co-authored with Boris Podolsky and Nathan Rosen, Einstein proposed a thought experiment which, he believed, demonstrated that quantum mechanics is not a complete theory of Nature.

The essence of the EPR argument is as follows. EPR were interested in what they termed 'elements of reality'. Their belief was that any such element of reality *must* be represented in any complete physical theory. The goal of the argument was to show that quantum mechanics is not a complete physical theory, by identifying elements of reality that were not included in quantum mechanics. The way they attempted to do this was by introducing what they claimed was a *sufficient condition* for a physical property to be an element of reality, namely, that it be possible to predict with certainty the value that property will have, immediately before measurement.

Box 2.7: Anti-correlations in the EPR experiment

Suppose we prepare the two qubit state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}, \tag{2.213}$$

a state sometimes known as the *spin singlet* for historical reasons. It is not difficult to show that this state is an entangled state of the two qubit system. Suppose we perform a measurement of spin along the $\vec{v}$ axis on both qubits, that is, we measure the observable $\vec{v} \cdot \vec{\sigma}$ (defined in Equation (2.116) on page 90) on each qubit, getting a result of $+1$ or -1 for each qubit. It turns out that no matter what choice of $\vec{v}$ we make, the results of the two measurements are always opposite to one another. That is, if the measurement on the first qubit yields $+1$, then the measurement on the second qubit will yield -1, and vice versa. It is as though the second qubit knows the result of the measurement on the first, no matter how the first qubit is measured. To see why this is true, suppose $|a\rangle$ and $|b\rangle$ are the eigenstates of $\vec{v} \cdot \vec{\sigma}$. Then there exist complex numbers $\alpha, \beta, \gamma, \delta$ such that

$$|0\rangle = \alpha|a\rangle + \beta|b\rangle \tag{2.214}$$

$$|1\rangle = \gamma|a\rangle + \delta|b\rangle. \tag{2.215}$$

Substituting we obtain

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}} = (\alpha\delta - \beta\gamma)\frac{|ab\rangle - |ba\rangle}{\sqrt{2}}. \tag{2.216}$$

But $\alpha\delta - \beta\gamma$ is the determinant of the unitary matrix $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, and thus is equal to a phase factor $e^{i\theta}$ for some real θ. Thus

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{|ab\rangle - |ba\rangle}{\sqrt{2}}, \tag{2.217}$$

up to an unobservable global phase factor. As a result, if a measurement of $\vec{v} \cdot \vec{\sigma}$ is performed on both qubits, then we can see that a result of $+1$ (-1) on the first qubit implies a result of -1 ($+1$) on the second qubit.

Consider, for example, an entangled pair of qubits belonging to Alice and Bob, respectively:

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{2.218}$$

Suppose Alice and Bob are a long way away from one another. Alice performs a measurement of spin along the $\vec{v}$ axis, that is, she measures the observable $\vec{v} \cdot \vec{\sigma}$ (defined in Equation (2.116) on page 90). Suppose Alice receives the result $+1$. Then a simple quantum mechanical calculation, given in Box 2.7, shows that she can predict with certainty that Bob will measure -1 on his qubit if he also measures spin along the $\vec{v}$ axis. Similarly, if Alice measured -1, then she can predict with certainty that Bob will measure $+1$ on his qubit. Because it is always possible for Alice to predict the value of the measurement result recorded when Bob's qubit is measured in the $\vec{v}$ direction, that physical property must correspond to an element of reality, by the EPR criterion, and should be represented in any complete physical theory. However, standard quantum mechanics, as we have presented it, merely tells one how to calculate the probabilities of the respective measurement outcomes if $\vec{v} \cdot \vec{\sigma}$ is measured. Standard quantum mechanics certainly does not include any fundamental element intended to represent the value of $\vec{v} \cdot \vec{\sigma}$, for all unit vectors $\vec{v}$.

The goal of EPR was to show that quantum mechanics is incomplete, by demonstrating that quantum mechanics lacked some essential 'element of reality', by their criterion. They hoped to force a return to a more classical view of the world, one in which systems could be ascribed properties which existed independently of measurements performed on those systems. Unfortunately for EPR, most physicists did not accept the above reasoning as convincing. The attempt to impose on Nature *by fiat* properties which she must obey seems a most peculiar way of studying her laws.

Indeed, Nature has had the last laugh on EPR. Nearly thirty years after the EPR paper was published, an *experimental test* was proposed that could be used to check whether or not the picture of the world which EPR were hoping to force a return to is valid or not. It turns out that Nature *experimentally invalidates* that point of view, while agreeing with quantum mechanics.

The key to this experimental invalidation is a result known as *Bell's inequality*. Bell's inequality is *not* a result about quantum mechanics, so the first thing we need to do is momentarily *forget* all our knowledge of quantum mechanics. To obtain Bell's inequality, we're going to do a thought experiment, which we will analyze using our common sense notions of how the world works – the sort of notions Einstein and his collaborators thought Nature ought to obey. After we have done the common sense analysis, we will perform a quantum mechanical analysis which we can show *is not consistent with the common sense analysis*. Nature can then be asked, by means of a real experiment, to decide between our common sense notions of how the world works, and quantum mechanics.

Imagine we perform the following experiment, illustrated in Figure 2.4. Charlie prepares two particles. It doesn't matter how he prepares the particles, just that he is capable of repeating the experimental procedure which he uses. Once he has performed the preparation, he sends one particle to Alice, and the second particle to Bob.

Once Alice receives her particle, she performs a measurement on it. Imagine that she has available two different measurement apparatuses, so she could choose to do one of two different measurements. These measurements are of physical properties which we shall label P_Q and P_R, respectively. Alice doesn't know in advance which measurement she will choose to perform. Rather, when she receives the particle she flips a coin or uses some other random method to decide which measurement to perform. We suppose for simplicity that the measurements can each have one of two outcomes, $+1$ or -1. Suppose Alice's particle has a value Q for the property P_Q. Q is assumed to be an *objective property* of Alice's particle, which is merely revealed by the measurement, much as we imagine the position of a tennis ball to be revealed by the particles of light being scattered off it. Similarly, let R denote the value revealed by a measurement of the property P_R.

Similarly, suppose that Bob is capable of measuring one of two properties, P_S or P_T, once again revealing an objectively existing value S or T for the property, each taking value +1 or −1. Bob does not decide beforehand which property he will measure, but waits until he has received the particle and then chooses randomly. The timing of the experiment is arranged so that Alice and Bob do their measurements *at the same time* (or, to use the more precise language of relativity, in a causally disconnected manner). Therefore, the measurement which Alice performs cannot disturb the result of Bob's measurement (or vice versa), since physical influences cannot propagate faster than light.

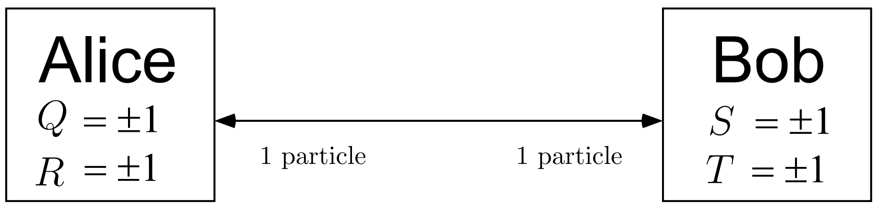

Figure 2.4. Schematic experimental setup for the Bell inequalities. Alice can choose to measure either Q or R, and Bob chooses to measure either S or T. They perform their measurements simultaneously. Alice and Bob are assumed to be far enough apart that performing a measurement on one system can not have any effect on the result of measurements on the other.

We are going to do some simple algebra with the quantity $QS + RS + RT - QT$. Notice that

$$QS + RS + RT - QT = (Q + R)S + (R - Q)T. \tag{2.219}$$

Because $R, Q = \pm 1$ it follows that either $(Q + R)S = 0$ or $(R - Q)T = 0$. In either case, it is easy to see from (2.219) that $QS + RS + RT - QT = \pm 2$. Suppose next that $p(q, r, s, t)$ is the probability that, before the measurements are performed, the system is in a state where $Q = q, R = r, S = s$, and $T = t$. These probabilities may depend on how Charlie performs his preparation, and on experimental noise. Letting $\mathbf{E}(\cdot)$ denote the mean value of a quantity, we have

$$\begin{aligned} \mathbf{E}(QS + RS + RT - QT) &= \sum_{qrst} p(q, r, s, t)(qs + rs + rt - qt) && (2.220) \\ &\leq \sum_{qrst} p(q, r, s, t) \times 2 && (2.221) \\ &= 2. && (2.222) \end{aligned}$$

Also,

$$\begin{aligned} \mathbf{E}(QS + RS + RT - QT) &= \sum_{qrst} p(q, r, s, t)qs + \sum_{qrst} p(q, r, s, t)rs \\ &\quad + \sum_{qrst} p(q, r, s, t)rt - \sum_{qrst} p(q, r, s, t)qt && (2.223) \\ &= \mathbf{E}(QS) + \mathbf{E}(RS) + \mathbf{E}(RT) - \mathbf{E}(QT). && (2.224) \end{aligned}$$

Comparing (2.222) and (2.224) we obtain the *Bell inequality*,

$$\mathbf{E}(QS) + \mathbf{E}(RS) + \mathbf{E}(RT) - \mathbf{E}(QT) \leq 2. \tag{2.225}$$

This result is also often known as the *CHSH inequality* after the initials of its four discoverers. It is part of a larger set of inequalities known generically as Bell inequalities, since the first was found by John Bell.

By repeating the experiment many times, Alice and Bob can determine each quantity on the left hand side of the Bell inequality. For example, after finishing a set of experiments, Alice and Bob get together to analyze their data. They look at all the experiments where Alice measured P_Q and Bob measured P_S. By multiplying the results of their experiments together, they get a sample of values for QS. By averaging over this sample, they can estimate $\mathbf{E}(QS)$ to an accuracy only limited by the number of experiments which they perform. Similarly, they can estimate all the other quantities on the left hand side of the Bell inequality, and thus check to see whether it is obeyed in a real experiment.

It's time to put some quantum mechanics back in the picture. Imagine we perform the following quantum mechanical experiment. Charlie prepares a quantum system of two qubits in the state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{2.226}$$

He passes the first qubit to Alice, and the second qubit to Bob. They perform measurements of the following observables:

$$Q = Z_1 \qquad S = \frac{-Z_2 - X_2}{\sqrt{2}} \tag{2.227}$$

$$R = X_1 \qquad T = \frac{Z_2 - X_2}{\sqrt{2}}. \tag{2.228}$$

Simple calculations show that the average values for these observables, written in the quantum mechanical $\langle \cdot \rangle$ notation, are:

$$\langle QS \rangle = \frac{1}{\sqrt{2}};\ \langle RS \rangle = \frac{1}{\sqrt{2}};\ \langle RT \rangle = \frac{1}{\sqrt{2}};\ \langle QT \rangle = -\frac{1}{\sqrt{2}}. \tag{2.229}$$

Thus,

$$\langle QS \rangle + \langle RS \rangle + \langle RT \rangle - \langle QT \rangle = 2\sqrt{2}. \tag{2.230}$$

Hold on! We learned back in (2.225) that the average value of QS plus the average value of RS plus the average value of RT minus the average value of QT can never exceed two. Yet here, quantum mechanics predicts that this sum of averages yields $2\sqrt{2}$!

Fortunately, we can ask Nature to resolve the apparent paradox for us. Clever experiments using photons – particles of light – have been done to check the prediction (2.230) of quantum mechanics versus the Bell inequality (2.225) which we were led to by our common sense reasoning. The details of the experiments are outside the scope of the book, but the results were resoundingly in favor of the quantum mechanical prediction. The Bell inequality (2.225) is *not* obeyed by Nature.

What does this mean? It means that one or more of the assumptions that went into the derivation of the Bell inequality must be incorrect. Vast tomes have been written analyzing the various forms in which this type of argument can be made, and analyzing the subtly different assumptions which must be made to reach Bell-like inequalities. Here we merely summarize the main points.

There are two assumptions made in the proof of (2.225) which are questionable:

(1) The assumption that the physical properties P_Q, P_R, P_S, P_T have definite values Q, R, S, T which exist independent of observation. This is sometimes known as the assumption of *realism*.
(2) The assumption that Alice performing her measurement does not influence the result of Bob's measurement. This is sometimes known as the assumption of *locality*.

These two assumptions together are known as the assumptions of *local realism*. They are certainly intuitively plausible assumptions about how the world works, and they fit our everyday experience. Yet the Bell inequalities show that at least one of these assumptions is not correct.

What can we learn from Bell's inequality? For physicists, the most important lesson is that their deeply held commonsense intuitions about how the world works are wrong. The world is *not* locally realistic. Most physicists take the point of view that it is the assumption of realism which needs to be dropped from our worldview in quantum mechanics, although others have argued that the assumption of locality should be dropped instead. Regardless, Bell's inequality together with substantial experimental evidence now points to the conclusion that either or both of locality and realism must be dropped from our view of the world if we are to develop a good intuitive understanding of quantum mechanics.

What lessons can the fields of quantum computation and quantum information learn from Bell's inequality? Historically the most useful lesson has perhaps also been the most vague: there is something profoundly 'up' with entangled states like the EPR state. A lot of mileage in quantum computation and, especially, quantum information, has come from asking the simple question: 'what would some entanglement buy me in this problem?' As we saw in teleportation and superdense coding, and as we will see repeatedly later in the book, by throwing some entanglement into a problem we open up a new world of possibilities unimaginable with classical information. The bigger picture is that Bell's inequality teaches us that entanglement is a fundamentally new resource in the world that goes essentially *beyond* classical resources; iron to the classical world's bronze age. A major task of quantum computation and quantum information is to exploit this new resource to do information processing tasks impossible or much more difficult with classical resources.

Problem 2.1: (Functions of the Pauli matrices) Let $f(\cdot)$ be any function from complex numbers to complex numbers. Let $\vec{n}$ be a normalized vector in three dimensions, and let θ be real. Show that

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma}. \tag{2.231}$$

Problem 2.2: (Properties of the Schmidt number) Suppose $|\psi\rangle$ is a pure state of a composite system with components A and B.

(1) Prove that the Schmidt number of $|\psi\rangle$ is equal to the rank of the reduced density matrix $\rho_A \equiv \mathrm{tr}_B(|\psi\rangle\langle\psi|)$. (Note that the rank of a Hermitian operator is equal to the dimension of its support.)
(2) Suppose $|\psi\rangle = \sum_j |\alpha_j\rangle|\beta_j\rangle$ is a representation for $|\psi\rangle$, where $|\alpha_j\rangle$ and $|\beta_j\rangle$ are (un-normalized) states for systems A and B, respectively. Prove that the