Eigenvalue Inequalities in Quantum Information Processing

Thesis by

Sumit Kumar Daftuar

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2004
(Defended September 25, 2003)

Acknowledgements

Many people have helped me in this endeavor.

My advisor, John Preskill, was gracious enough to accept me as a mathematics student into his research group. His wonderful course on quantum information theory introduced me to the subject. He also provided guidance on problems to consider answering, and where to look for help in solving them, on various occasions.

Michael Nielsen introduced me to the subject of majorization and its applications to quantum information theory. He suggested a lot of useful questions which got me started on the problem considered in Part I of this thesis. He also provided helpful encouragement and feedback on some of the work in Part I. In addition, he taught me some representation theory.

I collaborated with Matthew Klimesh on some of the work presented in Part I of this thesis (essentially, the last three sections of Chapter 2).

Patrick Hayden was my collaborator on Part II of this thesis. Perhaps "mentor" would be a better word to describe his role. He introduced me to the problem and recognized how to generalize my initial line of attack; from that point, he guided our joint efforts. Along the way, he explained many difficult concepts to me. In the context of all this, it seems hardly worth mentioning that he also provided extensive comments on a draft of this thesis and drew one of the figures for me. I cannot thank him enough.

Michael Hartl helped me learn LaTeX, and more recently, helped me with thesis-specific LaTeXissues. Long before that, he was my study partner in virtually every undergraduate physics course I took (and some math courses too), and has undoubtedly influenced my thinking in ways I don't even realize.

I also wish to thank Charlene Ahn, Michael Aschbacher, David Bacon, David Beckman, John Cortese, Christopher Fuchs, Jim Harrington, Rowan Killip, Allen Knutson, Andrew Landahl, Debbie Leung, Carlos Mochon, Benjamin Rahn, Eric Rains, Guifré Vidal, Clint White, and Richard Wilson, whose helpful discussions (in some cases, courses) enhanced my understanding of physics and/or mathematics during my time at Caltech.

Abstract

This thesis develops restrictions governing how a quantum system, jointly held by two parties, can be altered by the local actions of those parties, under assumptions about how they may communicate. These restrictions are expressed as constraints involving the eigenvalues of the density matrix of one of the parties. The thesis is divided into two parts.

Part I (Chapters 1–4) explores what is possible if the two parties may use only classical communication. A well-known result by M. Nielsen says that this is intimately connected to the *majorization* relation: if x is the vector of eigenvalues of the initial state, then y can be the vector of eigenvalues of the final state if and only if x is majorized by y. It was recently observed that it is possible for $x \otimes z$ to be majorized by $y \otimes z$, even if x is not majorized by y; physically, this means that the presence of a state with eigenvalues z is a *catalyst* that allows a certain transformation to occur. If such a z exists, then x is said to be *trumped* by y. Part I is mainly a study of the structure of this trumping relation, an extension of the majorization relation. Notably, we show that for almost all probability vectors $y \in \mathbb{R}^d$ where $d \geq 4$, there is no finite dimension n such that the set of vectors trumped by y can be determined by restricting attention to catalysts of dimension n. We also study some concrete examples to illustrate various aspects of the trumping relation.

Part II (Chapters 5–9) considers the question of how a state can change as a result of quantum communication between the parties; i.e., one party sends the other a portion of the jointly held quantum system. Given the spectrum of the initial state, it turns out that the possible spectra of the final state are given by the solutions to linear inequalities. We develop a method for deriving these inequalities, using a variational principle. In order to apply this principle, we need to know when certain subvarieties of a Grassmannian variety intersect, which can be a regarded as a problem in Grassmannian cohomology. We discuss this cohomology and derive the conditions for nontrivial intersection. Finally, we illustrate how these intersections give rise to the desired inequalities.

Contents

List of Figures

Part I

Mathematical Structure of Entanglement Catalysis

Chapter 1

Majorization

We begin by introducing the theory of *majorization*, a mathematical relation that has recently been shown to have striking applications to quantum information theory. Majorization constraints have been shown to govern transformations of quantum entanglement [1], to restrict the spectra of separable quantum states [2], and to characterize how quantum states change as a result of mixing or measurement [3]. It has even been suggested that all efficient quantum algorithms must respect a majorization principle [4]. Our purposes will be to introduce some background facts that will be useful to us, and to demonstrate various ways of characterizing the majorization condition. Because our main goal for Part I will be to study an extension of the majorization relation (known as *trumping*), such characterizations will serve as an illustration of the types of results we seek for the trumping relation. This chapter consists of background material that can be found in a reference such as [5] or [6].

1.1 Definition and Motivation

Let $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$. We will be most interested in the case where x and y are are d-dimensional probability vectors; in other words, their components are nonnegative and sum to unity. However, for most results in the theory of majorization, this restriction is not needed. Let $x^\downarrow$ denote the d-dimensional vector obtained by arranging the components of x in non-increasing order: $x^\downarrow = (x_1^\downarrow, \ldots, x_d^\downarrow)$, where $x_1^\downarrow \geq x_2^\downarrow \geq \cdots \geq x_d^\downarrow$. Then we say that x is *majorized* by y, written $x \prec y$, if the following relations hold:

$$\sum_{i=1}^{\ell} x_i^{\downarrow} \leq \sum_{i=1}^{\ell} y_i^{\downarrow} \qquad (1 \leq \ell < d) \tag{1.1}$$

and

$$\sum_{i=1}^{d} x_i^{\downarrow} = \sum_{i=1}^{d} y_i^{\downarrow}. \tag{1.2}$$

Intuitively, if x and y are probability vectors such that $x \prec y$, then x describes an unambiguously more random distribution than does y. For example, in $\mathbb{R}^2$, we have that $(0.5, 0.5) \prec (0.8, 0.2)$. In fact, $(0.5, 0.5)$ is majorized by every vector in $\mathbb{R}^2$ whose components sum to unity.

The majorization relation defines a partial order on d-dimensional real vectors, where $x \prec y$ and $y \prec x$ if and only if $x^{\downarrow} = y^{\downarrow}$. To see that majorization is not a complete relation, consider for instance $x = (0.5, 0.25, 0.25)$ and $y = (0.4, 0.4, 0.2)$; then $x \not\prec y$ and $y \not\prec x$.

Majorization was introduced to formalize the notion of what it means for one vector to be unambiguously less disordered (or alternatively, more unequal) than another. Some of the beginnings of the theory originate from economics, where it played a role in comparing income and wealth distributions. We will illustrate the meaning of majorization in terms of this idea, to motivate the definition given by Inequalities 1.1 and Equation 1.2. Consider two populations X and Y, each of d individuals. Let x_i be the wealth of individual i in population X, and let y_i be the wealth of individual i in population Y. Suppose for simplicity that the total amount of wealth in the two populations is the same, $\sum_i x_i = \sum_i y_i$ (we can divide each term x_i and y_i by $\sum_i x_i$ and $\sum_i y_i$, respectively, to normalize for differences in total wealth). Now, suppose that the richest individual in population Y has at least as much wealth as the richest individual in population X, the two richest individuals in population Y have at least as much combined wealth as the two richest individuals in population X, etc. (Note that because the total amount of wealth is equal in the two populations, this is equivalent to saying that the poorest individual in population

X has at least as much wealth as the poorest individual in population Y, the two poorest individuals in X have at least as much combined wealth as the two poorest individuals in Y, etc.) Then it is reasonable to say that $(x_1, \ldots, x_n)$ represents a more equal distribution of wealth than $(y_1, \ldots, y_n)$. This notion of inequality was introduced by M. O. Lorenz [7] in 1905. In our notation, this is saying precisely that

$$\sum_{i=1}^{\ell} x_i^{\downarrow} \leq \sum_{i=1}^{\ell} y_i^{\downarrow}, \tag{1.3}$$

i.e., that x is majorized by y.

Another way of evaluating wealth inequality is by considering the effects of transfers of wealth. Let i and j be two individuals in a population X, where without loss of generality we assume that $x_i \leq x_j$. A transfer of wealth is said to take place if j (the wealthier member) gives some wealth to i, but not so much that i is now wealthier than j used to be. Mathematically, (x_i, x_j) gets mapped to the convex combinations $(tx_i + (1-t)x_j, (1-t)x_i + tx_j)$, for some $t \in [0, 1]$. The effect of a transfer is to make the overall wealth distribution more equal; this suggests that we define one wealth distribution to be more equal than another, if it can be obtained from the other by a series of wealth transfers. This notion of inequality was suggested by E. C. Pigou [8] and H. Dalton [9] in the early 20th century. It turns out that these two notions of inequality are equivalent, a fact which we will prove in the next section.

1.2 T-transforms

Define a linear map T from $\mathbb{R}^d$ to $\mathbb{R}^d$ to be a *T-transform* if there exist $t \in [0, 1]$ and indices j, k such that

$$T(y) = (y_1, \ldots, y_{j-1}, ty_j + (1-t)y_k, y_{j+1}, \ldots, (1-t)y_j + ty_k, y_{k+1}, \ldots, y_d).$$

Then we have the following theorem:

Theorem 1.2.1 *Let x and y be vectors in $\mathbb{R}^d$. Then $x \prec y$ if and only if x can be* *obtained from y by a finite number of T-transforms.*

Proof It is easy to see that $T(y) \prec y$ for any T-transform T, so if $D = T_1 \dots T_r$ is a product of T-transforms, then $x = D(y) \prec y$. This proves one direction.

For the other direction, we will use induction on d, the dimension of the vector space of which x and y are elements. Clearly, the result holds for the base case $d = 2$.

Suppose the statement is true for a given dimension d, and that $x \prec y$ for vectors $x, y \in \mathbb{R}^{d+1}$. We may assume without loss of generality that $x = x^{\downarrow}$ and $y = y^{\downarrow}$. Since $x \prec y$, $y_{d+1} \leq x_{d+1} \leq x_1 \leq y_1$, so there must be a $k \in \{1, \dots, d+1\}$ such that $y_k \leq x_1 \leq y_{k-1}$. So there exists $t \in [0,1]$ such that $x_1 = ty_1 + (1-t)y_k$. Let T be the T-transform that maps y_1 to $ty_1 + (1-t)y_k$, and maps y_k to $ty_k + (1-t)y_1$:

$$
\begin{aligned}
Ty &= (ty_1 + (1-t)y_k, y_2, \dots, y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, y_{d+1}) && (1.4)\\
&= (x_1, y'), && (1.5)
\end{aligned}
$$

where

$$
y' = (y_2, \dots, y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, y_{d+1}). \tag{1.6}
$$

Define $x' = (x_2, x_3, \dots x_{d+1})$. Now x' and y' are d-dimesional vectors, so we will show that $x' \prec y'$ in order to apply the inductive hypothesis. Suppose first that $1 \leq \ell \leq k-2$. Then since $y_{k-1} \geq x_1$, we have that

$$
\begin{aligned}
\sum_{j=1}^{\ell} x'_j &= \sum_{j=2}^{\ell+1} x_j && (1.7)\\
&\leq \sum_{j=2}^{\ell+1} y_j && (1.8)\\
&= \sum_{j=1}^{\ell} y'_j && (1.9)\\
&\leq \sum_{j=1}^{\ell} (y'_j)^{\downarrow}. && (1.10)
\end{aligned}
$$

Next suppose that $k-1 \leq \ell \leq d$. Then we have

$$\begin{aligned}
\sum_{j=1}^{\ell} (y'_j)^{\downarrow} &\geq \sum_{j=1}^{\ell} y'_j & (1.11)\\
&= \sum_{j=2}^{k-1} y_j + [(1-t)y_1 + ty_k] + \sum_{j=k+1}^{\ell+1} y_j & (1.12)\\
&= \sum_{j=1}^{\ell+1} y_j - [ty_1 + (1-t)y_k] & (1.13)\\
&= \sum_{j=1}^{\ell+1} y_j - x_1 & (1.14)\\
&\geq \sum_{j=1}^{\ell+1} x_j - x_1 & (1.15)\\
&= \sum_{j=2}^{\ell+1} x_j & (1.16)\\
&= \sum_{j=1}^{\ell} x'_j. & (1.17)
\end{aligned}$$

We have thus shown that $x' \prec y'$. Therefore, there is a sequence $T_1, \ldots T_r$ of T-transforms on $\mathbb{R}^d$ such that $x' = T_1 \cdots T_r y'$. But we may regard each T_i as a transformation on $\mathbb{R}^{d+1}$ that fixes the first coordinate, so we have that $x = T_1 \ldots T_r T y$. $\square$

Corollary 1.2.2 *The two notions of wealth inequality given in the previous section are equivalent.*

1.3 Geometric Characterization

Recall that a $d \times d$ matrix A is said to be *doubly stochastic* if all of its entries are nonnegative, and each row and column of A sums to unity. For instance, is not hard to see that every T-transformation is a doubly stochastic map, and that products of doubly stochastic maps are doubly stochastic. The study of doubly stochastic matrices is well-known to be connected to the theory of majorization [10, 11]:

Theorem 1.3.1 *(a) A $d \times d$ real matrix A is doubly stochastic if and only if $Ay \prec y$ for all $y \in \mathbb{R}^d$.*

(b) $x \prec y$ if and only if there is a doubly stochastic matrix A such that $x = Ay$.

If we think of x and y as probability vectors, then Theorem 1.3.1 (a) tells us that the doubly stochastic matrices are precisely those matrices that map any probability distribution to one that is at least as mixed.

Given a vector $y \in \mathbb{R}^d$, define $S(y)$ to be the set of vectors $x \in \mathbb{R}^d$ such that $x \prec y$. By Theorem 1.3.1, $S(y) = \{Ay | A \text{ is doubly stochastic}\}$. In this section we will establish *Birkhoff's theorem*, which gives a geometric description of the doubly stochastic matrices, and use it to give a geometric description of $S(y)$.

We begin with the *marriage problem* from combinatorics [12]. Let B and G be two finite sets of the same cardinality, and let R be a relation on $B \times G$. We think of the elements of B and G as "boys" and "girls," respectively, and $R(b, g)$ as the relation that $b \in B$ and $g \in G$ love one another. A *compatible matching* is a pairing of each boy with one girl (distinct for each boy) such that only couples who love one another are paired up. The marriage problem is to determine when a compatible matching exists, given $B \times G$ and R. The solution is given by Hall's theorem:

Theorem 1.3.2 (Hall's Theorem) *A compatible matching for $B \times G$ and R exists if and only if every group of k boys loves at least k girls, for $k \in \{1, \ldots, |B|\}$.*

Proof Clearly, if a compatible matching exists, each group of k boys loves at least k girls (those girls chosen to be their matches).

For the reverse direction, we proceed by induction. The base case $|B| = 1$ is clear, so assume the statement is true when $|B| \leq n$; we wish to prove it for $|B| = n + 1$.

Suppose first that there exists $k \in \{1, \ldots, n\}$ such that there is a group β of k boys who love a group γ of exactly k girls. Then β and γ can be compatibly matched, by the inductive hypothesis. The complements β^c and γ^c can also be compatibly matched: if S is a subset of β^c containing h members, then by assumption, the set $\beta \cup S$ of $k + h$ boys loves at least $k + h$ girls, so that the h boys of S must love at least h girls in γ^c. This implies that β^c and γ^c can be compatibly matched, by the inductive hypothesis.

Now suppose that the assumption of the previous paragraph is false, meaning that for each $k \leq n$, all groups of k boys love at least $k+1$ girls. In this case we can simply take one boy and girl who love each other, and pair them together. The remaining n boys and n girls now satisfy the inductive hypothesis. □

Hall's theorem is equivalent to the following theorem on matrices. Given a $d \times d$ matrix A, define a *diagonal* of A to be a set $\{a_{1\pi(1)}, a_{2\pi(2)}, \ldots, a_{d\pi(d)}\}$, where π is a permutation of $\{1, \ldots, d\}$.

Corollary 1.3.3 (König-Frobenius Theorem) *A $d \times d$ matrix A contains a diagonal with no zero elements if and only if every $k \times l$ zero submatrix of A satisfies $k + l \leq d$.*

Proof We construct a marriage problem from the matrix A. The boys correspond to the rows of A, and the girls correspond to the columns; boy i and girl j love one another if and only if $A_{ij} \neq 0$. Then a compatible matching occurs if and only if A has a nonzero diagonal. By Hall's theorem, this happens if and only if each group of k boys loves at least k girls; i.e., for every $k \times l$ zero submatrix, $k \leq d - l$. □

We are now ready to prove Birkhoff's theorem.

Theorem 1.3.4 (Birkhoff's Theorem) *The set of $d \times d$ doubly stochastic matrices is a convex set whose extreme points are the $d \times d$ permutation matrices.*

Proof It is straightforward to check that the set of $d \times d$ doubly stochastic matrices is convex, and that the permutation matrices are extreme points of this set. So we must show that any doubly stochastic stochastic matrix D can be written as a convex sum of permutation matrices:

$$D = \sum_i p_i P_i. \tag{1.18}$$

Let $n(D)$ be the number of nonzero matrix elements of D. Because each row must have at least one nonzero entry, $n(D) \geq d$. We use induction on $n(D)$. For the base case $n(D) = d$, D has only one nonzero entry in each row and in each column, this nonzero entry must therefore be 1. It follows that D itself is a permutation matrix, so the statement is true for the base case.

For the inductive step, first note that the sum of all the elements of D must be equal to d. If D has a $k \times l$ submatrix, then the sum of the elements of the k rows corresponding to this submatrix, plus the sum of the elements of the l columns correspond to the submatrix, must be less than the sum of all elements of D, since no nonzero element is included more than once in the sum. Therefore, $k + l \leq d$. So we may apply the König-Frobenius theorem to conclude that there must be a diagonal of D with only nonzero elements. Choose any such diagonal, and let p be the smallest element on this diagonal, and P be the permutation matrix whose ones are on this diagonal. If $p = 1$, then D must be a permutation matrix, so we are done. Consider the case $0 < p < 1$. Let Q be the matrix defined by

$$Q = \frac{D - pP}{1 - p}. \tag{1.19}$$

Then Q is doubly stochastic and has fewer nonzero entries than D, so by the inductive hypothesis, we may write Q as a convex sum of permutation matrices:

$$Q = \sum_i p_i P_i. \tag{1.20}$$

But $D = (1 - p)Q + pP$, so

$$D = pP + \sum_i (1 - p) p_i P_i \tag{1.21}$$

is a convex sum of permutation matrices. □

Birkhoff's Theorem and Theorem 1.3.1 together imply the following:

Theorem 1.3.5 *For any $y \in \mathbb{R}^d$, $S(y)$ is a convex set whose extreme points are the elements of the set $\{Py | P$ is a $d \times d$ permutation matrix$\}$.*

1.4 Schur-convexity

Much of the power of majorization comes from the theory of *Schur-convexity*, which allows one to derive inequalities from an appropriate majorization condition. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *Schur-convex* if $f(x) \le f(y)$ whenever $x \prec y$. If $f(x) \ge f(y)$ whenever $x \prec y$, then f is said to be *Schur-concave*. While it is not obvious that interesting Schur-convex (or Schur-concave) functions should exist at all, the following theorem shows how to construct many such functions:

Theorem 1.4.1 *If $I \subset \mathbb{R}$ is an interval and $g : I \to \mathbb{R}$ is convex (concave), then the function*

$$\phi(x) = \sum_{i=1}^{n} g(x_i)$$

is Schur-convex (Schur-concave) on I^n.

Proof In view of Theorem 1.2.1, it is sufficient to show that $\phi(x) \le \phi(y)$ whenever $x = Ty$ for some T-transform T. Without loss of generality, suppose T acts non-trivially on the first two components of y, so that $x_1 = ty_1 + (1-t)y_2$, $x_2 = (1-t)y_1 + ty_2$, and $x_i = y_i$ for $i > 2$. Then $g(x_1) + g(x_2) = g(ty_1 + (1-t)y_2) + g((1-t)y_1 + ty_2) \le tg(y_1) + (1-t)g(y_2) + (1-t)g(y_1) + tg(y_2) = g(y_1) + g(y_2)$, so $\phi(x) \le \phi(y)$. □

One consequence of Theorem 1.4.1 is the connection between majorization and entropy, a more familiar measure of randomness. Because the function $g(p) = -p \log p$ is concave on the interval $[0,1]$, it follows that the entropy function $H(x) = -\sum_i x_i \log x_i$ is a Schur-concave function. That is, if $x \prec y$ (where x and y are probability vectors) then $H(x) \ge H(y)$. This agrees with our intuition that $x \prec y$ means that x describes a more random probability distribution than y does. Of course, majorization is a much stronger condition than the entropy criterion for determining relative randomness: there exist probability vectors x and y such that $x \not\prec y$, yet $H(x) \ge H(y)$. This is not hard to understand, when we consider that majorization is not a complete relation.

The notion of Schur-convexity has been used to derive inequalities in many branches of mathematics, notably linear algebra, geometry, and statistics. For example, it can be shown that the diagonal entries of a Hermitian matrix are majorized by its eigenvalues (this is an easy consequence of Ky Fan's Maximum Principle; see Theorem 6.1.1). Schur himself used this fact to give a proof of Hadamard's well-known determinant inequality:

Theorem 1.4.2 (Hadamard Determinant Inequality) *Let H be a positive definite Hermitian matrix. Then the determinant of H is less than or equal to the product of the diagonal entries.*

Proof Let $h = (h_{11}, h_{22}, \ldots, h_{nn})$ be the vector of diagonal entries of H, and let $\lambda(H) = (\lambda_1(H), \ldots, \lambda_n(H))$ be the vector of eigenvalues of H. Because the function $g(t) = \log t$ is concave, the function $\phi(x) = \sum_{i=1}^{d} \log t$ is Schur-concave. Since $h \prec \lambda(H)$, it follows that $\sum_{i=1}^{d} \log \lambda_i(H) \leq \sum_{i=1}^{d} \log h_{ii}$. This implies that the product of the eigenvalues is less than or equal to the product of the diagonal entries. □

The majorization relation itself can be defined in terms of Schur-convex functions. It is not hard to prove the following directly:

Theorem 1.4.3 *Let x, $y \in \mathbb{R}^d$. Then $x \prec y$ if and only if for all $t \in \mathbb{R}$,*

$$\sum_{i=1}^{d} |x_i - t| \leq \sum_{i=1}^{d} |y_i - t|. \tag{1.22}$$

Theorem 1.4.3 has limited use because it is easier to check the defining inequalities for majorization than to check that Inequalities 1.22 are satisfied. However, it has theoretical value because it shows that Schur-convex functions can be used to characterize majorization:

Theorem 1.4.4 *Let $x, y \in \mathbb{R}^d$. Then $x \prec y$ if and only if $f(x) \leq f(y)$ for all Schur-convex functions $f : \mathbb{R}^d \to \mathbb{R}$.*

Proof The function $g_t(s) = |s - t|$ is convex, so that for any $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$, $\phi_t(x) = \sum_{i=1}^{d} |x_i - t|$ is Schur-convex. So if $f(x) \leq f(y)$ for all Schur-convex functions

f, then in particular $\phi_t(x) \leq \phi_t(y)$ for all $t \in \mathbb{R}$, so $x \prec y$ by Theorem 1.4.3. The reverse direction follows from the definition of Schur-convex function. $\square$

1.5 Summary

We collect some useful properties of the majorization relation into the following list:

- Given two vectors x and y, it is easy to determine whether $x \prec y$ (the definition can be checked directly, for example).
- We can intepret $x \prec y$ as saying that x can be obtained from y via a series of simple mixing operations (transfers).
- The geometric structure of majorization is well-behaved; $x \prec y$ means that x lies in the convex hull of the vectors obtained by permuting the components of y.
- Majorization can also be characterized function-theoretically, in that there is a family of functions ϕ_t such that $\phi_t(x) \leq \phi_t(y)$ for all t is necessary and sufficient for $x \prec y$.

We will keep this list in mind in trying to analyze the related notion of *trumping*, defined in the next chapter.

Chapter 2

Introduction to Trumping

In this chapter, we introduce an extension of the majorization relation that will be the main focus of our study in Part I. Given probability vectors x and y, we ask when there exists a probabability vector z such that $x \otimes z \prec y \otimes z$. (It turns out that this situation may occur even if $x \not\prec y$.) This question arises naturally in studying what transformations of quantum entanglement are possible using only local operations and classical communcation. The mathematical notion may be accurately described as "tensor product induced majorization" but we will use the simpler term *trumping*, introduced by M. Nielsen [6]. The material in this chapter, and in the first section of the next chapter, was published previously by the author and a collaborator [13].

2.1 Entaglement Catalysis

Quantum entanglement exists when a quantum mechanical system, consisting of various subsystems, cannot be fully described simply by giving a complete local description of all the subsystems. Entanglement seems to play an essential role in numerous remarkable applications of quantum information science, including quantum cryptography [14, 15], quantum teleportation [16], and superdense coding [17]; because of this, it has come to be viewed as a fundamental resource that allows one to perform certain information-processing tasks. As with any physical resource, one wishes to measure how much entanglement is present in a given system, and to determine under what conditions it is possible to convert one form of entanglement to another. The problem of how to quantify and classify entanglement is one of the basic questions in the study of quantum information [18, 19].

The following theorem due to M. Nielsen shows that the structure of bipartite quantum entanglement is intimately related to majorization [1]:

Theorem 2.1.1 *Suppose Alice and Bob are in joint possession of a bipartite entangled quantum state* $|\psi\rangle$ *which they wish to transform into another bipartite entangled state* $|\phi\rangle$ *using only local operations and classical communication (LOCC). Let* $|\psi\rangle = \sum_{i=1}^{d} \sqrt{\alpha_i}|i_A\rangle|i_B\rangle$ *be a Schmidt decomposition of* $|\psi\rangle$*, and let* $|\phi\rangle = \sum_{i=1}^{d} \sqrt{\beta_i}|i'_A\rangle|i'_B\rangle$ *be a Schmidt decomposition of* $|\phi\rangle$ *. Then* $|\psi\rangle$ *can be converted to* $|\phi\rangle$ *by LOCC if and only if the vector* $\alpha = (\alpha_1, \ldots, \alpha_d)$ *is majorized by* $\beta = (\beta_1, \ldots, \beta_d)$*.*

Nielsen's theorem defines a partial order on the entangled bipartite pure states. If state $|\psi\rangle$ has x as its vector of Schmidt coefficients, and $|\phi\rangle$ has y as its vector of Schmidt coefficients, then we can transform $|\psi\rangle$ to $|\phi\rangle$ using LOCC if and only if $x \prec y$. Because our ability to transform one state to another depends only on their Schmidt coefficients, and not on the bases, we shall abuse nomenclature and refer to any vector of Schmidt coefficients as a "state."

Jonathan and Plenio have extended Nielsen's result by describing a phenomenon known as entanglement catalysis [20]. Suppose that $x = (0.4, 0.4, 0.1, 0.1)$ and $y = (0.5, 0.25, 0.25, 0)$. Then $x \not\prec y$. Now let $z = (0.6, 0.4)$. Then we have $x \otimes z \prec y \otimes z$. In other words, if Alice and Bob start only with state x (by which we mean a jointly entangled quantum state whose Schmidt coefficients are the components of x), they cannot transform it into state y using LOCC. But if they also have state z available, then they can turn $x \otimes z$ into $y \otimes z$. So they can "borrow" z, use it to help turn x into y, and "return" it after performing the transformation. We say that z is a *catalyst* for the transformation.

The phenomenon of catalysis illustrates that entanglement itself can be used as a resource to help perform transformations of entangled states. One naturally wishes to know when this is possible: given x and y, can we determine whether x can be transformed to y using LOCC in the presence of a catalyst? This is equivalent to asking whether there is a probability vector z such that $x \otimes z \prec y \otimes z$. Transformations using LOCC together with a catalyst are termed *entanglement-assisted LOCC transformations*, abbreviated as *ELOCC transformations*.

2.2 Definitions and Basic Properties

We will adopt the terminology and notation introduced by Nielsen [6] and say that x is *trumped* by y, written $x \prec_T y$, if there exists a catalyst z (of any dimension) such that $x \otimes z \prec y \otimes z$. For any given y, let $T(y)$ denote the set of all x such that x is trumped by y; and for any y and z, let $T(y, z)$ be the set of all x such that $x \otimes z \prec y \otimes z$. In addition, we introduce the following notation: for any d-dimensional probability vector y and any positive integer k, let $T_k(y) = \{x \mid \exists$ a k-dimensional probability vector z such that $x \otimes z \prec y \otimes z\}$.

In contrast to the situation with the majorization relation, the mathematical structure of the trumping relation is not well understood. One desires a necessary and sufficient condition for determining whether $x \prec_T y$ (or alternately, to determine the elements of the set $T(y)$ for any given y). Characterizing the trumping relation in this way would help us to better understand the structure of the bipartite entangled states. However, such a characterization is not yet known. Part I of this thesis describes progress made in learning about the structure of this relation.

Our results will rely heavily on the fact that the trumping relation involves vectors with all nonnegative components. Note that this is quite different from the situation with majorization, in which most results extend easily to vectors containing negative components.

The following proposition lists some elementary facts about the trumping relation.

Proposition 2.2.1 *Let x and y be d-dimensional probability vectors, let z be a probability vector (of any dimension), and let $S(y)$, $T(y)$, and $T_k(y)$ be defined as above. Then*

(a) $x \prec y \Rightarrow x \otimes z \prec y \otimes z$.

(b) $S(y) \subseteq T(y)$.

(c) $T(y) = \bigcup_{k=1}^{\infty} T_k(y)$.

(d) If $x \prec_T y$, *then* $x_1^{\downarrow} \leq y_1^{\downarrow}$ *and* $x_d^{\downarrow} \geq y_d^{\downarrow}$.

(e) $T(y)$ *is a convex set.*

(f) If $x \prec_T y$ *and* $y \prec_T x$, *then* $x^{\downarrow} = y^{\downarrow}$.

Proof Parts (a)-(d) follow easily from the definitions. For (e) suppose that $x_1, x_2 \in T(y)$, and $t \in [0, 1]$. Then $\exists z_1, z_2$ such that $x_1 \otimes z_1 \prec y \otimes z_1$ and $x_2 \otimes z_2 \prec y \otimes z_2$. From part (a), it follows that $x_1 \otimes z_1 \otimes z_2 \prec y \otimes z_1 \otimes z_2$ and $x_2 \otimes z_1 \otimes z_2 \prec y \otimes z_1 \otimes z_2$. Therefore, by convexity of $S(y \otimes z_1 \otimes z_2)$, $tx_1 \otimes z_1 \otimes z_2 + (1-t)x_2 \otimes z_1 \otimes z_2 \prec y \otimes z_1 \otimes z_2$, so $tx_1 + (1-t)x_2 \in T(y)$. For (f), suppose that $\exists z_1, z_2$ such that $x \otimes z_1 \prec y \otimes z_1$ and $y \otimes z_2 \prec x \otimes z_2$. Then

$$x \otimes z_1 \otimes z_2 \prec y \otimes z_1 \otimes z_2 \prec x \otimes z_1 \otimes z_2, \tag{2.1}$$

so that $(x \otimes z_1 \otimes z_2)^{\downarrow} = (y \otimes z_1 \otimes z_2)^{\downarrow}$ and hence $x^{\downarrow} = y^{\downarrow}$. □

2.3 A Key Lemma

The following lemma and its corollary will be useful to us in proving additional results, and are also interesting in their own right:

Lemma 2.3.1 *Let* $x = (x_1, \ldots, x_d)$ *and* $y = (y_1, \ldots, y_d)$ *be* d*-dimensional probability vectors, whose components we assume to be arranged in non-increasing order:* $x_1 \geq x_2 \geq \cdots \geq x_d$, *and similarly for* y. *Suppose that* $x \prec y$, $y_1 > x_1$, *and* $y_d < x_d$. *Then* x *is in the interior of* $T(y)$.

Note that when we say x is in the interior of $T(y)$ we mean the interior relative to the space of d-dimensional probability vectors; that is, for any x there must exist an ϵ such that if x' is a probability vector for which $\|x' - x\| < \epsilon$ (in the Euclidean norm, for instance), then $x' \in T(y)$.

We remark that the conclusion is obvious if x is in the interior of $S(y)$; the important fact is that the result holds when x is on the boundary of $S(y)$.

Proof Note that $x_d > 0$. Pick an α satisfying $\alpha < 1$, $\alpha > x_1/y_1$, and $\alpha > y_d/x_d$. Let k be an integer for which $x_1\alpha^{k-1} < x_d$. Now let z be the k-dimensional vector

$$z = (1, \alpha, \ldots, \alpha^{k-1}).$$

(Of course z is not a probability vector, but it can easily be normalized. For convenience in the proof, we neglect the normalization.)

We will show that x is in the interior of $T(y, z)$. Since $T(y, z) \subset T(y)$, this will establish the result.

Let $(y \otimes z)_i^{\downarrow}$ denote the ith component of $y \otimes z$ when its components are arranged in non-increasing order. We will show that for $1 \leq \ell \leq dk - 1$,

$$\sum_{i=1}^{\ell} (x \otimes z)_i^{\downarrow} < \sum_{i=1}^{\ell} (y \otimes z)_i^{\downarrow}. \tag{2.2}$$

Note that since $x \otimes z$ must be majorized by $y \otimes z$, we already know that (2.2) must hold for $0 \leq \ell \leq dk$ if "$<$" is replaced by "$\leq$" (and this fact is used later in the proof). Showing that (2.2) holds for $1 \leq \ell \leq dk - 1$ will complete the proof since it is then clear that any sufficiently small perturbations to x (within the probability space) will not cause (2.2) to be violated for any $1 \leq \ell \leq dk - 1$.

For the remainder of the proof we fix ℓ as an arbitrary integer satisfying $1 \leq \ell \leq dk - 1$. Consider the terms that the left hand sum of (2.2) will contain. For $1 \leq i \leq d$, let r_i denote the number of these terms which are of the form $x_i\alpha^j$, with $0 \leq j < k$. (In case of repeated values of components of $x \otimes z$, we regard terms with smaller i to be included in the sum first.) Note that these r_i terms must be $x_i, x_i\alpha, \ldots, x_i\alpha^{r_i - 1}$, since these are the largest of this form. The sum (which we denote by s_x) can thus be written

$$s_x = \sum_{i=1}^{d} \sum_{j=0}^{r_i - 1} x_i \alpha^j \tag{2.3}$$

Note that $0 \le r_i \le k$ and in addition $r_1 > 0$ and $r_d < k$.

Consider the sum

$$s_y = \sum_{i=1}^{d} \sum_{j=0}^{r_i - 1} y_i \alpha^j. \tag{2.4}$$

The terms of this sum may or may not be the ℓ largest components of $y \otimes z$, but if $s_x < s_y$ then we are done because s_y is less than or equal to the right hand sum in (2.2). The fact that $x \prec y$ implies that $s_x \le s_y$; this follows from comparing the terms in the sums with a fixed j. Thus we need only consider the case $s_x = s_y$.

Let m_y be the minimum of the terms included in the sum in (2.4) and let M_y be the maximum of those components of $y \otimes z$ which are *not* included in this sum. Define m_x and M_x analagously. If $M_y > m_y$ then we are done, since the largest term not in the sum in (2.4) can be swapped with the smallest one in the sum, implying (2.2). We assume that $M_y \le m_y$ and show that a contradiction will follow.

There are two cases to consider. We first consider the case where $r_1 < k$ (that is, $r_1 \ne k$). Note that our current assumptions (including $M_y \le m_y$) imply $m_y \le m_x$, since otherwise we would have

$$\sum_{i=1}^{\ell-1} (x \otimes z)_i^{\downarrow} > \sum_{i=1}^{\ell-1} (y \otimes z)_i^{\downarrow}.$$

It follows that

$$m_y \le m_x \le x_1 \alpha^{r_1 - 1} < y_1 \alpha^{r_1} \le M_y, \tag{2.5}$$

where we have used one of our requirements on α as well as the facts that $x_1 \alpha^{r_1 - 1}$ is in the sum in (2.3) and $y_1 \alpha^{r_1}$ is not in the sum in (2.4). But (2.5) contradicts our assumption that $M_y \le m_y$, so the first case is complete.

In the other case $r_1 = k$, so $m_x \le x_1 \alpha^{k-1}$. But $x_1 \alpha^{k-1} < x_d$ by our choice of k, so we must have $r_d > 0$. Our assumptions imply that $M_y \ge M_x$, since otherwise we would have

$$\sum_{i=1}^{\ell+1}(x \otimes z)_i^{\downarrow} > \sum_{i=1}^{\ell+1}(y \otimes z)_i^{\downarrow}.$$

Therefore,

$$M_y \geq M_x \geq x_d \alpha^{r_d} > y_d \alpha^{r_d - 1} \geq m_y$$

by reasoning similar to that yielding (2.5). Again our assumption that $M_y \leq m_y$ is contradicted. Thus the proof is complete. □

Corollary 2.3.2 *Suppose x and y are d-dimensional probability vectors, with components arranged in non-increasing order, such that $x \prec_T y$ and $y_1 > x_1$ and $y_d < x_d$. Then x is in the interior of $T(y)$.*

Proof By definition there exists a z such that $x \otimes z \prec y \otimes z$. Since $y_1 > x_1$ and $y_d < x_d$ we must have $(x \otimes z)_1^{\downarrow} < (y \otimes z)_1^{\downarrow}$ and $(x \otimes z)_{dk}^{\downarrow} > (y \otimes z)_{dk}^{\downarrow}$, where k is the dimension of z.

We can thus apply Lemma 2.3.1 and conclude that $x \otimes z$ is in the interior of $T(y \otimes z)$. Since $x \mapsto x \otimes z$ is a continuous function, it follows that x is in the interior of $\{x \mid x \otimes z \in T(y \otimes z)\}$. But $\{x \mid x \otimes z \in T(y \otimes z)\} = T(y)$, so we are done. □

2.4 When Is Catalysis Useful?

If $T(y) = S(y)$, then catalysis is of no help in producing the state y. This is obviously the case when $y = (1, 0, \ldots, 0)$, for then all vectors in R^d are in both $S(y)$ and $T(y)$. Jonathan and Plenio have shown [20] that if $d \leq 3$ then $x \prec_T y \Rightarrow x \prec y$; in other words, $S(y) = T(y)$ if y is at most three-dimensional. The following theorem shows that for almost all vectors y of four or more dimensions, $S(y) \neq T(y)$:

Theorem 2.4.1 *Let $y = (y_1, \ldots, y_d)$ be a d-dimensional probability vector whose components are in non-increasing order. Then $T(y) \neq S(y)$ if and only if $y_1 \neq y_l$ and $y_m \neq y_d$ for some l, m with $1 < l < m < d$.*

This theorem says that $S(y) \neq T(y)$ if and only if y has at least two components that are distinct from both its smallest and largest components.

Proof Suppose that there exist such l and m. Let d_1 be the number of components of y equal to y_1, and let d_2 be the number of components of y equal to y_d. Then $d_1 + d_2 + 2 \leq d$. Let x be the d-dimensional vector whose first $d_1 + 1$ components are each equal to the average of the first $d_1 + 1$ components of y, whose last $d_2 + 1$ components are each equal to the average of the last $d_2 + 1$ components of y, and which matches y in any other components. Then it is easily checked that $x \prec y$. In fact x is on the boundary of $S(y)$ since $\sum_{i=1}^{d_1+1} x_i = \sum_{i=1}^{d_1+1} y_i$. However, by Corollary 2.3.2, x is in the interior of $T(y)$; thus $S(y) \neq T(y)$.

Conversely, assume that there are no l, m such that $l < m$, $y_1 \neq y_l$, and $y_m \neq y_d$. Again let d_1 be the number of components of y equal to y_1, and d_2 the number of components equal to y_d. Let $x \in T(y)$ and assume the components of x are arranged in decreasing order. Then $x_1 \leq y_1$, so $\sum_{i=1}^{j} x_i \leq \sum_{i=1}^{j} y_i$ for $j \in \{1, \ldots, d_1\}$. Also $x_d \geq y_d$, so $\sum_{i=j+1}^{d} x_i \geq \sum_{i=j+1}^{d} y_i$, and therefore $\sum_{i=1}^{j} x_i \leq \sum_{i=1}^{j} y_i$, for $j \in \{d - d_2, \ldots, d - 1\}$. But our assumptions imply that $d_1 + d_2 + 1 \geq d$, so in fact $\sum_{i=1}^{j} x_i \leq \sum_{i=1}^{j} y_i$ for all $j \in \{1, \ldots, d - 1\}$, and so $x \prec y$. Thus in this case $S(y) = T(y)$. $\square$

In applying this theorem, it should be noted that the dimension of y is somewhat arbitrary, as one can append zeroes to the vector y and thereby increase its dimension without changing the underlying quantum state. If y has at least three nonzero components, but exactly two distinct nonzero components, then appending zeroes will result in a vector y' such that $S(y') \neq T(y')$, although $S(y) = T(y)$. The reason for this phenomenon is that we only consider vectors x with the same dimension as that of y; by increasing the dimension of y, we increase the allowed choices for x as well. Thus, the dimension of the initial states x under consideration may determine whether $S(y) = T(y)$.

2.5 Catalysts of Arbitrarily High Dimension Must Be Considered

We will now show that for most y, there is no k such that $T_k(y) = T(y)$. In other words, there is no limit to the dimension of the catalysts that must be considered, in trying to determine which vectors are trumped by a given vector y. Our proof will proceed as follows: First we will show that $T_k(y)$ is a closed set for any k and all y, and then we will show that $T(y)$ is in general not closed. It follows that $T_k(y) \neq T(y)$.

The results of the previous section, and of this section, give a precise characterization of when $S(y) = T(y)$, and when there exists a k such that $T_k(y) = T(y)$. While it is clear that the former situation implies the latter, it turns out that the converse is true as well.

Proposition 2.5.1 *$T_k(y)$ is closed.*

Proof For a given d-dimensional probability vector y, let

$$h(x,z) = \max_{1\leq j<dk} \sum_{i=1}^{j} \left((x \otimes z)_i^{\downarrow} - (y \otimes z)_i^{\downarrow} \right),$$

where x and z are probability vectors of d and k dimensions, respectively. Observe that h is a composition of continuous functions (including the maximum of a finite set of expressions, and the function $x \mapsto x^{\downarrow}$), and so is continuous in x and z.

Let

$$f(x) = \min_z h(x,z),$$

where the minimum is over all k-dimensional probability vectors z; this minimum exists since $h(x,z)$ is continuous in z and the minimization is over a compact set. Observe that $x \in T_k(y)$ if and only if $f(x) \leq 0$.

Suppose now that $x \notin T_k(y)$. Then $f(x) > \epsilon$ for some $\epsilon > 0$. Let x' be given with $\|x - x'\| < \epsilon/d$. Let z be an arbitrary k-dimensional probability vector, let j_0 be a maximizing value of j in $h(x,z)$ and π be a permutation for which $(x \otimes z)_i^{\downarrow} = (x \otimes z)_{\pi(i)}$ for each i. Let v be the d-dimensional vector $(\epsilon/d, \ldots, \epsilon/d)$ and note that $x'_i > x_i - v_i$

for each i. We then have

$$\begin{aligned} h(x', z) - h(x, z) &\geq \sum_{i=1}^{j_0} \Big((x' \otimes z)_i^{\downarrow} - (x \otimes z)_i^{\downarrow} \Big) \\ &\geq \sum_{i=1}^{j_0} \big((x' \otimes z)_{\pi(i)} - (x \otimes z)_{\pi(i)} \big) \\ &> \sum_{i=1}^{j_0} \big(((x - v) \otimes z)_{\pi(i)} - (x \otimes z)_{\pi(i)} \big) \\ &= -\sum_{i=1}^{j_0} (v \otimes z)_{\pi(i)} \\ &\geq -\sum_{i=1}^{dk} (v \otimes z)_{\pi(i)} \\ &= -\epsilon. \end{aligned}$$

Therefore $h(x', z) > 0$ for all z, so $f(x') > 0$. We thus see that $x' \notin T_k(y)$ for x' in a neighborhood of x. Therefore $T_k^c(y)$ is open, so $T_k(y)$ is closed. □

Theorem 2.5.2 *Let $y = (y_1, \ldots, y_d)$ be a d-dimensional probability vector, with components in non-increasing order, such that $T(y) \neq S(y)$. Then $T(y)$ is not closed. In particular, for all k, $T_k(y) \neq T(y)$.*

Proof. By Theorem 2.4.1, the hypothesis is equivalent to the existence of l, m such that $1 < l < m < d$, $y_1 > y_l$, $y_m > y_d$. For convenience, we redefine l to be the index of the first component of y that is not equal to y_1, and m to be the index of the last component of y that is not equal to y_d; clearly we still have $l < m$. Let $\Delta = \min\{y_1 - y_l, y_m - y_d\}$ and let x be the d-dimensional vector given by $x_l = y_l + \Delta$, $x_m = y_m - \Delta$, and $x_i = y_i$ for $i \notin \{l, m\}$. It is easily checked that $y \prec x$ but $x \not\prec y$; therefore $x \not\prec_T y$. Let $w = (\frac{1}{d}, \ldots, \frac{1}{d})$ and note that $w \in S(y)$.

Suppose $T(y)$ is closed. Since $T(y)$ is convex, the set $\{t \in [0, 1] \mid tx + (1 - t)w \in T(y)\}$ is a closed interval not containing 1, say $[0, t_0]$. So $T(y)$ contains $t_0 x + (1 - t_0)w$ as a boundary point. But $t_0 x + (1 - t_0)w$ satisfies the hypotheses of Corollary 2.3.2 and is thus an interior point of $T(y)$. This is a contradiction, so $T(y)$ cannot be closed. As Theorem 2.5.1 says that each $T_k(y)$ is closed, we must have $T_k(y) \neq T(y)$.

□

So whenever catalysis is useful in producing y (i.e., $S(y) \neq T(y)$), catalysts of arbitrarily high dimension must be considered. In other words, when $S(y) \neq T(y)$, then for any k there is a $k' > k$ such that $T_k(y)$ is a strict subset of $T_{k'}(y)$. However, we do not know whether increasing the catalyst dimension by one will necessarily give an improvement. That is, it is unknown whether there is any vector y and $k \geq 1$ such that $S(y) \neq T_k(y)$ but $T_k(y) = T_{k+1}(y)$.

In the study of ELOCC transformations, one hoped-for phenomenon is the existence of a easily described *universal set of catalysts*. This is a set S of states z such that if $x \prec_T y$, then there exists $z \in S$ such that $x \otimes z \prec y \otimes z$. Such a set would be interesting both theoretically, and also useful from a practical perspective, as it would limit which states might be needed in a laboratory in order to perform certain transformations. However, one consequence of Theorem 2.5.2 is that no finite set can be a universal set of catalysts:

Corollary 2.5.3 *Any universal set of catalysts must be an infinite set.*

Proof Let y be any vector for which $S(y) \neq T(y)$. If S is a finite set, let k be the highest dimension of any state in S. Then if S is universal, $T_k(y) = T(y)$, contradicting Theorem 2.5.2. □

Chapter 3

Additional Properties

In this chapter, we derive various additional properties of the trumping relation. We show that virtually all states are useful as catalysts, and we study the generalization of catalysis to probabilistic LOCC transformations. We also examine how the notion of Schur-convexity applies to trumping.

3.1 Which states Can Be catalysts?

One interesting question is that of which states are potentially useful as catalysts. If a vector z is *uniform*, meaning that its nonzero components are all identical, then it is easily seen that z is not capable of acting as a catalyst: if $x \otimes z \prec y \otimes z$, then $x \prec y$ so z served no use as a catalyst. In [6] Nielsen conjectured that all nonuniform vectors are potentially useful as catalysts. In this section, we show that this conjecture is true.

Before we proceed, let us consider the implications of this conjecture. We know already that a uniform z cannot act as a catalyst. A uniform z with k nonzero components corresponds to a maximally entangled quantum state of Schmidt number k; if $k = 1$ then the state is unentangled. So we have the following situation: if z is a maximally entangled state, then z cannot be used as a catalyst; but for any other entangled state z, the conjecture says that z can serve as a catalyst. In using entanglement as a resource, it is possible to have too much as well as too little.

Theorem 3.1.1 *Let $z = (z_1, \ldots, z_k)$ be a non-uniform probability vector. Then there exist probability vectors $x, y \in R^4$ such that $x \otimes z \prec y \otimes z$, but $x \not\prec y$.*

Proof We may assume without loss of generality that $z_1 \geq z_2 \geq \cdots \geq z_k > 0$. Define α and β by the relations

$$\frac{z_1}{z_k} = \frac{\alpha}{\beta}$$

and

$$\alpha + \beta = 1.$$

By non-uniformity of z, $\alpha > \beta$.

Let $x_1 = x_2 = \frac{1}{2}\alpha + \frac{1}{4}\beta$, and $x_3 = x_4 = \frac{1}{4}\beta$. Let $y_1 = \alpha$, let $y_2 = y_3 = \frac{1}{2}\beta$, and let $y_4 = 0$. Let $x = (x_1, x_2, x_3, x_4)$, $y = (y_1, y_2, y_3, y_4)$. Note that $x \prec y$, so obviously $x \otimes z \prec y \otimes z$. Our goal is to show that all the majorization inequalities between $x \otimes z$ and $y \otimes z$ are strict; in other words, for all $\ell \in \{1, 2, \ldots, 4k-1\}$,

$$\sum_{i=1}^{\ell} (x \otimes z)_i^{\downarrow} < \sum_{i=1}^{\ell} (y \otimes z)_i^{\downarrow}. \tag{3.1}$$

We will show first that the inequalities are strict when ℓ is even; so for now, assume that ℓ is even. There are five cases to consider.

Case 1: $1 \leq \ell \leq k$. We have

$$\sum_{i=1}^{\ell} (x \otimes z)_i^{\downarrow} = (\alpha + \frac{1}{2}\beta) \sum_{i=1}^{\ell/2} z_i,$$

while

$$\sum_{i=1}^{\ell} (y \otimes z)_i^{\downarrow} = \alpha \sum_{i=1}^{\ell} z_i.$$

Thus

$$\begin{aligned}\sum_{i=1}^{\ell}(y\otimes z)_i^{\downarrow} - \sum_{i=1}^{\ell}(x\otimes z)_i^{\downarrow} &= \alpha \sum_{i=\ell/2+1}^{l} z_i - \frac{1}{2}\beta\sum_{i=1}^{\ell/2} z_i \\ &= \sum_{i=1}^{\ell/2}(\alpha z_{\ell/2+i} - \frac{1}{2}\beta z_i).\end{aligned}$$

This last quantity is a sum of positive terms (by the definition of α and β), so the inequality (3.1) is strict.

Case 2: $k+1 \le \ell < 2k$. We have

$$\sum_{i=1}^{\ell}(x\otimes z)_i^{\downarrow} = (\alpha + \frac{1}{2}\beta)\sum_{i=1}^{\ell/2} z_i$$

and

$$\sum_{i=1}^{\ell}(y\otimes z)_i^{\downarrow} \ge \alpha + \frac{1}{2}\beta\sum_{i=1}^{\ell-k} z_i.$$

The difference thus satisfies

$$\sum_{i=1}^{\ell}(y\otimes z)_i^{\downarrow} - \sum_{i=1}^{\ell}(x\otimes z)_i^{\downarrow} \ge \alpha \sum_{i=\ell/2+1}^{k} z_i - \frac{1}{2}\beta \sum_{i=\ell-k+1}^{\ell/2} z_i.$$

Note that the sums on the right hand side each contain $k-\ell/2$ terms. Since $\alpha z_i > \frac{1}{2}\beta z_j$ for any i, j, the difference is positive, and again (3.1) holds.

Case 3: $\ell = 2k$. In this case

$$\sum_{i=1}^{\ell}(x\otimes z)_i^{\downarrow} = \alpha + \frac{1}{2}\beta$$

and

$$\sum_{i=1}^{\ell}(y\otimes z)_i^{\downarrow} \ge \alpha + \frac{1}{2}\beta\sum_{i=1}^{k-1} z_i + \frac{1}{2}\beta z_1$$

$$= \alpha + \frac{1}{2}\beta + \frac{1}{2}\beta(z_1 - z_k) > \alpha + \frac{1}{2}\beta,$$

so the inequality 3.1 is strict.

Case 4: $2k+1 \leq \ell \leq 3k$. We have

$$\sum_{i=1}^{\ell}(x \otimes z)_i^{\downarrow} = \alpha + \frac{1}{2}\beta + \frac{1}{2}\beta \sum_{i=1}^{\ell/2-k} z_i$$

while

$$\sum_{i=1}^{\ell}(y \otimes z)_i^{\downarrow} \geq \alpha + \frac{1}{2}\beta + \frac{1}{2}\beta \sum_{i=1}^{\ell-2k} z_i.$$

The second quantity is clearly larger, so the inequality 3.1 is strict.

Case 5: $3k+1 \leq \ell < 4k$. This case is trivial because the sum for $y \otimes z$ is 1 (because there are no more nonzero terms to be added), and the sum for $x \otimes z$ is less than 1.

We have shown that (3.1) holds when ℓ is even (and in the proper range). Now suppose ℓ is odd. From the even cases, it is easily verified that

$$\sum_{i=1}^{\ell-1}(x \otimes z)_i^{\downarrow} + \sum_{i=1}^{\ell+1}(x \otimes z)_i^{\downarrow} < \sum_{i=1}^{\ell-1}(y \otimes z)_i^{\downarrow} + \sum_{i=1}^{\ell+1}(y \otimes z)_i^{\downarrow} \tag{3.2}$$

when $\ell \in \{1, 3, \ldots, 4k-1\}$. Based on the fact that the components of $(y \otimes z)^{\downarrow}$ are non-increasing, $\sum_{i=1}^{\ell}(y \otimes z)_i^{\downarrow}$ is greater than or equal to the average of the two sums in the right side of (3.2). However, $\sum_{i=1}^{\ell}(x \otimes z)_i^{\downarrow}$ is *equal* to the average of the sums in the left side of (3.2), since the components of $(x \otimes z)^{\downarrow}$ appear in pairs. We therefore see that (3.1) holds when ℓ is odd.

Thus, the majorization inequalities are strict for all ℓ between 1 and $4k-1$ inclusive, so for sufficiently small ϵ, $(x_1+\epsilon, x_2+\epsilon, x_3-\epsilon, x_4-\epsilon) \otimes z \prec y \otimes z$. However, $(x_1+\epsilon, x_2+\epsilon, x_3-\epsilon, x_4-\epsilon) \not\prec y$, so our theorem is proved. □

3.2 Probabilistic Catalysis

If $x \not\prec y$, then Theorem 2.1.1 tells us that there is no LOCC protocol that performs the transformation $x \rightarrow y$. However, it may still be possible to produce y given x, using only local operations and classical communication, if we are willing to accept some probability of failure. This situation is considered in [21], where a protocol optimizing the probability of success is presented. Let $P(x \to y)$ be the maximum probability of success of transforming x to y using LOCC. Then we have the following result [21].

Theorem 3.2.1 $P(x \to y) = \min_{\ell} \frac{\sum_{i=\ell}^{d} x_i^{\downarrow}}{\sum_{i=\ell}^{d} y_i^{\downarrow}}$.

Note that if $x \prec y$, then the numerator in the expression of Theorem 3.2.1 is always greater than or equal to the denominator, with equality when $\ell = 1$, so the theorem reduces to the statement that $P(x \to y) = 1$ in this case.

Theorem 3.2.1 suggests that we consider *probabilistic catalysis*: situations where $P(x \otimes z \to y \otimes z) > P(x \to y)$, even though $x \not\prec_T y$. The following result is analogous to Theorem 2.5.2:

Theorem 3.2.2 *Suppose $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d)$ are probability vectors with components in non-increasing order. Suppose $x_1 \leq y_1$ and $x_d \geq y_d$. Then either (1) $x \prec_T y$ or (2) There is no z (of any dimension) such that $P(x \otimes z \to y \otimes z)$ is maximized.*

Proof Let $p = P(x \to y) < 1$ (if $p = 1$, we are done). Define $\Delta = 1 - p\sum_{i=2}^{d} y_i$, and let $y'(p) = (\Delta, py_2, py_3, \ldots, py_d)$. Since for any $\ell \geq 2$, $\sum_{i=\ell}^{d} x_i \geq p\sum_{i=\ell}^{d} y_i = \sum_{i=\ell}^{d} y_i'(p)$, $x \prec y'(p)$. Also, it is easy to see that $x_1 < y_1'(p)$ and $x_d > y_d'(p)$.

By Lemma 2.3.1, this implies that there exists a catalyst z (of dimension, say, n) such that for all $\ell \in \{2, \ldots, nd\}$, $\sum_{i=\ell}^{nd}(x \otimes z)_i^{\downarrow} > \sum_{i=\ell}^{nd}(y'(p) \otimes z)_i^{\downarrow}$. Since $(y'(p) \otimes z) \prec (y \otimes z)'(p)$, it follows that for every $\ell \in \{2, \ldots, d\}$, $\sum_{i=\ell}^{nd}(x \otimes z)_i^{\downarrow} > \sum_{i=\ell}^{nd}((y \otimes z)'(p))_i^{\downarrow} = p\sum_{i=\ell}^{nd}(y \otimes z)_i^{\downarrow}$. Therefore, we have that $P(x \otimes z \to y \otimes z) > p$.

We have shown that whenever x and y satisfy the conditions of the lemma with $P(x \to y) < 1$, there must exist a catalyst z such that $P(x \otimes z \to y \otimes z) > P(x \to y)$. But if x and y satisfy the hypotheses of the lemma, then so do $x \otimes z$ and $y \otimes z$, so (assuming that $P(x \otimes z \to y \otimes z) < 1$) there is another catalyst w such that $P(x \otimes z \otimes w \to y \otimes z \otimes w) > P(x \otimes z \to y \otimes z)$. In other words, there can be no z

that maximizes the probability of transformation (unless this probability is one, i.e., x is trumped by y). $\square$

Similar results hold if the requirement $x_1 \leq y_1$ and $x_d \geq y_d$ are relaxed:

Theorem 3.2.3 *Suppose $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d)$ are probability vectors with components in non-increasing order.*

(a) If $x_d < y_d$, then if $\max_z P(x \otimes z \to y \otimes z)$ exists, it is equal to $\frac{x_d}{y_d}$.

(b) If $x_d \geq y_d$, then if $\max_z P(x \otimes z \to y \otimes z)$ exists, it is equal to 1.

Proof We divide the analysis into three cases: $x_d < y_d$ and $x_1 \leq y_1$, $x_1 > y_1$ and $x_d \geq y_d$, and $x_1 > y_1$ and $x_d < y_d$. (The case $x_1 \leq y_1$ and $x_d \geq y_d$ was proven in the previous theorem.)

Suppose that $x_d < y_d$ and $x_1 \leq y_1$. It follows that $p \equiv P(x \to y) \leq \frac{x_d}{y_d} \equiv q$. The interesting case is where $p < q$, so let's assume that. Suppose that z maximizes $P(x \otimes z \to y \otimes z) \equiv p* < q$ (the interesting case is $p* < q$). Then note that $P(x \otimes z \to (y \otimes z)'(q)) = \frac{p*}{q}$. So, as before there must be a w such that $P(x \otimes z \otimes w \to (y \otimes z)'(q) \otimes w) > \frac{p*}{q}$, and hence $P(x \otimes z \otimes w \to (y \otimes z \otimes w)'(q)) > \frac{p*}{q}$. But, since $P((y \otimes z \otimes w)'(q) \to y \otimes z \otimes w) = q$, it follows that $P(x \otimes z \otimes w \to y \otimes z \otimes w) > \frac{p*}{q} q = p*$, contradicting the assumption that z maximized the probability. So there is no z that maximizes the probability of transformation.

Next suppose that $x_1 > y_1$ and $x_d \geq y_d$. Suppose that z maximizes $P(x \otimes z \to y \otimes z) \equiv p* < 1$. Let $q_1 = \frac{1 - x_1 z_1}{1 - y_1 z_1}$, and note that without loss of generality, we may assume $p* < q_1$. Now $P(x \otimes z \to (y \otimes z)'(q_1)) = \frac{p*}{q_1}$, so there exists a w such that $P(x \otimes z \otimes w \to (y \otimes z)'(q_1) \otimes w) > \frac{p*}{q_1}$, which implies that $P(x \otimes z \otimes w \to (y \otimes z \otimes w)'(q_1)) > \frac{p*}{q_1}$. It follows that $P(x \otimes z \otimes w \to y \otimes z \otimes w) > p*$, a contradiction. So there can be no such z.

Finally, suppose that $x_1 > y_1$ and $x_d < y_d$. Suppose that z maximizes $P(x \otimes z \to y \otimes z) \equiv p*$. Let q_1 be as before, and let $q_2 = \frac{x_d}{y_d}$. Without loss of generality, we may assume that $q_1 > q_2$. Now $P(x \otimes z \to (y \otimes z)'(q_2)) = \frac{p*}{q_2}$. Applying the same reasoning as before, we get a contradiction unless $p = q_2$. $\square$

3.3 Additive Schur-Convexity

A subclass of Schur-convex functions can be used to give necessary conditions for $x \prec_T y$. Let S_d be the set of probability vectors in $\mathbb{R}^d$. A family of functions $f_d : S_d \subset \mathbb{R}^d \to \mathbb{R}$ is said to be *additive* if the following holds: If $x \in \mathbb{R}^{d_1}$ and $x' \in \mathbb{R}^{d_2}$ then $f_{d_1 d_2}(x \otimes x') = f_{d_1}(x) + f_{d_2}(x')$. Then we have [6]

Theorem 3.3.1 *Let $\{f_d\}_{d=1}^{\infty}$ be an additive family of functions, each of which is Schur-convex. Then each f_d has the property that if x and y are probability vectors in $\mathbb{R}^d$ such that $x \prec_T y$, then $f_d(x) \leq f_d(y)$.*

Proof If $x \prec_T y$, then there exists some positive integer k and probability vector $z \in \mathbb{R}^k$ such that $x \otimes z \prec y \otimes z$. Thus, $f_{dk}(x \otimes z) \leq f_{dk}(y \otimes z) \Rightarrow f_d(x) + f_k(z) \leq f_d(y) + f_k(z) \Rightarrow f_d(x) \leq f_d(y)$. $\square$

Just as Schur-convex functions respect the majorization relation, so additive Schur-convex functions respect the trumping relation. However, the situation is not completely analogous because there may be functions respecting the trumping relation that do not fall into an additive Schur-convex family of functions. The following is a list of the known families of additive Schur-convex functions:

- The negative of the rank function, $f_d(x) = -$ (the number of nonzero components of x).
- The max function, $f_d(x) =$ the largest component of x.
- The negative of the min function, $f_d(x) = -$ (the smallest component of x).
- The negative of the entropy function, $f_d(x) = \sum_{i=1}^{d} x_i \log x_i$.
- The log of the product function, $f_d(x) = \sum_{i=1}^{d} \log x_i$ (only if all $x_i \neq 0$, otherwise $f_d(x) = -\infty$).
- The log of the power sums: for any real $\alpha \notin [0, 1]$, $f_d(x) = \log \sum_{i=1}^{d} x_i^{\alpha}$ (where the sum is defined to be $-\infty$ if $k \leq 0$ and any $x_i = 0$), and for $\alpha \in (0, 1)$, $f_d(x) = -\log \sum_{i=1}^{d} x_i^{\alpha}$.

Instead of using the log of the power sums, one may just as well use the power sums themselves, $f_d(x) = \sum_{i=1}^{d} x_i^d$; because the log function is monotonic, this is equivalent to using the logs of the power sums. The fact that the other functions on the list (the rank, max, negative min, negative entropy, and log-product functions) respect the trumping relation is a consequence of the fact that the power sum functions do. The negative rank, max, and negative min functions can be considered to be limiting cases of the power sums when α goes to 0 (from above), ∞, and $-\infty$, respectively. That the negative entropy and log of the product functions must respect the trumping relation can be seen by taking the derivative with respect to α of the power sum function at $\alpha = 0$ and $\alpha = 1$, respectively, and noting that this derivative must be positive (because equality holds at $\alpha = 0$ for vectors of the same rank, and at $\alpha = 1$ for all probability vectors). Thus, all known additive Schur-convex functions can be thought of as special cases of the power sum functions. In light of this, M. Nielsen has conjectured that $x \prec_T y$ if and only if for all real $\alpha < 0$ or $\alpha > 1$,

$$\sum_{i=1}^{d} x_i^{\alpha} \leq \sum_{i=1}^{d} y_i^{\alpha}, \tag{3.3}$$

and for all real $\alpha \in (0, 1)$,

$$\sum_{i=1}^{d} x_i^{\alpha} \geq \sum_{i=1}^{d} y_i^{\alpha}. \tag{3.4}$$

This intruiging conjecture has not yet been settled.

Chapter 4

Examples

This chapter gives concrete examples that illustrate various features of the trumping relation. Many of these examples were found in attempts to prove conjectures made by the author or others. In a sense, the results presented here are disappointing, since they often highlight ways in which the trumping relation is not as well-behaved as one might wish for it to be.

4.1 The Simplest Non-trivial Case

From Theorem 2.4.1, it follows that $T(y) = S(y)$ when y is of dimension three or smaller (in [20], this fact is proven directly). Furthermore, it is clear that catalysis cannot occur unless the catalyst state has dimension at least two. So the simplest (lowest-dimensional) case of catalysis occurs when y is four-dimensional and the catalyst is two-dimensional. We will analyze this simplest case to suggest properties of $T_k(y)$ and $T(y)$ in general.

In [22], P. H. Anspach gives a categorization of $T_2(y)$, when y is four-dimensional. We will use this result extensively, so we state it here. Let $y = (y_1, y_2, y_3, y_4)$, $x = (x_1, x_2, x_3, x_4)$, with components arranged in non-increasing order. In order for $x \not\prec y$

yet $x \prec_T y$, there must exist $\epsilon_1 > 0$, $\epsilon_2 > 0$, $\epsilon_3 \geq 0$ such that [22]

$$\begin{aligned} y_1 &= x_1 + \epsilon_1 && (4.1)\\ y_2 &= x_2 - \epsilon_1 - \epsilon_2 && (4.2)\\ y_3 &= x_3 + \epsilon_2 + \epsilon_3 && (4.3)\\ y_4 &= x_4 - \epsilon_3. && (4.4) \end{aligned}$$

However, this necessary condition is not sufficient. Anspach's result is the following [22]:

Theorem 4.1.1 *For x and y as above, let*

$$m = \max\left(\frac{y_2 + \epsilon_2}{y_1}, \frac{y_4}{y_3 - \epsilon_2}, \frac{\epsilon_2}{\epsilon_1}\right) \tag{4.5}$$

$$M = \min\left(\frac{y_3 - \epsilon_2}{y_2 + \epsilon_2}, \frac{\epsilon_3}{\epsilon_2}\right) \tag{4.6}$$

Then $x \not\prec y$, but $x \in T_2(y) \iff m \leq M$. Moreover if $m \leq M$, then $z = (p, 1-p)$ (where $p \geq 0.5$) will be a catalyst iff $m \leq \frac{1-p}{p} \leq M$.

This concrete description allows us to determine some properties of $T_2(y)$ when y is four-dimensional. In the next section, for example, we will use it to show that $T_2(y)$ is convex in this case. We also have the following result, answering the question of whether there is a universal set of catalyts for the case of $T_2(y)$, where y is four dimensional:

Theorem 4.1.2 *Let $y = (0.5, 0.25, 0.25, 0)$. Then there is no countably infinite set of two-dimensional catalysts $\{z_i\}_{i \in \mathbb{Z}}$ such that $T_2(y) = \bigcup_{i=1}^{\infty} T(y, z)$.*

In other words, there is no countably infinite set of two-dimensional catalysts that is universal for determining $T_2(y)$.

Proof For $\epsilon \in (0.029, 0.031)$, choose $x(\epsilon) = (x_1(\epsilon), x_2(\epsilon), x_3(\epsilon), x_4(\epsilon))$ as follows:

$$\begin{aligned} x_1(\epsilon) &= 0.45 \\ x_2(\epsilon) &= 0.30 + \epsilon \\ x_3(\epsilon) &= 0.25 - \epsilon - 20\epsilon^2 \\ x_4(\epsilon) &= 20\epsilon^2 \end{aligned}$$

Then

$$\begin{aligned} m &= \max\left(\frac{0.25+\epsilon}{0.5}, 0, \frac{\epsilon}{0.05}\right) = 20\epsilon \\ M &= \min\left(\frac{0.25+\epsilon}{0.25+\epsilon}, 20\epsilon\right) = 20\epsilon \end{aligned}$$

Thus, $m = M \implies$ the catalyst for this transformation is *unique*. So every state of this form requires its own unique catalyst, for $\epsilon \in (0.029, 0.031)$; therefore, no countable set of catalysts will be sufficient. □

4.2 Convexity and Catalysis

Convexity is a useful notion in determining what state transformations are possible under LOCC, because the set $S(y)$ can be described as the convex hull of a finite set of points (by Theorem 1.3.5). Since $T(y)$ is also a convex set, one naturally wishes to find its extreme points. Unfortunately, Theorem 2.5.2 tells us that in general, $T(y)$ is not a closed set, suggesting that this program will be far more difficult than it was for $S(y)$, as we know that $T(y)$ will not simply be the convex hull of some finite number of its elements. Thus, we may wish to attack this problem by considering the sets $T_k(y)$ (which we know to be closed, at least, by Proposition 2.5.1). If each $T_k(y)$ has a tractable description, it may lead to a nice characterization of $T(y) = \bigcup_k T_k(y)$. Thus, we wish to know whether the sets $T_k(y)$ are convex in general.

For the simplest non-trivial case, $T_k(y)$ is indeed convex:

Theorem 4.2.1 *Let y be four-dimensional. Then $T_2(y)$ is convex.*

Proof Let x and x' be elements of $T_2(y)$, and let $\lambda \in [0,1]$. If x and x' are both in $S(y)$, then so is $\lambda x + (1-\lambda)x'$, because $S(y)$ is convex. If only one of x and x' is in $S(y)$, then without loss of generality assume $x \in S(y)$. Choose a two-dimensional z such that $x' \otimes z \prec y \otimes z$. Then since $x \prec y$, $x \otimes z \prec y \otimes z$, so (by convexity of $S(y \otimes z)$) it follows that $(\lambda x + (1-\lambda)x') \otimes z \prec y \otimes z$, so $(\lambda x + (1-\lambda)x') \in T_2(y)$.

Finally, suppose that $x \notin S(y)$ and $x' \notin S(y)$. This is the situation where Theorem 4.1.1 applies to both x and x' (and is far more involved to analyze than the previous two situations). We need to show that if $x, x' \not\prec y$ but $x \in T_2(y)$ and $x' \in T_2(y)$, then for all $\lambda \in (0,1)$, $\lambda x + (1-\lambda)x' \in T_2(y)$. So suppose there exist $\epsilon_1 > 0$, $\epsilon_2 > 0$, $\epsilon_3 \geq 0$ such that

$$\begin{aligned} y_1 &= x_1 + \epsilon_1 && (4.7) \\ y_2 &= x_2 - \epsilon_1 - \epsilon_2 && (4.8) \\ y_3 &= x_3 + \epsilon_2 + \epsilon_3 && (4.9) \\ y_4 &= x_4 - \epsilon_3 && (4.10) \end{aligned}$$

and similarly, that there exist $\delta_1 > 0$, $\delta_2 > 0$, $\delta_3 \geq 0$ such that

$$\begin{aligned} y_1 &= x'_1 + \delta_1 && (4.11) \\ y_2 &= x'_2 - \delta_1 - \delta_2 && (4.12) \\ y_3 &= x'_3 + \delta_2 + \delta_3 && (4.13) \\ y_4 &= x'_4 - \delta_3. && (4.14) \end{aligned}$$

Note that taking a convex combination of x and x' involves taking a convex combination of the difference terms ϵ_i and δ_i. That is, let $w = \lambda x + (1-\lambda)x'$. Let

$\gamma_i = \lambda\epsilon_i + (1-\lambda)\delta_i$, for $i = 1, 2, 3$. Then

$$\begin{aligned}
y_1 &= w_1 + \gamma_1 && (4.15)\\
y_2 &= w_2 - \gamma_1 - \gamma_2 && (4.16)\\
y_3 &= w_3 + \gamma_2 + \gamma_3 && (4.17)\\
y_4 &= w_4 - \gamma_3. && (4.18)
\end{aligned}$$

Now Theorem 4.1.1 can be restated as follows. In order for $x \in T_2(y)$, the following inequalities must hold:

$$\begin{aligned}
y_2^2 + 2y_2\epsilon_2 + \epsilon_2^2 &\leq y_1y_3 - y_1\epsilon_2 && (4.19)\\
y_2\epsilon_2 + \epsilon_2^2 &\leq y_1\epsilon_3 && (4.20)\\
y_2y_4 + y + 4\epsilon_2 &\leq y_3^2 - 2y_3\epsilon_2 + \epsilon_2^2 && (4.21)\\
y_4\epsilon_2 &\leq y_3\epsilon_3 - \epsilon_2\epsilon_3 && (4.22)\\
y_2\epsilon_2 + \epsilon_2^2 &\leq y_3\epsilon_1 - \epsilon_1\epsilon_2 && (4.23)\\
\epsilon_2^2 &\leq \epsilon_1\epsilon_3 && (4.24)
\end{aligned}$$

and if $x' \in T_2(y)$, then Inequalities (4.19-4.24) must hold if we replace each ϵ_i with the corresponding δ_i.

We need to show that if Inequalities (4.19-4.24) hold for $\epsilon_1, \epsilon_2, \epsilon_3$, and if they also hold when these are replaced with $\delta_1, \delta_2, \delta_3$, respectively, then they also hold when replaced with $\lambda\epsilon_1 + (1-\lambda)\delta_1, \lambda\epsilon_2 + (1-\lambda)\delta_2, \lambda\epsilon_3 + (1-\lambda)\delta_3$, respectively.

Before examining each inequality individually, we need the following notion. Two real vectors $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are said to be *similarly ordered* if for any indices $i, j \in \{1, \ldots, n\}$, $(a_i - a_j)(b_i - b_j) \geq 0$. It is a well-known fact (see, for example, Chapter 10 of [23]) that if $a(\pi) = (a_{\pi(1)}, \ldots, a_{\pi(n)})$ and $b(\sigma) = (b_{\sigma(1)}, \ldots, b_{\sigma(n)})$ are permutations of the vectors a and b, then the dot product $a(\pi) \cdot b(\sigma)$ is maximized when $a(\pi)$ and $b(\sigma)$ are similarly ordered.

Case 1: We must show that

$$y_2^2 + 2y_2(\lambda\epsilon_2 + (1-\lambda)\delta_2) + (\lambda\epsilon_2 + (1-\lambda\delta_2)^2 \leq y_1y_3 - y_1(\lambda\epsilon_2 + (1-\lambda)\delta_2,$$

using the corresponding inequalities for ϵ and δ,

$$y_2^2 + 2y_2\epsilon_2 + \epsilon_2^2 \leq y_1y_3 - y_1\epsilon_2$$

and

$$y_2^2 + 2y_2\delta_2 + \delta_2^2 \leq y_1y_3 - y_1\delta_2.$$

For convenience we will bring all our terms to one side, i.e., we must show that

$$y_1y_3 - y_1(\lambda\epsilon_2 + (1-\lambda)\delta_2) - (y_2^2 + 2y_2(\lambda\epsilon_2 + (1-\lambda)\delta_2) + (\lambda\epsilon_2 + (1-\lambda)\delta_2)^2) \geq 0.$$

Define $f(t) = y_1y_3 - y_1t - y_2^2 - 2y_2t - t^2$. Then f is a decreasing function of t, for t between δ_2 and ϵ_2. So if $f(\delta_2) \geq 0$ and $f(\epsilon_2) \geq 0$, then $f(\lambda\epsilon_2 + (1-\lambda)\delta_2) \geq 0$, QED.

Case 2: We must show that

$$y_1(\lambda\epsilon_3 + (1-\lambda)\delta_3) - y_2(\lambda\epsilon_2 + (1-\lambda)\delta_2) - (\lambda\epsilon_2 + (1-\lambda)\delta_2)^2 \geq 0.$$

By assumption, we have that

$$\lambda(y_1\epsilon_3 - y_2\epsilon_2 - \epsilon_2^2) + (1-\lambda)(y_1\delta_3 - y_2\delta_2 - \delta_2^2) \geq 0.$$

Comparing these two inequalities, we see that the first one is satisfied if

$$(\lambda\epsilon_2 + (1-\lambda)\delta_2)^2 \leq \lambda\epsilon_2^2 + (1-\lambda)\delta_2^2,$$

which follows from the Cauchy-Schwarz Inequality (applied to the vectors $(\sqrt{\lambda}, \sqrt{1-\lambda})$, $(\epsilon_2\sqrt{\lambda}, \delta_2\sqrt{1-\lambda})$).

Case 3: It follows from Eq. (4.3) that $\epsilon_2 \leq y_3$ (and also $\delta_2 \leq y_3$). Define $g(t) =$

$y_3^2 - 2y_3t + t^2 - y_2y_4 + y_4\epsilon_2$. Then $g'(t) = -2(y_3 - t) - y_4 \leq 0$ when t is between δ_2 and ϵ_2. Therefore, since $g(\delta_2) \geq 0$ and $g(\epsilon_2) \geq 0$, it follows that $g(\lambda\epsilon_2 + (1-\lambda)\delta_2) \geq 0$.

Case 4: Suppose first that the vectors (δ_2, ϵ_2) and (δ_3, ϵ_3) are not similarly ordered. Define $h(t_1, t_2) = y_3t_2 - y_4t_1 - t_1t_2$, $(t_1, t_2) \in [0, y_3] \times [0, y_3]$. Then h is decreasing in t_1 and increasing in t_2. It follows that h is monotonic on the line connecting (δ_2, δ_3) and (ϵ_2, ϵ_3). So if $h(\epsilon_2, \epsilon)$ and $h(\delta_2, \delta_3)$ are both positive, then so is $h(\lambda\epsilon_2 + (1-\lambda)\delta_2, \lambda\epsilon_3 + (1-\lambda)\delta_3)$, as desired.

Now suppose that (δ_2, ϵ_2) and (δ_2, ϵ_3) are similarly ordered. By assumption, we have that

$$\lambda(y_3\epsilon_3 - y_4\epsilon_2 - \epsilon_2\epsilon_3) + (1-\lambda)(y_3\delta_3 - y_4\delta_2 - \delta_2\delta_3) \geq 0.$$

From this, our desired inequality

$$y_3(\lambda\epsilon_3 + (1-\lambda)\delta_3) - (\lambda\epsilon_2 + (1-\lambda)\delta_2)(\lambda\epsilon_3 + (1-\lambda)\delta_3) - y_4(\lambda\epsilon_2 + (1-\lambda)\delta_2) \geq 0$$

will follow provided that $\epsilon_2\epsilon_3 + \delta_2\delta_3 \geq \epsilon_2\delta_3 + \delta_2\epsilon_3$. But this follows from the fact that (δ_2, ϵ_2) and (δ_3, ϵ_3) are similarly ordered.

Case 5: If (δ_1, ϵ_1) and (δ_2, ϵ_2) are not similarly ordered, then an argument identical to the one used in the previous case shows that the desired inequality holds. So suppose that (δ_2, ϵ_2) and (δ_3, ϵ_3) are similarly ordered. By assumption, we have that

$$\lambda(y_3\epsilon_1 - \epsilon_1\epsilon_2 - y_2\epsilon_2 - \epsilon_2^2) + (1-\lambda)(y_3\delta_1 - \delta_1\delta_2 - y_2\delta_2 - \delta_2^2) \geq 0.$$

From this, our desired inequality

$$y_3(\lambda\epsilon_1 + (1-\lambda)\delta_1) - (\lambda\epsilon_1 + (1-\lambda)\delta_1)(\lambda\epsilon_2 + (1-\lambda)\delta_2) - y_2(\lambda\epsilon_2 + (1-\lambda\delta_2) + (\lambda\epsilon_2 + (1-\lambda)\delta_2)^2 \geq 0$$

will follow if $\epsilon_1\epsilon_2 + \delta_1\delta_2 \geq \epsilon_1\delta_2 + \delta_1\epsilon_2$ and $\epsilon_2^2 + \delta_2^2 \geq 2\epsilon_2\delta_2$. The first of these inequalities is a consequence of (δ_1, ϵ_1) and (δ_2, ϵ_2) being similarly ordered; the second follows from $(\epsilon_2 - \delta_2)^2 \geq 0$.

Case 6: From $\epsilon_2^2 \leq \epsilon_1\epsilon_3$ and $\delta_2^2 \leq \delta_1\delta_3$, we get

$$\epsilon_2\delta_2 \leq \sqrt{(\epsilon_1\delta_3)(\delta_1\epsilon_3)} \leq \frac{1}{2}(\epsilon_1\delta_3 + \delta_1\epsilon_3)$$

where the last step follows from the Arithmetic Mean-Geometric Mean Inequality. Thus,

$$(\lambda\epsilon_2 + (1-\lambda)\delta_2)^2 = \lambda^2\epsilon_2^2 + (1-\lambda)^2\delta_2^2 + 2\lambda(1-\lambda)\epsilon_2\delta_2$$

$$\leq \lambda^2\epsilon_1\epsilon_3 + (1-\lambda)^2\delta_1\delta_3 + \lambda(1-\lambda)(\epsilon_1\delta_3 + \delta_1\epsilon_3) = (\lambda\epsilon_1 + (1-\lambda)\delta_1)(\lambda\epsilon_3 + (1-\lambda)\delta_3).$$

□

Using the description of $T_2(y)$ for four-dimensional y provided by Theorem 4.1.1, we were able to show that $T_2(y)$ is convex. In higher dimensions (of either the target state or the catalyst), however, the following examples suggest that characterizing $T_k(y)$ will be quite difficult.

Example 4.2.2 *For $y = (0.5, 0.25, 0.25, 0)$, the set $T_3(y)$ is not convex.*

To see this, let

$$x_1 = (0.455, 0.335, 0.185, 0.025),$$

$$x_2 = (0.405, 0.403, 0.178, 0.014),$$

$$z_1 = (0.412, 0.336, 0.252),$$

$$z_2 = (0.498, 0.309, 0.193).$$

Then direct calculation confirms that $x_1 \otimes z_1 \prec y \otimes z_1$ and $x_2 \otimes z_2 \prec y \otimes z_2$, so $x_1, x_2 \in T_3(y)$. However, if we set $\lambda = 0.3$, then $\lambda x_1 + (1-\lambda)x_2 \not\prec_{T_3} y$. The proof is by contradiction; one assumes a catalyst $z = (p_1, p_2, p_3)$ exists and uses the majorization inequalities to show that there can be no such p_1, p_2, p_3. However, it is mostly tedious simple arithmetic, and we omit it here. □

Example 4.2.3 *Let $y = (0.4, 0.25, 0.2, 0.15, 0)$. Then $T_2(y)$ is not convex.*

To see this, let

$$x_1 = (0.373, 0.295, 0.1696888, 0.1501556, 0.0121556),$$

$$x_2 = (0.392, 0.264, 0.1876896, 0.1531552, 0.0031552),$$

$$z_1 = (0.597, 0.403),$$

$$z_2 = (0.569, 0.431).$$

Then $x_1 \otimes z_1 \prec y \otimes z_1$ and $x_2 \otimes z_2 \prec y \otimes z_2$. However, if $\lambda = 0.1$, then $\lambda x_1 + (1-\lambda)x_2 \not\prec_{T_2} y$. Once again, we omit the tedious proof. □

Examples 4.2.2 and 4.2.3 lead us to make the following conjecture:

Conjecture 4.2.4 *If $S(y) \neq T_k(y)$, then $T_k(y)$ is not convex, except when $k = 2$ and y is four-dimensional.*

4.3 Infinite-dimensional Catalysts

In defining $T(y)$, we allow the dimension of catalyst states to be arbitrarily large. What if the dimension were actually infinite? Can we achieve more than we could with catalysts of arbitrarily large but finite dimension? The answer to this question is yes, as shown by the following example.

Example 4.3.1 *Let $x = (0.4, 0.4, 0.2), y = (0.5, 0.25, 0.25), \alpha = 2^{-\frac{1}{8}}$. Then $x \not\prec_T y$, but if $z = \frac{1}{1-\alpha}(1, \alpha, \alpha^2, \ldots, \alpha^n, \ldots)$, then $x \otimes z \prec y \otimes z$.*

Proof Because $x_3 < y_3$, it is impossible for $x \prec_T y$. It is straightforward to verify that for any $\ell > 0$, $\sum_{i=1}^{\ell}(x \otimes z)_i^{\downarrow} < \sum_{i=1}^{\ell}(y \otimes z)_i^{\downarrow}$, so $x \otimes z \prec y \otimes z$. □

4.4 Probability and Catalysis

Another question we may ask is how the probability of transforming one state to another via LOCC relates to our ability to catalyze such a transformation. An interesting result in this area was provided by Z. Zhou and G. Guo, who showed that [24]

Theorem 4.4.1 *If $x \prec_T y$, then for $n \geq 1$, $P(x^{\otimes n} \to y^{\otimes n}) \geq P(x \to y)$.*

In other words, if $x \to y$ under ELOCC, then in the absence of a catalyst, the success probability of transforming multiple copies of x into multiple copies of y under LOCC is at least as large as the probability of transforming one copy of x into one copy of y.

This result suggests that we ask the following question: can we place any bounds on $P(x \to y)$, given that $x \prec_T y$? It may seem that if $x \otimes z \prec y \otimes z$, then the probability $P(x \to y)$ should not be "too low." However, the following example shows that this intuition is incorrect.

Example 4.4.2 *Let $0 < c_1 \ll c_2 \ll 1$. Set $x = (\frac{1}{2} - c_1, \frac{1}{2} - c_1, c_1, c_1), y = (1 - 2c_1 - c_2, c_1 + \frac{1}{2}c_2, c_1 + \frac{1}{2}c_2, 0) \in R^4$. Then for any $\epsilon > 0$, we can choose c_1, c_2 such that $x \prec_T y$, but $P(x \to y) < \epsilon$.*

Proof To show this, note that $x, y \in \mathbb{R}^4$. Define $\epsilon_1, \epsilon_2, \epsilon_3$ as in Inequalities (4.1-4.4):

$$\epsilon_1 = y_1 - x_1 = \frac{1}{2} - c_1 - c_2 \tag{4.25}$$

$$\epsilon_2 = (x_1 + x_2) - (y_1 + y_2) = \frac{1}{2}c_2 - c_1 \tag{4.26}$$

$$\epsilon_3 = x_4 - y_4 = c_1 \tag{4.27}$$

Then we compute m and M as in Theorem 4.1.1, and find that $M = \frac{2c_1}{c_2}$, while $m = \max(\frac{c_2}{1-2c_1-c_2}, \frac{\frac{1}{2}c_2-c_1}{\frac{1}{2}-c_1-c_2}) \leq 2c_2 - 4c_1$ if $c_1, c_2 < \frac{1}{3}$. Meanwhile, $P(x \to y) = \frac{2c_1}{c_1+\frac{1}{2}c_2} < \frac{4c_1}{c_2}$. So let $\frac{c_1}{c_2} < \frac{\epsilon}{4}$, and let $c_2 < \frac{2c_1}{c_2}$. Then we will have that $m \leq M$, so that $x \in T_2(y)$ by Theorem 4.1.1 and therefore $x \prec_T y$, while $P(x \to y) < \epsilon$. □

The previous example shows that it is possible to find x, y such that $x \prec_T y$ (in fact, $x \in T_2(y)$) and $P(x \to y)$ is as small as desired. The next example shows that the probability enhancement achievable with a catalyst does not vary continuously with the catalyst.

Observation 4.4.3 *For fixed probability vectors $x, y \in \mathbb{R}^d$, and $z \in \mathbb{R}^n$, define $g_{x,y}(z) = P(x \otimes z \to y \otimes z)$. Then it is not true in general that $g_{x,y}$ is a continuous function of z.*

Proof We illustrate this with the following. Let $x = (0.4, 0.4, 0.1, 0.1)$, $y = (0.5, 0.25, 0.2, 0.05)$, and for $\epsilon \in [0, 0.4]$, define $z(\epsilon) = (0.6, 0.4 - \epsilon, \epsilon)$. Then it is easy to check that

$$\lim_{\epsilon \longrightarrow 0} g_{x,y}(\epsilon) = \lim_{\epsilon \longrightarrow 0} P(x \otimes z(\epsilon) \to y \otimes z(\epsilon)) = 0.8,$$

while

$$g_{x,y}(0) = P(x \otimes z(0) \to y \otimes z(0)) = \frac{20}{23} \approx 0.869.$$

$\square$

In general, the probability achievable with the aid of a catalyst becomes ill-behaved at points where the catalyst's Schmidt number is changing (i.e., where one of its components goes to zero).

We close with a conjecture relating the trumping relation to probabilistic catalysis. Define $T'(y) \equiv \{x | \sup_z P(x \otimes z \to y \otimes z) = 1\}$, where the supremum is taken over probability vectors z of any dimension. We conjecture the following:

Conjecture 4.4.4 $T'(y) = \overline{T(y)}$, *the closure of the set* $T(y)$.

Part II

On the Spectrum of a Partial Trace

Chapter 5

Introduction to Part II

In this chapter we describe the mathematical problem considered in Part II, and its physical significance. We also discuss a related problem, known as *Horn's problem*, which has been recently solved. Finally, we give a physical application of the solution to Horn's problem.

5.1 The Problem

Let $A = \mathbb{C}^{d_A}$, $B = \mathbb{C}^{d_B}$, and let ρ_{AB} be an operator on $A \otimes B$. We identify ρ_{AB} with its matrix in the standard basis, which has entries

$$\rho_{AB}^{ij,kl} = \langle i_A| \otimes \langle j_B|\rho_{AB}|k_A\rangle \otimes |l_B\rangle. \tag{5.1}$$

Define the *partial trace* $\rho_A = \operatorname{Tr}_B \rho_{AB}$ of ρ_{AB} to be the operator

$$\rho_A = \sum_k \langle k_B|\rho_{AB}|k_B\rangle \tag{5.2}$$

on A. The matrix entries of ρ_A are

$$\rho_A^{ij} = \sum_k \langle i_A| \otimes \langle k_B|\rho_{AB}|j_A\rangle \otimes |k_B\rangle. \tag{5.3}$$

Equivalently, given the matrix ρ_{AB}, we can define ρ_A to be the unique matrix such that

$$\mathrm{Tr}(\rho_{AB} X \otimes I_B) = \mathrm{Tr}(\rho_A X) \tag{5.4}$$

for all X on A, where I_B is the identity on B.

Our present work will focus on the following question: What is the relationship between the spectrum of ρ_{AB} and the spectrum of ρ_A? We generally adopt the point of view that the spectrum of ρ_{AB} is given and we wish to deduce what possible spectra of ρ_A may occur. (However, our final results will allow one to reason in the other direction as well; given the spectrum of ρ_A, one can deduce the possible spectra of ρ_{AB}.) We let $\mathcal{H}_{AB}(\lambda) = \{\rho_{AB} : \mathrm{Spec}(\rho_{AB}) = \lambda\}$ be the set of Hermitian matrices on $A \otimes B$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{d_A d_B}$; then our problem is to fully characterize the set $\mathcal{S}_A(\lambda) = \{\mathrm{Spec}(\rho_A) : \rho_{AB} \in \mathcal{H}_{AB}(\lambda)\}$. (We adopt the convention that $\mathrm{Spec}(X)$, the vector of eigenvalues of an operator X, is always written with components in non-increasing order.)

5.2 Physical Interpretation

Determining the possible spectra of a partial trace has a number of physical applications. The usual situation is to regard ρ_{AB} as the density matrix of a quantum system AB, a composite of subsystems A and B; ρ_A is then the density matrix of subsystem A. In this context, we are asking which quantum-mechanical descriptions of a subsystem of a quantum system are compatible with the description of the whole system.

Understanding the relationship between a density operator and its partial trace also allows us to characterize what state transformations are achievable using quantum communication. To illustrate, suppose two parties, Alice and Bob, share a state between them that can be described by a state vector $|\varphi_{ABC}\rangle \in A \otimes B \otimes C$, where Alice holds quantum systems A and C and Bob holds the system B. First, we assume that there will be only one round of quantum communication, from Alice to Bob. Al-

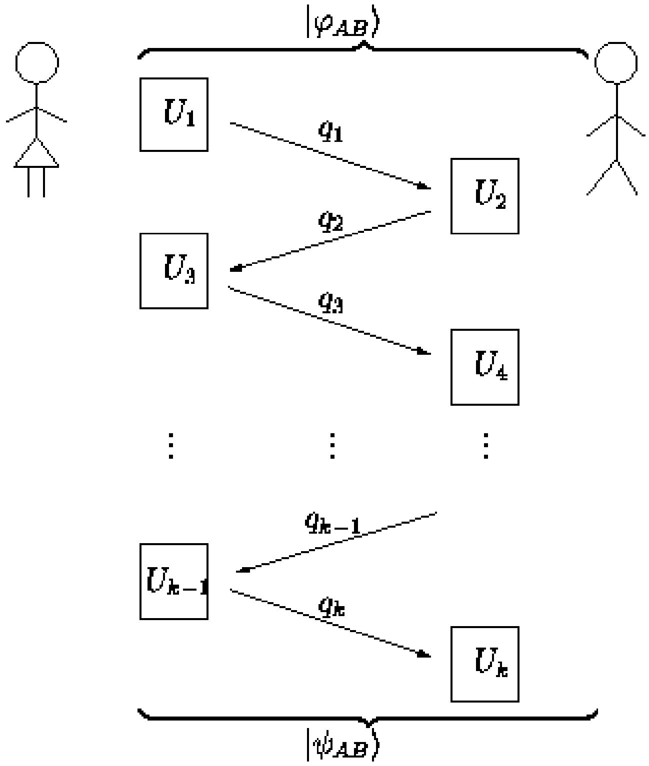


Figure 5.1 A many-round quantum communication protocol. Two parties, Alice and Bob, initially share a joint system $|\phi_{AB}\rangle$. Alice applies a local unitary operator U_1 and then sends q_1 quantum bits to Bob, who performs a local unitary U_2 and then sends q_2 quantum bits to Alice, etc.; in the end, they share system $\psi_{AB}\rangle$. By Theorem 5.2.1, this is equivalent to a protocol in which there is only one round of quantum communication.

ice's initial description of her subsystem (her "reduced density operator") is given by $\varphi_{AC} = \mathrm{Tr}_B\,|\varphi_{ACB}\rangle\langle\varphi_{ACB}|$. If she then sends Bob the system C through a "quantum channel," her new density operator becomes $\varphi_A = \mathrm{Tr}_{CB}\,|\varphi_{ACB}\rangle\langle\varphi_{ACB}|$. Thus, understanding how quantum systems change as a result of quantum communication is equivalent to understanding how a density matrix is related to its partial trace. This connection was in fact the original motivation for studying this problem.

If many rounds of communication are allowed in a quantum communication protocol, it may seem that the analysis should become more complicated (see Figure 5.1). Happily, this turns out not to be the case. In fact, the following result [27] shows that it is enough to consider one-round protocols:

Theorem 5.2.1 *Suppose there exists a bipartite quantum communication protocol that transforms the state $|\varphi_{AB}\rangle$ to the state $|\psi_{AB}\rangle$, requiring a total of q qubits of communication. Then there is a one-round protocol that accomplishes the same transformation $|\varphi_{AB}\rangle \to |\psi_{AB}\rangle$, also requiring q qubits of information.*

Proof The proof involves showing that at any round of the protocol, any communication from Bob to Alice can be replaced by communication from Alice to Bob; it then follows that all communication can be taken to be in one direction. The effect of Bob sending a qubit to Alice is to transform a state $\sum_i \sqrt{\lambda_i}|i_A, i_B\rangle$ to a state $\sum_i \sqrt{\lambda_{i'}}|i'_A, i'_B\rangle$, where the prior and posterior states are written in their Schmidt decompositions. But by the symmetry of the Schmidt decomposition, the swap operator exchanging Alice's and Bob's systems is equivalent to applying some local unitaries $U_A \otimes U_B$ on their joint system. Thus, instead of having Bob send a qubit to Alice, they can apply $U_A \otimes U_B$ and then have Alice send a qubit to Bob (and finally apply some local unitaries $U'_A \otimes U'_B$ to swap Alice and Bob back again) to accomplish the same transformation. $\square$

While the problem of comparing the spectrum of a matrix to that of its partial trace has a natural application to density matrices, it may be applied to other settings as well. For example, given the spectrum of an observable for a certain quantum system, one may wish to ask what the spectrum of that observable may be for a subsystem of the given system. In this context ρ_{AB} is the matrix of the observable, rather than a density matrix.

5.3 Horn's Problem

Horn's problem is the following: Given the spectra of $n \times n$ Hermitian matrices X and Y, what are the possible spectra of $Z = X + Y$? This problem was first seriously attacked by H. Weyl in 1912 [28], but the complete solution has only been achieved recently [29–34]. We shall see that Horn's problem is intimately connected with the problem of relating the spectrum of a matrix to that of its partial trace; much of the mathematical machinery employed to solve Horn's problem can be adapted to the latter problem, and the form of the solution is the same in each case. In this section, we give a brief history of Horn's problem and its solution.

Early attempts at Horn's problem involved finding inequalities that the eigenvalues of X, Y, and Z had to satisfy, in order for $Z = X + Y$. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be the eigenvalues of X, $\beta = (\beta_1, \ldots, \beta_n)$ be the eigenvalues of Y, and $\gamma = (\gamma_1, \ldots, \gamma_n)$ be the eigenvalues of Z. (As usual, we assume the eigenvalues are written in descending order: $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$, etc.) One basic constraint that α, β and γ must satisfy is the *trace condition*

$$\sum_{i=1}^{n} \gamma_i = \sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n} \beta_i. \tag{5.5}$$

Besides this equality condition, all other constraints on the eigenvalues involved linear inequalities among the eigenvalues; in fact, they all had the form

$$\sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j. \tag{5.6}$$

where I, J, and K are all subsets of $\{1, \ldots, n\}$ of the same cardinality r. Such inequalities were systematically analyzed by A. Horn in 1962 [35]. He found conditions on triples of index sets (I, J, K) for which he conjectured that inequalities of the form of Ineq. (5.6) would be necessary and sufficient.

Horn defined sets T^n_r of triples (I, J, K), corresponding to the (conjectured) necessary and sufficient inequalities, inductively as follows. For each positive integer n and $r \leq n$, let

$$U^n_r = \{(I, J, K) | \sum_{i \in I} i + \sum_{j \in J} j = \sum_{k \in K} k + r(r+1)/2\}. \tag{5.7}$$

Then for $r = 1$, let $T^n_1 = U^n_1$. For $r > 1$, let

$$\begin{aligned} T^n_r \quad = \quad & \{(I, J, K) \in U^n_r | \text{ for all } p < r \text{ and all } (F, G, H) \in T^r_p, \\ & \sum_{f \in F} i_f + \sum_{g \in G} j_g \leq \sum_{h \in H} k_h + p(p+1)/2\}. \end{aligned}$$

Horn's conjecture can then be stated:

Conjecture 5.3.1 (Horn) *A triple (α, β, γ) can be the eigenvalues of $n \times n$ Hermitian matrices X, Y, and Z, where $Z = X + Y$, if and only if the trace condition holds, and*

$$\sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j$$

for all $(I, J, K) \in T_r^n$, for all $r < n$.

Horn showed that his conjecture was valid for $n = 3$ and $n = 4$ (the case $n = 2$ was already known), and asserted that his proof could be extended for $n \leq 8$. However, the general case proved elusive. In 1982, B .V. Lidskii [36] announced that he had verified Horn's conjecture, but his proof sketch was very incomplete, and the details have never appeared. The problem was finally solved in the past five years by Klyachko [29, 30], with important contributions from Tao, Totaro, Woodward, and Belkale [31–34]:

Theorem 5.3.2 *Horn's conjecture is true. More generally, for each positive n and N there exists a finite set L and index sets $\{K_l\} \subset \{1, \ldots, N\}$ and $\{J_{il}\} \subset \{1, \ldots, N\}$, where $l \in L$ and $i \in \{1, \ldots, N\}$, such that the following holds: An $n \times n$ Hermitian matrix A can be written as the sum of N Hermitian $n \times n$ matrices with respective spectra $\lambda^1, \lambda^2, \ldots, \lambda^N$ if and only if*

$$\sum_{k \in K_l} (\mathrm{Spec}(A))_k \leq \sum_{i=1}^{N} \sum_{j \in J_{il}} \lambda_j^i \tag{5.8}$$

holds for all $l \in L$.

In the above theorem, we assume that the spectra λ^i are each written in non-increasing order.

The second sentence of Theorem 5.3.2 was actually known to Horn, although he did not know how to generate the sets $\{K_l\}$ and $\{J_{il}\}$. Algorithms for generating these sets are now known. A somewhat unexpected complication that arises is that Horn's list of inequalities is redundant for $n > 5$; as n increases, the number of redundant inequalities grows rapidly. So it is natural to desire a minimal set of inequalities that are necessary and sufficient for (α, β, γ) to be the spectra of Hermitian matrices X, Y, and $X+Y$. This issue has been resolved as well; Knutson and Tao have developed combinatorial gadgets called "honeycombs" that can be used to determine which of Horn's inequalities is redundant.

5.4 An Application to LOCC Protocols

Besides serving as a motivation for the present work, Horn's problem may itself yield insights into problems in quantum information theory. We now present an application demonstrating that it is sufficient to consider protocols of a special type in performing transformations using local operations and classical communication (LOCC.)

First, using Theorem 5.3.2, we derive a theorem for representing a matrix as a convex combination of isospectral matrices.

Theorem 5.4.1 *Suppose a matrix σ can be written as a convex combination of unitary conjugations of a fixed Hermitian matrix ρ:*

$$\sigma = \sum_{i-1}^{N} p_i U_i \rho U_i^\dagger, \tag{5.9}$$

where each U_i is unitary, $p_i \geq 0$ and $\sum_i p_i = 1$. Set $p = (p_1, \ldots, p_N)$, and suppose that q is a probability distribution such that $q \prec p$. Then there exist unitary matrices $\{V_i\}_{i=1}^N$ such that

$$\sigma = \sum_{i=1}^{N} q_i V_i \rho V_i^\dagger. \tag{5.10}$$

Proof Let $\mu = \mathrm{Spec}(\sigma)$ and $\lambda = \mathrm{Spec}(\rho)$, so that $p_i\lambda = \mathrm{Spec}(p_i U_i \rho U_i^\dagger)$. By Theorem 5.3.2 there is a list of inequalities, each of the form

$$\sum_{k \in K} \mu_k \leq \sum_{i=1}^{N} \sum_{J \in J_i} p_i \lambda_j, \tag{5.11}$$

that must be satisfied in order for Equation (5.9) to hold. By the symmetry of interchanging the order of the summands in Equation (5.9), it must be true for each $\pi \in S_N$ that

$$\sum_{k\in K} \mu_k \leq \sum_{i=1}^{N} \sum_{j\in J_i} p_{\pi(i)} \lambda_j. \tag{5.12}$$

Now since $q \prec p$, it follows from Theorem 1.3.5 that there exist coefficients $c_\pi \geq 0$, $\sum_{\pi\in S_N} c_\pi = 1$, such that for all $i \in \{1, \ldots, N\}$,

$$q_i = \sum_{\pi\in S_N} c_\pi p_{\pi(i)}. \tag{5.13}$$

Now we take a convex sum of Inequalities (5.12) over $\pi \in S_N$:

$$\begin{aligned}
\sum_{k\in K} \mu_k &= \sum_{\pi\in S_N} c_\pi \sum_{k\in K} \mu_k && (5.14)\\
&\leq \sum_{\pi\in S_N} c_\pi \sum_{i=1}^{N} \sum_{j\in J_i} p_{\pi(i)} \lambda_j \text{ by Inequalities (5.12)} && (5.15)\\
&= \sum_{i=1}^{N} \sum_{j\in J_i} \lambda_j \sum_{\pi\in S_N} c_\pi p_{\pi(i)} && (5.16)\\
&= \sum_{i=1}^{N} \sum_{j\in J_i} q_i \lambda_j. && (5.17)
\end{aligned}$$

In other words, if an inequality of the form of Inequality (5.11) holds for values p_i, then it also holds when every p_i is replaced by q_i. Applying Theorem 5.3.2, we conclude that there must be unitary matrices $\{V_i\}_{i=1}^{N}$ such that

$$\sigma = \sum_{i=1}^{N} q_i V_i \rho V_i^\dagger. \tag{5.18}$$

□

In particular, we have

Corollary 5.4.2 *Suppose a matrix σ can be written as a convex combination of uni-*

tary conjugations of a fixed Hermitian matrix ρ:

$$\sigma = \sum_{i=1}^{N} p_i U_i \rho U_i^\dagger, \tag{5.19}$$

where each U_i is unitary, $p_i \geq 0$ and $\sum_i p_i = 1$. Then there exist unitary matrices $\{V_i\}_{i=1}^N$ such that

$$\sigma = \frac{1}{N} \sum_{i=1}^{N} V_i \rho V_i^\dagger. \tag{5.20}$$

Proof Set $q = (\frac{1}{N}, \ldots, \frac{1}{N})$ in Theorem 5.4.1. □

In [1], M. Nielsen describes how to transform a quantum state $|\varphi_{AB}\rangle$, jointly held by two parties, into another bipartite quantum state $|\psi_{AB}\rangle$, using only local operations and classical communication; this is possible whenever

$$\mathrm{Spec}(\varphi_A) \prec \mathrm{Spec}(\psi_A). \tag{5.21}$$

It follows easily from Ky Fan's Maximum Principle (see Section 6.1) that Condition (5.21) holds if φ_A can be written as a convex sum

$$\varphi_A = \sum_{i=1}^{N} p_i U_i \psi_A U_i^\dagger \tag{5.22}$$

where each U_i is unitary. Nielsen shows that if $|\phi_{AB}\rangle$ can be tranformed into $|\psi_{AB}\rangle$ via LOCC, then Equation (5.22) holds, by presenting a protocol (using $\log N$ bits of classical communication) that exhibits this representation. In the protocol, one party performs a measurement with N possible outcomes, where p_i is the probability of the ith outcome, to her portion of the joint system. The outcome i is communicated to the other party, who then performs a unitary U_i to his portion of the system. Any such protocol carries out the transformation $|\varphi_{AB}\rangle \rightarrow |\psi_{AB}\rangle$, so Corollary 5.4.2 has the following consequence.

Corollary 5.4.3 *In Nielsen's protocol for transforming quantum states via LOCC, all measurement outcomes may be taken to be equiprobable without increasing the* *number of bits of classical communication required.*

Chapter 6

Variational Principle

We use a variational principle argument to show that inequalities between the eigenvalues of ρ_{AB} and of ρ_A arise whenever a certain Grassmannian intersection is nonempty. We also show explicitly that when $d_A = 2$, these inequalities are sufficient.

6.1 Some Basic Inequalities

In this section we use a simple argument to derive some inequalities that the spectra of ρ_{AB} and ρ_A must satisfy. Although these inequalities will subsumed by our later results, the proof illustrates the strategy behind the general method. We will make use of the following well-known fact from linear algebra [26]:

Theorem 6.1.1 (Ky Fan's Maximum Principle) *Let A be an $n \times n$ Hermitian matrix with spectrum λ, where we assume as usual that the components of λ are in non-increasing order. Then for all $k \in \{1, \ldots, n\}$,*

$$\sum_{j=1}^{k} \lambda_j = \max \sum_{j=1}^{k} \langle x_j, A x_j \rangle \tag{6.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{C}^n$ and the maximum is taken over all orthonormal k-tuples of vectors $\{x_1, \ldots, x_k\}$ in $\mathbb{C}^n$.

Proof Let $v_1, \ldots, v_n$ be an orthonormal eigenbasis for A, ordered so that $\langle v_i, A v_i \rangle = \lambda_i$. For any k, if we choose $\{v_1, \ldots, v_k\}$ as our k-tuple, then $\sum_{j=1}^{k} \lambda_j = \sum_{j=1}^{k} \langle v_j, A v_j \rangle$.

Now let $\{x_1, \ldots, x_k\}$ be any other orthonormal k-tuple. Let V be the span of $\{x_1, \ldots, x_k\}$, let F be the span of $\{v_1, \ldots, v_k\}$, and let $W = V \cap F$. Let $V' = V \cap F^{\perp}$ and let F' be the orthogonal complement of W in F. Suppose W is d-dimensional. Choose an orthonormal basis $\{w_1, \ldots, w_d\}$ for W, an orthonormal basis $\{v'_1, \ldots, v'_{k-d}\}$ for V', and an orthonormal basis $\{f'_1, \ldots, f'_{k-d}\}$ for F. Now since $V' \leq F^{\perp}$, we must have that $\langle v'_i, Av'_i\rangle \leq \lambda_k$ for all i; and since $F' \in F$, it follows that $\langle f'_i, Af'_i\rangle \geq \lambda_k$ for all i. So we have

$$\begin{aligned}
\sum_{j=1}^{k}\langle x_j, Ax_j\rangle &= \sum_{j=1}^{d}\langle w_j, Aw_j\rangle + \sum_{j=1}^{k-d}\langle v'_j, Av'_j\rangle && (6.2)\\
&\leq \sum_{j=1}^{d}\langle w_j, Aw_j\rangle + \sum_{j=1}^{k-d}\langle f'_j, Af'_j\rangle && (6.3)\\
&= \sum_{j=1}^{k}\langle v_j, Av_j\rangle && (6.4)\\
&= \sum_{j=1}^{k}\lambda_j. && (6.5)
\end{aligned}$$

□

We will use the following notation. Let $\mathrm{Gr}_k(A)$ denote the k-dimensional Grassmannian on the vector space A; that is, $\mathrm{Gr}_k(A)$ is the space of all k-dimensional subspaces of A. For $V \leq \mathbb{C}^n$, let P_V denote the projection operator onto the subspace V. Given a vector $v \in \mathbb{C}^d$ and a positive integer n, we define $\Sigma_n(v)$ to be the vector whose components are obtained by summing successive blocks of n components of v:

$$\Sigma_n(v) = (v_1 + \cdots + v_n, v_{n+1} + \cdots + v_{2n}, \ldots, v_{[d/n](n-1)+1} + \cdots + v_d). \tag{6.6}$$

Recall that we denoted the dimensions of system A and B by d_A and d_B, respectively; and that all vectors of matrix spectra are assumed to be with components in non-increasing order. We will use these conventions throughout.

Theorem 6.1.2 *Let λ be the spectrum of ρ_{AB}, and $\tilde{\lambda}$ be the spectrum of its partial* *trace ρ_A. Then for every $k \in \{1, \ldots, d_A\}$, the inequality*

$$\sum_{i=1}^{k} \tilde{\lambda}_i \leq \sum_{i=1}^{d_B k} \lambda_i \tag{6.7}$$

must hold. We may write the d_A inequalities succinctly as the majorization relation

$$\tilde{\lambda} \prec \Sigma_{d_B}(\lambda). \tag{6.8}$$

Proof

$$\begin{aligned} \sum_{i=1}^{k} \tilde{\lambda}_i &= \max_{\{V \in \operatorname{Gr}_k(A)\}} \operatorname{Tr}(\rho_A P_V) \\ &= \max_{\{V \in \operatorname{Gr}_k(A)\}} \operatorname{Tr}(\rho_{AB} P_{V \otimes B}) \\ &\leq \max_{\{V \in \operatorname{Gr}_{kd_B}(A \otimes B)\}} \operatorname{Tr}(\rho_{AB} P_V) \\ &= \sum_{i=1}^{kd_B} \lambda_i, \end{aligned} \tag{6.9}$$

where the first and last equalities follow from Ky Fan's Maximum Principle, the second equality comes from the definition of partial trace, and the inequality follows because the maximum is being taken over a larger set of projection operators than in the previous expression. □

Note the basic idea behind the proof. We expressed the sum of eigenvalues for each matrix in terms of a variational principle on subspaces, and then we looked for an intersection between subspaces in order to relate the variational expressions. This idea will be developed further in the next section.

6.2 General Method

Let A be an $n \times n$ Hermitian matrix with spectrum λ, and let V be a subspace of $\mathbb{C}^n$. Let P_V be orthogonal projection from $\mathbb{C}^n$ onto V. Then $P_V \circ A$ can be regarded as a map from V to V. Define the *Rayleigh trace* of A on V to be the trace of this map:

$$R_A(V) = \operatorname{Tr}(V \hookrightarrow \mathbb{C}^n \xrightarrow{A} \mathbb{C}^n \overset{P_V}{\twoheadrightarrow} V). \tag{6.10}$$

Observe that if B is another Hermitian matrix, then $R_{A+B}(V) = R_A(V) + R_B(V)$. Also note that

$$\max_{V,\ \dim\ V=r} R_A(V) = \sum_{i=1}^{r} \lambda_i \tag{6.11}$$

(this is a restatement of Ky Fan's Maximum Principle), and similarly

$$\min_{V,\ \dim\ V=r} R_A(V) = \sum_{i=n-r+1}^{n} \lambda_i. \tag{6.12}$$

Let A_r denote the r-dimensional vector space spanned by eigenvectors corresponding to the r largest eigenvalues of A (if A is degenerate with $\lambda_r = \lambda_{r+1}$, then choose any such A_r.) Now given a binary sequence π of length n and weight r (sometimes written $\pi \in \binom{n}{r}$), the *Schubert cell* in the r-Grassmannian corresponding to π is defined as

$$S_\pi(A) = \{V \leq \mathbb{C}^n | \dim(V \cap A_i)/(V \cap A_{i-1}) = \pi(i), 1 \leq i \leq n\}, \tag{6.13}$$

where $\pi(i)$ is the ith term in the sequence π. Then $\pi(i) = 1$ for r values of i; label these values $i_1 < i_2 < \cdots i_r$. The following variational principle is due to Hersch and Zwahlen [37].

Theorem 6.2.1

$$\min_{V \in S_\pi(A)} R_A(V) = \sum_i \pi(i)\lambda_i. \tag{6.14}$$

Equality occurs when V is the span of eigenvectors corresponding to the eigenvalues $\lambda_{i_1}, \ldots, \lambda_{i_r}$.

Proof Let $V \in S_\pi(A)$, and choose orthogonal unit vectors $u_1, u_2, \ldots, u_r$ such that $u_k \in V \cap A_{i_k}$. Now A_{i_k} is spanned by eigenvectors of A with eigenvalue greater than or equal to λ_{i_k}, so $\langle Au_k, u_k\rangle \geq \lambda_{i_k}$. It follows that

$$R_A(V) = \sum_{k=1}^{r} \langle Au_k, u_k\rangle \geq \sum_{k=1}^{r} \lambda_{i_k} = \sum_i \pi(i)\lambda_i. \tag{6.15}$$

Now suppose V is the span of eigenvectors corresponding to eigenvalues $\lambda_{i_1}, \lambda_{i_r}$. In this case u_k is an eigenvector of A with eigenvalue λ_{i_k}, so that $R_A(V) = \sum_i \pi(i)\lambda_i$. □

For any $k \leq d_A$, define the map $\phi : \mathrm{Gr}_k(A) \to \mathrm{Gr}_{d_B k}(A \otimes B)$ by $\phi(V) = V \otimes B$. Let $\{y_1, \ldots, y_{d_B}\}$ be an orthonormal basis of $\mathbb{C}^{d_B} = B$, and let I_B denote the identity operator on B. Then for any operator X_A on A, and any $v \in \mathbb{C}^{d_A}$, we have that

$$\begin{aligned}
\sum_{i=1}^{d_B} \langle v \otimes y_i, \frac{1}{d_B} X_A \otimes I_B (v \otimes y_i)\rangle &= \frac{1}{d_B} \sum_{i=1}^{d_B} \langle v, X_A v\rangle \langle y_i, I_B y_i\rangle \\
&= \frac{1}{d_B} \sum_{i=1}^{d_B} \langle v, X_A v\rangle \\
&= \langle v, X_A v\rangle.
\end{aligned}$$

It follows that $R_{X_A}(V) = R_{\frac{1}{d_B} X_A \otimes I_B}(\phi(V))$.

The following theorem was motivated by an analogous argument, due to Johnson [38], Totaro [31], and Helmke and Rosenthal [39], used in the solution of Horn's problem.

Theorem 6.2.2 *Let X_A be an operator on A and Y_{AB} be an operator on $A \otimes B$ such that $X_A = -\mathrm{Tr}_B(Y_{AB})$. Let $\tilde{\lambda}$ be the spectrum of X_A and λ be the spectrum of Y_{AB}. If $\phi(S_\pi(X_A)) \cap S_\sigma(Y_{AB}) \neq \emptyset$, then*

$$\sum_{i=1}^{d_A} \pi(i)\tilde{\lambda}_i + \sum_{i=1}^{d_A d_B} \sigma(i)\lambda_i \leq 0. \tag{6.16}$$

Inequality 6.16 also holds if $\phi(\overline{S_\pi(X_A)}) \cap \overline{S_\sigma(Y_{AB})} \neq \emptyset$.

Proof Let $W \otimes B \in \phi(S_\pi(X_A)) \cap S_\sigma(Y_{AB})$. Then we have

$$
\begin{aligned}
& \textstyle\sum_{i=1}^{d_A} \pi(i)\tilde{\lambda}_i + \sum_{i=1}^{d_A d_B} \sigma(i)\lambda_i && (6.17)\\
= \quad & \min_{V \in S\pi(X_A)} R_{X_A}(V) + \min_{V' \in S_\sigma(Y_{AB})} R_{Y_{AB}}(Y_{AB}) && (6.18)\\
= \quad & \min_{V \in S\pi(X_A)} R_{\frac{1}{d_B} X_A \otimes I_B}(\phi(V)) + \min_{V' \in S_\sigma(Y_{AB})} R_{Y_{AB}}(Y_{AB}) && (6.19)\\
= \quad & \min_{V \in \phi(S\pi(X_A))} R_{\frac{1}{d_B} X_A \otimes I_B}(V) + \min_{V' \in S_\sigma(Y_{AB})} R_{Y_{AB}}(Y_{AB}) && (6.20)\\
\leq \quad & R_{\frac{1}{d_B} X_A \otimes I_B}(W \otimes B) + R_{Y_{AB}}(W \otimes B) && (6.21)\\
= \quad & R_{\frac{1}{d_B} X_A \otimes I_B + Y_{AB}}(W \otimes B) && (6.22)\\
= \quad & \mathrm{Tr}(P_{W \otimes B}(\tfrac{1}{d_B} X_A \otimes I_B + Y_{AB}) P_{W \otimes B}) && (6.23)\\
= \quad & \mathrm{Tr}(P_{W \otimes B}(\tfrac{1}{d_B} X_A \otimes I_B + Y_{AB})) && (6.24)\\
= \quad & \mathrm{Tr}((P_W \otimes I_B)(\tfrac{1}{d_B} X_A \otimes I_B + Y_{AB})) && (6.25)\\
= \quad & \mathrm{Tr}(P_W \, \mathrm{Tr}_B(\tfrac{1}{d_B} X_A \otimes I_B + Y_{AB})) && (6.26)\\
= \quad & \mathrm{Tr}(P_W(X_A + \mathrm{Tr}_B(Y_{AB}))) && (6.27)\\
= \quad & 0. && (6.28)
\end{aligned}
$$

This shows the inequality in the case that $\phi(S_\pi(X_A)) \cap S_\sigma(Y_{AB}) \neq \emptyset$. If $\phi(\overline{S_\pi(X_A)}) \cap \overline{S_\sigma(Y_{AB})} \neq \emptyset$, then Theorem 6.2.1, along with the fact that the Rayleigh trace is continuous, implies that $\min_{V \in \overline{S_\pi(A)}} R_A(V) = \sum_i \pi(i)\lambda_i$, and the argument for the case $\phi(S_\pi(X_A)) \cap S_\sigma(Y_{AB}) \neq \emptyset$ applies equally to this case. □

Theorem 6.2.2 yields inequalities that must be satisfied by the spectra of a matrix and its partial trace, from intersections of Schubert cells. As we will discuss in the next chapter, the closures of the Schubert cells are generators of the homology of the Grassmannian; thus, we can regard the inequalities as coming from nonzero products in cohomology. Determining which of these products are nonzero and translating these nonzero products into the appropriate inequalities will be the focus of the next three chapters.

6.3 Solution for $d_A = 2$

When $d_A = 2$, the relationship between the spectrum of ρ_{AB} and that of $\mathrm{Tr}(\rho_{AB}) = \rho_A$ is particularly simple: the only inequalities restricting the spectra are those given by Theorem 6.1.2. Moreover, this is the only situation for which we are able to explicitly construct matrices demonstrating that the inequalities are sufficient. (If we interpret our problem in terms of quantum communication protocols, the $d_A = 2$ case corresponds to the situation where Alice sends to Bob her entire quantum system except for one qubit.) We give the solution for this case here.

Theorem 6.3.1 *If $d_A = 2$, the inequalities given by Theorem 6.1.2 are sufficient. That is, given a vector $\lambda \in \mathbb{R}^{2d_B}$ and a vector $\tilde{\lambda} \in \mathbb{R}^2$, each with components in non-increasing order, satisfying $\tilde{\lambda} \prec (\sum_{i=1}^{d_B} \lambda_i, \sum_{i=d_B+1}^{2d_B} \lambda_i)$, there exist matrices ρ_{AB} and ρ_A such that the spectrum of ρ_{AB} is λ, the spectrum of ρ_A is $\tilde{\lambda}$, and $\rho_A = \mathrm{Tr}_B(\rho_{AB})$.*

Proof Let $\lambda = (\lambda_{0,0}, \lambda_{0,1}, \ldots, \lambda_{0,d_B-1}, \lambda_{1,0}, \lambda_{1,1}, \ldots, \lambda_{1,d_B-1})$, let $\{|0_A\rangle, |1_A\rangle\}$ and $\{|0_B\rangle, \ldots, |(j-1)_B\rangle$ be orthonormal bases for A and B, respectively, and set

$$\sigma_{AB} = \sum_{i=0}^{1} \sum_{j=0}^{d_B-1} \lambda_{i,j} |i_A\rangle |j_B\rangle \langle i_A| \langle j_B| \tag{6.29}$$

For $t \in [0, 2\pi]$, let

$$\begin{aligned} U(t) = \; & \sum_{i=0}^{1} \sum_{j=0}^{d_B-1} \cos t |i_A\rangle |j_B\rangle \langle i_A| \langle j_B| \\ & + \sum_{j=0}^{d_B-1} \sin t |0_A\rangle |(j-1)_B\rangle \langle 1_A| \langle j_B| \\ & - \sum_{j=0}^{d_B-1} \sin t |1_A\rangle |j_B\rangle \langle 0_A| \langle (j-1)_B|, \end{aligned} \tag{6.30}$$

where the subtraction in the labels of the bra and ket vectors is done modulo d_B. Now $U(t)$ is unitary (in fact, it is real orthogonal) for all t, so the spectrum of $U(t)\sigma_{AB}U(t)^\dagger$

is λ. A direct calculation verifies that

$$\begin{aligned}U(t)\sigma_{AB}U(t)^\dagger = & \sum_{i=0}^{1}\sum_{j=0}^{d_B-1}\lambda_{i,j}\cos^2 t|i_A\rangle|j_B\rangle\langle i_A|\langle j_B| \\ & +\sum_{j=0}^{d_B-1}(\lambda_{1,j}-\lambda_{0,j-1})\sin t\cos t|0_A\rangle|(j-1)_B\rangle\langle 1_A|\langle j_B| \\ & +\sum_{j=0}^{d_B-1}(\lambda_{1,j}-\lambda_{0,j-1})\sin t\cos t|1_A\rangle|j_B\rangle\langle 0_A|\langle (j-1)_B| \\ & +\sum_{j=0}^{d_B-1}\sin^2 t(\lambda_{0,j-1}|0_A\rangle|j_B\rangle\langle 0_A|\langle j_B| + \lambda_{1,j}|1_A\rangle|j_B\rangle\langle 1_A|\langle j_B|),\end{aligned} \tag{6.31}$$

so that

$$\begin{aligned}\mathrm{Tr}_B(U(t)\sigma_{AB}U(t)^\dagger) = & \left(\sum_{j=0}^{d_B-1}\lambda_{0,j}\cos^2 t + \sum_{j=0}^{d_B-1}\lambda_{1,j}\sin^2 t\right)|0_A\rangle\langle 0_A| \\ & +\left(\sum_{j=0}^{d_B-1}\lambda_{1,j}\cos^t + \sum_{j=0}^{d_B-1}\lambda_{0,j}\sin^2 t\right)|1_A\rangle\langle 1_A|.\end{aligned} \tag{6.32}$$

Let $\alpha_1 = \sum_{j=0}^{d_B-1}\lambda_{0,j}$, $\alpha_2 = \sum_{j=0}^{d_B-1}\lambda_{1,j}$. If we let $\rho_{AB}(t) = U(t)\sigma_{AB}U(t)^\dagger$, then the spectrum of the partial trace of $\rho_{AB}(t)$ is $(\alpha_1\cos^2 t + \alpha_2\sin^2 t, \alpha_1\sin^2 t + \alpha_2\cos^2 t)$. By choosing the appropriate value of $t \in [0, 2\pi]$, any convex combination of α_1 and α_2 can be achieved for the eigenvalues of $\mathrm{Tr}_B(\rho_{AB}(t))$. $\square$

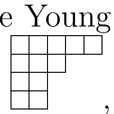
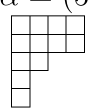
Chapter 7

Schubert Calculus

This chapter describes arithmetic in the cohomology ring of the Grassmannian. It consists of background material and our treatment follows the discussions in [40], [41], and [42].

7.1 Symmetric Polynomials

In this section we give some background on the ring Λ_n of symmetric polynomials in n variables with integer coefficients. A certain class of such polynomials, the Schur polynomials, will be of particular interest, due to its relationship with Grassmannian cohomology. The Schur polynomials (as well as the Grassmannian cohomology classes) are indexed by partitions of integers, so we begin with some terminology relating to partitions.

A *partition* of an integer n is a finite sequence $\alpha = (\alpha_1, \ldots \alpha_l)$ of nonnegative integers, with $n = \sum_i \alpha_i$, arranged in non-increasing order: $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_l \geq 0$. These integers $\alpha_1, \ldots, \alpha_l$ are called the *parts*, and the *length* $\ell(\alpha)$ is the number of nonzero parts. The integer $n = \sum_i \alpha_i$ is the *weight* of the partition, denoted $|\alpha|$. To any partition α we may associate a Young diagram, whose ith row has length α_i. The *conjugate partition* α^* is obtained by interchanging rows and columns in the Young diagram of α. For instance, if $\alpha = (5, 3, 2, 2)$, then the Young diagram of α is  , so the Young diagram of α^* is  and $\alpha^* = (4, 4, 2, 1, 1)$.

Now let Λ_n be the ring of symmetric polynomials with integer coefficients in n variables. There are a number of computationally useful bases for Λ_n. Perhaps the most natural basis is given by the *monomial symmetric functions.* These are functions obtained by starting with a monomial $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and symmetrizing it, to obtain a polynomial

$$m_\alpha = \sum_{\beta \in S_n(\alpha)} x^\beta. \tag{7.1}$$

In this notation, S_n permutes the coefficients of α. Note that the sum is not over all permutations in S_n, but over the image of these permutations; thus, any given monomial appears only once in the sum.

Theorem 7.1.1 *The polynomials m_α, where α ranges over partitions with at most n parts, form a basis over $\mathbb{Z}$ for the ring Λ_n.*

Proof Given a polynomial $P(x_1, \ldots, x_n) = \sum c_\alpha x^\alpha \in \Lambda_n$, let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be the maximal n-tuple (with respect to the lexicographic ordering) such that $c_\alpha \neq 0$. Because $P(x_1, \ldots, x_n)$ is symmetric, α must be a partition. Now $P(x_1, \ldots, x_n) - c_\alpha m_\alpha$ is also a symmetric polynomial, but one whose leading monomial is smaller than x^α with respect to the lexicographic ordering. Because $\alpha_i \geq 0$, the lexicographic ordering is a well-ordering, so it follows by induction that $P(x_1, \ldots, x_n)$ can be written as an integer combination of terms m_α.

Now suppose $\sum c_\alpha m_\alpha = 0$. Again, let α be the maximal n-tuple with respect to the lexicographic ordering such that $c_\alpha \neq 0$. Then the coefficient of x^α in the polynomial $\sum c_\alpha m_\alpha$ is c_α, a contradiction. □

We will make reference to the following two classes of symmetric polynomials. The *elementary symmetric polynomials* are a subset of the monomial symmetric functions, corresponding to partitions such that all parts are equal to one:

$$e_k = \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}, \tag{7.2}$$

for $1 \leq k \leq n$. The *complete symmetric polynomials* are

$$h_k = \sum_{1 \leq i_i \cdots \leq i_k \leq n} x_{i_1} \cdots x_{i_k}, \tag{7.3}$$

for $1 \leq k \leq n$. (If $k = 0$, then set $e_0 = h_0 = 1$.) We label products of elementary symmetric polynomials, as well as products of complete symmetric polynomials, by partitions α: $e_\alpha = e_{\alpha_1} \cdots e_{\alpha_l}$, and $h_\alpha = h_{\alpha_1} \cdots h_{\alpha_l}$.

Both the elementary symmetric polynomials and complete symmetric polynomials are important objects in the study of the ring Λ_n. The *fundamental theorem of symmetric polynomials* states that every symmetric polynomial can be written as a polynomial in the elementary symmetric polynomials [43]; in other words, the polynomials e_α, where α ranges through partitions with parts less than or equal to n, form a basis over $\mathbb{Z}$ of the ring Λ_n. We will make use of the following relationship between the polynomials e_k and h_k.

Proposition 7.1.2 *Let $\omega : \Lambda_n \to L_n$ be the ring homomorphism defined by $\omega(e_k) = h_k$. Then ω is an involution.*

Proof The formal generating series

$$e(t) = \sum_{k \geq 0} e_k t^k = \prod_{i=1}^{n} (1 + tx_i) \tag{7.4}$$

and

$$h(t) = \sum_{k \geq 0} h_k t^k = \prod_{i=1}^{n} (1 - tx_i)^{-1} \tag{7.5}$$

satisfy the relation $e(t)h(-t) = 1$, so that

$$\sum_{i+j=k} (-1)^i e_i h_j = 0 \tag{7.6}$$

for all integers $k > 0$. Applying ω to the above equation, we have that

$$\sum_{i+j=k} (-1)^i h_i \omega(h_j) = 0 \tag{7.7}$$

for all $k > 0$, so it follows by induction that $\omega(h_j) = e_j$. □

It follows from the fundamental theorem of elementary symmetric polynomials and Proposition 7.1.2 that the polynomials h_α form a $\mathbb{Z}$-basis of Λ_n.

We now describe another basis for the ring Λ_n: the *Schur polynomials*, which will be a greater focus of our study. In order to do so, we make some observations about the ring of antisymmetric polynomials in n variables. These polynomials have a basis obtained from antisymmetrizing monomials: if γ is an n-tuple of natural numbers, then let

$$a_\gamma = \sum_{w \in S_n} \varepsilon(w) x^{w(\gamma)}, \tag{7.8}$$

where $\varepsilon(w)$ is the sign of the permutation w. Note that if γ has two equal components, then $a_\gamma = 0$. Thus, we restrict our attention to the case where γ is a strictly decreasing partition. Then γ has the form $\gamma = \alpha + \delta$, where α is a partition and $\delta = (n-1, n-2, \ldots, 1, 0)$. An argument similar to the proof of Theorem 7.1.1 shows that the polynomials $a_{\alpha+\delta}$, where α ranges over partitions with at most n parts, form a basis for the ring of antisymmetric polynomials with integer coefficients.

Next, note that every antisymmetric polynomial must be divisible by $(x_i - x_j)$ for all $i \neq j$, and so must be divisible by the Vandermonde determinant $\det(x_i^{n-j})_{1 \leq i,j \leq n} = \prod_{1 \leq i < j \leq n} (x_i - x_j)$. It is not hard to see that multiplying a symmetric polynomial by the Vandermonde determinant produces an antisymmetric polynomial, and that dividing an antisymmetric polynomial by the Vandermonde determinant yields a symmetric polynomial. Thus, multiplication by the Vandermonde determinant gives an isomorphism between symmetric and antisymmetric polynomials. The *Schur polynomials* are obtained by dividing the polynomials a_γ by the Vandermonde determinant (which is the same as a_δ):

$$s_\alpha = \frac{a_{\alpha+\delta}}{a_\delta} = \frac{\det(x_i^{\alpha_j+n-j})_{1\le i,j\le n}}{\det(x_i^{n-j})_{1\le i,j\le n}}. \tag{7.9}$$

By the isomorphism between symmetric and antisymmetric polynomials, we have proven the following theorem.

Theorem 7.1.3 *The Schur polynomials s_α, as α ranges over all partitions with at most n parts, form a basis over $\mathbb{Z}$ of the ring Λ_n.*

Given a partition α and integer k, let $\alpha \otimes k$ denote the set of partitions obtained by adding k boxes to (the Young diagram of) α, at most one box per column. Let $\alpha \otimes 1^k$ denote the set of partitions obtained by adding k boxes to α, at most one box per row.

Theorem 7.1.4 (Pieri formulas) *With the above notation,*

$$s_\alpha e_k = \sum_{\beta\in\alpha\otimes 1^k} s_\beta, \tag{7.10}$$

and

$$s_\alpha h_k = \sum_{\beta\in\alpha\otimes k} s_\beta. \tag{7.11}$$

Proof We have

$$\frac{a_{\alpha+\delta}e_k}{a_\delta} = \frac{1}{a_\delta}\sum_{w\in S_n}\sum_{i_1<\cdots<i_k} \varepsilon(w)x^{w(\alpha+\delta)}x_{w(i_1)}\cdots x_{w(i_k)} \tag{7.12}$$

$$= \frac{1}{a_\delta}\sum_{\beta\in\{0,1\}^n} a_{\alpha+\beta+\delta} \tag{7.13}$$

$$= \sum_{\beta\in\alpha\otimes 1^k} s_\beta, \tag{7.14}$$

where the last equality follows because $a_{\alpha+\beta+\delta} = 0$ unless $\alpha+\beta$ is a partition. A similar expansion shows that

$$\frac{a_{\alpha+\delta}h_k}{a_\delta} = \frac{1}{a_\delta}\sum_{|\beta|=k} a_{\alpha+\beta+\delta}. \tag{7.15}$$

We need to show that the right-hand sum of Equation 7.15 is equal to a sum over partitions obtained by adding at most one box in any column of α. If $\alpha+\beta$ is a partition that differs from α by two or more boxes in the same column, then there must be some integer i such that $\beta_{i+1} > \alpha_i - \alpha_{i+1}$ (and conversely). In this case let η be a sequence defined as follows: $\eta_i = \beta_{i+1} = (\alpha_i - \alpha_{i+1} + 1)$, $\eta_{i+1} = \beta_i + (\alpha_i = \alpha_{i+1} + 1)$, and $\eta_j = \beta_j$ for $j \neq i, i+1$. (Note that $\eta_{i+1} > \alpha_i - \alpha_{i+1}$ iff $\beta_{i+1} > \alpha_i - \alpha_{i+1}$.) Then the n-tuple $a_{\alpha+\beta+\delta}$ differs by a transposition from $a_{\alpha+\eta+\delta}$, so $a_{\alpha+\beta+\delta} = -a_{\alpha+\eta+\delta}$. After cancelling these terms in the sum, we obtain the desired result. $\square$

Theorem 7.1.5 (Jacobi-Trudi formula) *Let α be a partition with at most n parts. Then*

$$s_\alpha = \det(h_{\alpha_i - i + j})_{1\le i,j\le n}. \tag{7.16}$$

Proof Let l be the length of α. Because $h_0 = 1$, $\det(h_{\alpha_i-i+j})_{1\le i,j\le n} = \det(h_{\alpha_i-i+j})_{1\le i,j\le l}$. Expand $\det(h_{\alpha_i-i+j})_{1\le i,j\le l}$ along the last column, using induction on l:

$$\det(h_{\alpha_i-i+j})_{1\le i,j\le l} = \sum_{i=1}^{l}(-1)^{l-i}s_{\lambda_1,\ldots,\lambda_{i-1},\lambda_{i+1}-1,\ldots,\lambda_l-1} \times h_{\lambda_i+l-i}. \tag{7.17}$$

Now it follows from Theorem 7.1.4 that the ith term of the above sum may be written as

$$\sum_{\beta\in J_i} s_\beta + \sum_{\beta\in J_{i+1}} s_\beta, \tag{7.18}$$

where J_i is the set of partitions β having the same weight as α, satisfying the conditions $\alpha_j \le \beta_j \le \alpha_{j-1}$ for $j < i$, and $\alpha_{j+1} - 1 \le \beta_j \le \alpha_j - 1$ for $j \ge i$. Therefore, the right hand sum of Equation 7.17 telescopes to give us the desired formula. $\square$

7.2 Grassmannians

Let E be an n-dimensional complex vector space. Recall that the *Grassmannian* $\mathrm{Gr}_k(E)$ is the set of k-dimensional vector subspaces of E. We shall also use the notation $\mathrm{Gr}(k, n)$ to denote the set of k-dimensional subspaces of an n-dimensional complex vector space.

Given $V \in \mathrm{Gr}_k(E)$, let $v_1, \dots, v_k$ be a basis of V. Then we may represent V by a $k \times n$ matrix whose row vectors are the vectors v_i. Obviously this representation is not unique; given two $k \times n$ matrices A and B of rank k, they represent the same element of $\mathrm{Gr}_k(E)$ if and only if $A = gB$ for some $g \in \mathrm{GL}_k$.

For any $I = \{i_1, \dots, i_k\}$ a subset of $\{1, \dots, n\}$ with cardinality k, define U_I to be set of all $V \in \mathrm{Gr}_k(E)$ such that there exists a matrix representative A for V whose Ith $k \times k$ minor is nonsingular. Note that if this is true for one matrix representative of V, it is true for any representative of V. Any $V \in U_I$ can be uniquely represented by a matrix V^I such that the Ith $k \times k$ minor is the identity matrix. For example, if $n = 7$, $k = 3$, and $I = \{1, 2, 3\}$, then any $V \in U_I$ has a unique representation by a matrix of the form

$$\begin{pmatrix} 1 & 0 & 0 & * & * & * & * \\ 0 & 1 & 0 & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * \end{pmatrix}, \tag{7.19}$$

where the stars denote arbitrary entries. And conversely, any matrix of this form represents a $V \in \mathrm{Gr}_k(E)$, so these representations give us a bijection

$$\varphi_I : U_I \longrightarrow \mathbb{C}^{k(n-k)} \tag{7.20}$$

for each I. Obviously, any $V \in \mathrm{Gr}_k(E)$ is in U_I for some I, and $\varphi_I(U_I \cap U_{I'})$ is open in $\mathbb{C}^{k(n-k)}$ for all I, I'. Moreover, if $V^I_{I'}$ denotes the I'th $k \times k$ minor of V^I, then

$$V^{I'} = (V^I_{I'})^{-1} V^I. \tag{7.21}$$

It follows that $\varphi_I \circ \varphi_{I'}^{-1}$ is holomorphic on $U_I \cap U_{I'}$, so the maps φ_I define a complex manifold structure on the Grassmannian.

If V is a k-dimensional subspace of E, then $\wedge^k V$ is a line in $\wedge^k E$, giving us a map

$$\phi : \mathrm{Gr}_k(E) \to \mathbb{P}(\wedge^k E). \tag{7.22}$$

Let $A = (a_{ij})$ be a $k \times n$ matrix representing V, so that V is the span of the rows of A. Then a set of homogeneous coordinates in $\phi(V)$ is given by the determinants of the $k \times k$ minors of this matrix: if I is a subset of $\{1, \ldots, n\}$ of cardinality k, then define the coordinate

$$x_I = \det A_I, \tag{7.23}$$

where A_I denotes the Ith $k \times k$ minor of A. These coordinates are known as *Plücker coordinates*, and the map ϕ is called the *Plücker embedding*. It can be shown [40] that the Plücker embedding is indeed an embedding of the Grassmannian $\mathrm{Gr}_k(E)$ into the projective space $\mathbb{P}(\wedge^k E)$, and that the homogeneous coordinates are the solutions of a set of (quadratic) polynomial equations, giving $\mathrm{Gr}_k(E)$ the structure of a projective algebraic variety.

7.3 Schubert Varieties of Grassmannians

Define a *(complete) flag* $F_\bullet$ on E to be a nested sequence

$$F_\bullet : 0 = F_0 \subset F_1 \subset F_2 \subset \ldots \subset F_n = E \tag{7.24}$$

with $\dim(F_i) = i$. For any such flag, we obtain a cell decomposition of $\mathrm{Gr}_k(E)$, as follows. Let α be a partition contained in a $k \times (n-k)$ rectangle (this means that α has length at most k and that all parts are less than or equal to $n-k$). To each such α we associate the *Schubert cell*

$$\Omega_\alpha = \{V \in \mathrm{Gr}_k(E) | \dim(V \cap F_j) = i \text{ if } n-k+i-\alpha_i \leq j \leq n-k+i-\alpha_{i+1}\}. \tag{7.25}$$

and the *Schubert variety*

$$X_\alpha = \{V \in \mathrm{Gr}_k(E) | \dim(V \cap F_{n-k+i-\alpha_i}) \geq i\}. \tag{7.26}$$

This definition of Schubert cell differs from the one given in the previous chapter, but the two definitions refer to the same object, as we now show. Given any binary string π of length n and weight k, associate to it a partition α_π as follows. Let a_i be the number of zeroes that appear in π before the ith one. Then let $\alpha_\pi = (a_k, a_{k-1}, \ldots, a_1)$. For instance if $\pi = 010011$, then $\alpha_\pi = (3, 3, 1)$. It is not hard to see that this gives a one-to-one correspondence between binary strings of length n and weight k, and partitions contained in a $k \times (n-k)$ rectangle, and that $S_\pi = \Omega_{\alpha_\pi}$.

When we wish to emphasize the flag, we write $\Omega_\alpha(F_\bullet)$ and $X_\alpha(F_\bullet)$ for Ω_α and X_α, respectively. Schubert varieties corresponding to partitions with only one nonzero part are called *special Schubert varieties*

$$X_l = \{V \in \mathrm{Gr}_k(E) | V \cap F_{n-k+1-l} \neq 0\}. \tag{7.27}$$

We now show that Schubert varieties are indeed algebraic varieties. Note that $\dim(V \cap F_i) \geq j$ if and only if the rank of the map

$$V \hookrightarrow \mathbb{C}^n \twoheadrightarrow \mathbb{C}^n / F_i \tag{7.28}$$

is less than or equal to $k-j$. This means that, in local coordinates, all minors of order $k-j+1$ of the matrix of this map must have vanishing determinant, a requirement governed by polynomial equations. The Schubert varieties are therefore algebraic subvarieties of $\mathrm{Gr}_k(E)$.

In what follows, let $f_1, \ldots, f_n$ be a basis respecting the flag $F_\bullet$ of E; in other words, these vectors are such that $F_i = \langle f_1, \ldots f_i \rangle$ for all i.

Let α be a partition contained in a $k \times (n-k)$ rectangle. In terms of the basis $\langle f_1, \ldots, f_n \rangle$, any $V \in \Omega_\alpha$ can be expressed in terms of a unique basis, consisting of the rows of a $k \times (n-k)$ matrix with the following properties: the ith row contains a 1 in the $(n-k+i-\alpha_i)$th position, and zeros in all subsequent positions; and all other entries in the $(n-k+i-\alpha_i)$th column are zero. For instance, if $n=7, k=3$, and $\alpha=(3,2,1)$, such matrices are of the form

$$\begin{pmatrix} * & 1 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & * & 1 & 0 & 0 & 0 \\ * & 0 & * & 0 & * & 1 & 0 \end{pmatrix}, \tag{7.29}$$

where the stars denote arbitrary entries. Clearly any such matrix corresponds to a $V \in \Omega_\alpha$, so we have a homeomorphism of Ω_α with $\mathbb{C}^{k(n-k)-|\alpha|}$. In general, V can written (not uniquely) as the span of the rows of any $k \times (n-k)$ matrix with a nonzero entry in the $(n-k+i-\alpha_i)$th position of the ith row, and zeros afterwards. Using our example $n=7, k=3$, and $\alpha=(3,2,1)$, such matrices can be written as

$$\begin{pmatrix} * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & 0 & 0 & 0 \\ * & * & * & * & * & * & 0 \end{pmatrix}, \tag{7.30}$$

where the last star in each row represents any nonzero term, and all other stars represent arbitrary terms. From this representation, we see that if $\alpha \subset \beta$ (this means that the Young diagram of α is contained in the diagram of β), then $\Omega_\beta \subset \overline{\Omega_\alpha}$.

The following theorem tells how to determine the incidence of Schubert varieties.

Theorem 7.3.1 *For all partitions $\alpha \subset k \times (n-k)$,*

(a) $X_\alpha = \overline{\Omega_\alpha} = \coprod_{\beta \supset \alpha} \Omega_\beta$, and

(b) $X_\beta \subset X_\alpha$ if and only if $\alpha \subset \beta$.

Proof For any $V \in \mathrm{Gr}_k(E)$, consider the dimensions of the successive intersections $V \cap F_i$. If $i=0$, this dimension is zero, while if $i=n$, this dimension is k; furthermore, as we go from i to $i+1$, the dimension of intersection cannot increase by more than 1. So, there must exist k values of i for which the dimension increases; these determine a partition $\alpha \subset k \times (n-k)$ such that the ith dimension increase occurs at $n-k+i-\alpha_i$. It follows that

$$\mathrm{Gr}_k(E) = \coprod_{\alpha \subset k \times (n-k)} \Omega_\alpha. \tag{7.31}$$

Now if $\dim(V \cap F_{n-k+i-\alpha_i}) \geq i$, then the first i increases in dimension must have occurred before $n-k+i-\alpha_i$, so this number must be greater than or equal to $n-k+i-\beta_i$. We conclude that

$$X_\alpha = \coprod_{\beta \supset \alpha} \Omega_\beta. \tag{7.32}$$

Now since X_α is closed, we have that

$$\begin{aligned} X_\alpha &= \overline{X_\alpha} & (7.33)\\ &= \overline{\coprod_{\beta \supset \alpha} \Omega_\beta} & (7.34)\\ &= \coprod_{\beta \supset \alpha} \overline{\Omega_\beta} & (7.35)\\ &= \overline{\Omega_\alpha}. & (7.36) \end{aligned}$$

Finally, it follows from Equation 7.31 and 7.32, and the fact that the Schubert cells are nonempty, that $X_\beta \subset X_\alpha$ if and only if $\alpha \subset \beta$. □

We have thus shown that the Schubert cells Ω_α form a cellular decomposition of the Grassmannian. Therefore, the fundamental classes of their closures are a basis of the integral cohomology of $\mathrm{Gr}_k(E)$. (Because all cells are of even real dimension, the integral cohomology is torsion-free.) For any Schubert variety X_α, let $\sigma_\alpha = [X_\alpha]$ denote its class in cohomology, called a *Schubert class*. The results of this section then imply the following theorem.

Theorem 7.3.2 *The integral cohomology of the Grassmannian $Gr_k(E)$ has a basis given by the Schubert classes σ_α, where α ranges over all partitions contained in a*

$k \times (n-k)$ rectangle:

$$H^*(Gr_k(E)) = \bigoplus_{\alpha \subset k \times (n-k)} \mathbb{Z}\sigma_\alpha. \tag{7.37}$$

The Schubert class σ_α is an element of $H^{2|\alpha|}(Gr_k(E))$.

7.4 Intersections of Varieties

Let us now determine when two Schubert varieties must intersect. Given a flag $F_\bullet$, let $\tilde{F}_\bullet$ be the opposite flag to $F_\bullet$. That is, if $\{f_1, \ldots, f_n\}$ is a basis for E such that $F_k = \langle f_1, \ldots, f_k \rangle$, then $\tilde{F}_k = \langle f_{n-k+1}, \ldots, f_n \rangle$. For any partition α with at most k rows and $n-k$ columns, let $\Omega_a = \Omega_\alpha(F_\bullet)$ and let $\tilde{\Omega}_\alpha = \Omega_\alpha(\tilde{F}_\bullet)$. Because $GL(E)$ acts transitively on the flags, Ω_α and $\tilde{\Omega}_\alpha$ have the same fundamental class, denoted σ_α.

We have seen that any element of Ω_α can be written as the span of the rows of a unique $k \times (n-k)$ matrix of the form

$$\begin{pmatrix} * & \ldots & * & 1 & 0 & \ldots & 0 & 0 & \ldots & \ldots & 0 & \ldots & \ldots & \ldots \\ * & \ldots & * & 0 & * & \ldots & * & 1 & 0 & \ldots & 0 & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ * & \ldots & * & 0 & * & \ldots & * & 0 & * & * & 1 & 0 & \ldots & 0 \end{pmatrix}, \tag{7.38}$$

where the ith row has a 1 in the $(n-k+i-\alpha_i)$th position. Similarly, each element of $\tilde{\Omega}_\beta$ can be written in terms of a basis whose elements are the rows of a unique $k \times (n-k)$ matrix of the form

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & * & * & 0 & * & \ldots & * & 0 & * & \ldots & * \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & 0 & \ldots & 0 & 1 & * & \ldots & * & 0 & * & \ldots & * \\ \ldots & \ldots & \ldots & 0 & \ldots & \ldots & 0 & 0 & \ldots & 0 & 1 & * & \ldots & * \end{pmatrix}, \tag{7.39}$$

where the i row has a 1 in position $\beta_{n-k-i+1} + i$.

If $\Omega_\alpha \cap \tilde{\Omega}_\beta \neq \emptyset$, then there must be a k-plane W such that each of the two above matrices determines a basis for W. Now, the first row of the first matrix cannot be a linear combination of rows of the second unless $\beta_{n-k}+1 \leq n+1-\alpha_1 \Longrightarrow \alpha_1+\beta_{n-k} \leq n$. In general, in order for the i row of the first matrix to be a linear combination of rows of the second matrix, but not a linear combination of the first $i-1$ rows of the second matrix, we must have that $\alpha_i + \beta_{n-k-i+1} \leq n$.

For any partition α contained in an $k \times (n-k)$ rectangle, define $\hat{\alpha}$ to be the complementary partition of α in the rectangle: that is, $\hat{\alpha}_i = n - \alpha_{n-k-i+1}$. (If the Young diagram of $\hat{\alpha}$ is turned upside down, it fits perfectly with the diagram of α to form a $k \times (n-k)$ rectangle.) The argument of the previous paragraph shows that $\Omega_\alpha \cap \tilde{\Omega}_\beta = \emptyset$ unless $\beta \subset \hat{\alpha}$. We now have

Theorem 7.4.1 *Suppose α and β are two partitions with at most k rows and $n-k$ columns, and that $|\alpha| + |\beta| = k(n-k)$. Then the cup product in cohomology of the fundamental classes corresponding to α and β is zero unless $\beta = \hat{\alpha}$, in which case it is one; that is,*

$$\sigma_\alpha \cup \sigma_\beta = \delta_{\beta,\hat{\alpha}}. \tag{7.40}$$

The classes σ_α and $\sigma_{\hat{\alpha}}$ are therefore said to be dual.

Proof We have seen that $\Omega_\alpha \cap \tilde{\Omega}_\beta = \emptyset$ unless $\alpha_i + \beta_{n-k+i-1} \leq n$ for all i. Since $|\alpha| + |\beta| = k(n-k)$, we must have equality hold in all these inequalities in order for them to be simultaneously satisfied, and so $\Omega_\alpha \cap \tilde{\Omega}_\beta = \emptyset$ unless $\beta = \hat{\alpha}$. It follows that if $\beta \neq \hat{\alpha}$, then the intersection of Schubert varieties $X_\alpha \cap \tilde{X}_\beta = \emptyset$, so $\sigma_\alpha \cup \sigma_b = 0$. On the other hand, if $\beta = \hat{\alpha}$, then $X_\alpha \cap \tilde{X}_\beta = \Omega_\alpha \cap \tilde{\Omega}_\beta$. The above parametrizations of Ω_α and $\tilde{\Omega}_\beta$ in terms of matrices show that Ω_α intersects $\tilde{\Omega}_\beta$ in exactly one point, determined by the basis vectors corresponding to the positions of the 1's in both of these matrices. Now the stars in the matrices correspond to local coordinates of Ω_α and Ω_β; taking all the stars together yields coordinates for a neighborhood of the intersection in the Grassmannian. The intersection is obtained at the point where all coordinates are equal to zero, so it follows that the intersection of Ω_α and Ω_β is transverse at that point. Therefore, $\sigma_\alpha \cup \sigma_{\hat{\alpha}} = 1$. $\square$

For an integer l between 1 and $n-k$, let σ_l denote the Schubert class corresponding to the special Schubert variety X_l. Then the Pieri rule holds for Schubert classes:

Theorem 7.4.2 (Pieri rule for Schubert classes) *Let a be a partition contained in an $k \times (n-k)$ rectangle, and let l be an integer between 1 and $n-k$. Then*

$$\sigma_\alpha \cup \sigma_l = \sum_{\nu \subset k \times (n-k), \nu \in \lambda \otimes k} \sigma_\nu. \tag{7.41}$$

Proof Since $\sigma_\alpha \in H^{2|\alpha|}$ and $\sigma_l \in H^{2l}$, their product $\sigma_\alpha \cup \sigma_l \in H^{2|\alpha|+2l}$. So we may write

$$\sigma_\alpha \cup \sigma_l = \sum_{\nu \subset k \times (n-k), |\nu| = |\alpha| + l} c_\nu \sigma_\nu \tag{7.42}$$

for some constants c_ν. But Theorem 7.4.1 then implies that

$$c_\nu = (\sigma_\alpha \cup \sigma_l) \cup \sigma_{\hat{\nu}}. \tag{7.43}$$

Thus, we must show that both sides of Equation 7.41 have the same intersection number with all classes σ_β, where $\beta = k(n-k) - |\alpha| - l$. If the diagram of α is put in the top left corner of a $k \times (n-k)$ rectangle, and the diagram of β is turned upside down and put in the bottom right corner of this rectangle, then the formula says that $\sigma_\beta \cup \sigma_\alpha \cup \sigma_l = 1$ when the diagrams do not overlap and none of the boxes of the rectangle that are in neither diagram are in the same column; and that $\sigma_\beta \cup \sigma_\alpha \cup \sigma_l = 0$ otherwise. The asserted condition for $\sigma_\beta \cup \sigma_\alpha \cup \sigma_l = 1$ is then equivalent to the inequalities

$$n - k - \alpha_k \geq \beta_1 \geq n - k - \alpha_{k-1} \geq \beta_2 \geq \cdots \geq n - k - \alpha_1 \geq \beta_k \geq 0. \tag{7.44}$$

Now define the sets

$$\begin{aligned} A_i &= F_{n-k+i-\alpha_i}, & (7.45)\\ B_i &= \tilde{F}_{n-k+i-\beta_i}, & (7.46)\\ C_i &= A_i \cap B_{k+1-i}. & (7.47) \end{aligned}$$

(Let $A_0 = B_0 = 0$.) We make use of the following lemma.

Lemma 7.4.3 *Let C be the subspace of E spanned by the spaces $C_1, \dots, C_k$. Then*

(a) $C = \bigcap_{i=0}^{k}(A_i + B_{k-i})$.

(b) $\sum_{i=1}^{k} \dim(C_i) = k + l$.

(c) The space C is a direct sum of subspaces C_i, each nonempty, if and only if Inequalities 7.44 hold.

(d) If $V \in Gr_k(E)$ is in $\Omega_\alpha \cap \tilde{\Omega}_\beta$, then $V \in C$. Furthermore, if the subspaces $C_1, \dots, C_k$ are linearly independent, then $\dim(V \cap C_i) = 1$ *for all i, and $V = V \cap C_1 \oplus \cdots \oplus V \cap C_k$.*

Suppose at least one of the Inequalities 7.44 does not hold. By the lemma we have that C is not a direct sum of the C_i, and so $\dim C \leq k + l - 1$. Therefore a generic subspace L of dimension $n-k+1-l$ will not intersect C except at the origin. Now it follows from the lemma that if $V \in \Omega_\alpha \cap \tilde{\Omega}_\beta$, then $V \notin \Omega_k(L)$, so $\Omega_\alpha \cap \tilde{\Omega}_\beta \cap \Omega_k(L) = \emptyset$.

Now suppose that Inequalities 7.44 all hold. Then $C = \bigoplus C_i$, and a generic L intersects C in a line spanned by a vector v, where we may write $v = u_1 + \ldots + u_k$, u_i a nonzero vector in C_i. Now the conditions that V intersects L in at least a line, and that $V \subset C$, imply that $v \in V$. But $V = \bigoplus V \cap C_i$ so each $u_i \in V$; thus, V is the subspace spanned by $u_1, \dots, u_k$. So the three Schubert varieties intersect in a single point; by locally identifying the Schubert cells with affine spaces, we see that the intersection is transversal. □

Proof of Lemma 7.4.3 First, note that C_i is spanned by the vectors f_j such that

$$i + \beta_{k+1-i} \leq j \leq n - k + i - \alpha_i. \tag{7.48}$$

(a) It suffices to show that the two expressions contain the same basis vectors f_p of the flag (and dual flag). If $f_p \in C$, then it is in some C_j, so

$$j + \beta_{k+1-j} \leq p \leq n - k + j - \alpha_j \text{ for some } 1 \leq j \leq k; \tag{7.49}$$

while $f_p \in \bigcap_{i=0}^{k}(A_i + B_{k-i})$ means that for all $0 \le i \le k$,

$$p \le n - k + i - \alpha_i \text{ or } p > i + \beta_{k-i}, \tag{7.50}$$

where we set $\alpha_0 = \beta_0 = n - k$. Now suppose $f_p \in C_j$. For $i < j$, $i + \beta_{k-i} < j + \beta_{k+1-j} \le p$, while for $i \ge j$, $p \le n-k+j-\alpha_j \le n+i-\alpha_i$, so $f_p \in \bigcap_{i=0}^{k}(A_i+B_{k-i})$. On the other hand, suppose $f_p \in \bigcap_{i=0}^{k}(A_i + B_{k-i})$. Find the smallest j such that $p \le n - k + j - \alpha_j$. Then since j is smallest, $p > (j-1) + \beta_{k-(j-1)}$, so $f_p \in C_j \subset C$.

(b) $\sum_{i=1}^{k} \dim(C_i) = \sum_{i=1}^{k}(n-k+1-\alpha_i-\beta_{k+1-i}) = k+l$, since $\sum_{i=1}^{k}(\alpha_i+\beta_{k+1-i}) = k(n-k) - l$.

(c) If the inequalities hold, then the vectors f_j spanning each C_i (given by Inequalities 7.48) are distinct.

(d) It suffices (by part (a)) to show that $V \subset A_i + B_{k-i}$ for all i. If $A_i \cap B_{k-i} \neq 0$, then $A_i = B_{k-i} = E$, so $V \subset A_i + B_{k-i}$ in this case. Next supppose that $A_i + B_{k-i} = 0$. We know that $\dim(V \cap A_i) \ge i$ and $\dim(V \cap B_{k-i}) \ge r - i$. But $\dim V = k$, so V is the direct sum of $V \cap A_i$ and $V \cap B_{k-i}$, so indeed we have $V \subset A_i + B_{k-i}$.

Now $\dim(V \cap A_i) \ge i$ and $\dim(V \cap B_{k+1-i} \ge k + 1 - i$, so $\dim(V \cap C_i) \ge 1$. If the C_i are linearly independent, then $\bigoplus(V \cap C_i) \subset V$, but the dimension of $\bigoplus(V \cap C_i) \ge k$, so it follows that $V = \bigoplus(V \cap C_i)$ and each $V \cap C_i$ has dimension exactly one. □

Because the Schubert classes in cohomology satisfy the Pieri rule, we have the following result.

Corollary 7.4.4 *The map $\Lambda_k \longrightarrow H^*(Gr_k(E))$, which sends the Schur function s_α to the Schubert class σ_α if α is a partition contained in a $k \times (n-k)$ rectangle, and sends s_α to zero otherwise, is a surjective ring homomorphism.*

Chapter 8

Computing ϕ^*

By Theorem 6.2.2, we can obtain inequalities relating an operator ρ_{AB} and its partial trace ρ_A whenever there is a non-empty intersection of the Schubert variety $X_\beta(F)$ with $\phi(X_\alpha(F'))$, where F and F' are the flags determined by eigenbases of ρ_{AB} and ρ_A, respectively. The condition that there must be a nonzero intersection corresponds cohomologically to there being nonzero product of the Schubert classes, $\sigma_\alpha \cup \phi^*(\sigma_\beta) \neq 0$, where $\phi^* : H^*(\mathrm{Gr}_{d_B k}(A \otimes B)) \longrightarrow H^*(\mathrm{Gr}_k(A))$ is the map on cohomology induced by ϕ. In order to compute when this product is nonzero, we wish to know the behavior of ϕ^*. This behavior is easier to determine using another presentation for the ring $H^*(\mathrm{Gr}(k,n))$, in terms of Chern classes of vector bundles. In this chapter we develop this presentation, show how it corresponds to the previous description of $H^*(\mathrm{Gr}(k,n))$ in terms of fundamental classes of Schubert varieties, and use it to describe how ϕ^* acts on $H^*(\mathrm{Gr}_{d_B k}(A \otimes B))$.

8.1 Vector Bundles

Recall that if M is a manifold, then a d-dimensional complex *vector bundle* is a map $p : E \to M$ such that the fiber $E_p \equiv p^{-1}(b)$ is an d-dimensional complex vector space for each $b \in M$, and the following local triviality condition is satisfied: there is an open cover $\{U_\alpha\}$ of M, together with homeomorphisms

$$h_\alpha : p^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}^d \tag{8.1}$$

that are vector space isomorphisms on each fiber. Often the total space E is referred to as the vector bundle, with the rest of the bundle structure implicit. If $d = 1$, then E is also referred to as a *line bundle*.

We will use several standard constructions of bundles:

(1) For any manifold M, and any d, there is the *trivial* or *product* bundle $E = M \times \mathbb{C}^d$, where p is the projection onto the first factor.

(2) If E and E' are bundles, then their direct sum $E \oplus E'$, their tensor product $E \otimes E'$, and the dual E^* are all defined in a natural way [44].

(3) Let M and N be manifolds and $p : E \to M$ a vector bundle over M. Then if $f : N \to M$ is a (continuous) map, it induces a vector bundle $f^*(E)$ on N, given by the following subset of $N \times E$:

$$\{(n, e) : f(n) = p(e)\}. \tag{8.2}$$

This bundle $f^*(E)$, called the *pullback of E by f*, is the unique maximal subset of $N \times E$ that makes the following diagram commute:

$$\begin{tikzcd} f^*(E) \arrow[r] \arrow[d] & E \arrow[d, "p"] \\ N \arrow[r, "f"'] & M. \end{tikzcd}$$

(4) Let V be a d-dimensional complex vector space and let $P(V)$ be its projectivization, that is, $P(V) = \mathrm{Gr}_1(V)$ is the set of one-dimensional subspaces of V. Let $\hat{V}$ be the product bundle $P(V) \times V$. Then the *universal subbundle S* is the subbundle of V given by

$$S = \{(\ell, v) \in P(V) \times V | v \in \ell\}, \tag{8.3}$$

also called the *tautological line bundle*; and the *universal quotient bundle Q* is defined by the exact sequence

$$0 \to S \to \hat{V} \to Q \to 0. \tag{8.4}$$

This is known as the *tautological exact sequence* over $P(V)$. The dual S^* is called the *hyperplane bundle.*

We will also use the following fact [45].

Proposition 8.1.1 *Let $0 \to A \to B \to C \to 0$ be an exact sequence of vector bundles. Then B is isomorphic as a bundle to $A \oplus C$.*

Instead of requiring the fiber of each point of a manifold M to be a vector space in our definition, we may have it be any topological space F, thus obtaining a *fiber bundle* with fiber F [45]. The main example of this will be the *projective bundle* $P(E) \to B$ associated to any d-dimensional vector bundle $E \to B$. The fiber at each point of $P(E)$ is isomorphic to the complex projective space $\mathbb{P}^{d-1}$, and the local trivializations of $P(E)$ are induced by those of E [44]. If we let p denote the projection from $P(E)$ to M, then we may pull back E by p to obtain a bundle $p^*(E)$ over $P(E)$, whose fiber at any point ℓ_p is E_p. As in example (4) above, this pullback bundle has a universal subbundle $S = \{(\ell_p, v) \in p^*(E) | v \in \ell_p\}$ and a universal quotient bundle Q defined by exactness of the sequence $0 \to S \to p^*(E) \to Q \to 0$.

8.2 Chern Classes

We now introduce *Chern classes*, which are integral cohomology classes associated to complex vector bundles. We will need the following fact. Let $\mathbb{P}^d$ be the d-dimensional complex projective space. Since PGL_{d+1} is a connected group acting transitively on the hyperplanes of $\mathbb{P}^d$, the fundamental class in cohomology associated to a hyperplane H does not depend on the chosen hyperplane. Let h denote this class, which we call the *hyperplane class.*

Chern classes can be defined axiomatically as follows [45]:

Theorem 8.2.1 *There are unique functions $c_1, c_2, \ldots$ on complex vector bundles $E \to N$, with $c_i(E) \in H^{2i}(M)$, that depend only on the isomorphism type of E and satisfy the following properties:*

(a) (functoriality) For any continuous map $f : N \to M$, $c_i(f^*(E)) = f^*(c_i(E))$.

(b) (Whitney sum formula) Writing $c = 1 + c_1 + c_2 + \ldots$, we have $c(E_1 \oplus E_2) = c(E_1) \cup c(E_2)$.

(c) If $i > \dim E$, then $c_i(E) = 0$.

(d) (normalization) For the tautological line bundle S on $\mathbb{P}^d$, $c_1(S) = -h$, the negative of the hyperplane class.

These classes $c_i(E)$ are called Chern classes *of the vector bundle E, and $c(E) = \sum_k c_k(E)$ is called the* total Chern class *of E (setting $c_0(E) = 1$).*

We note that the Whitney sum formula may be written as

$$c_k(E \oplus F) = \sum_{i+j=k} c_i(E) \cup c_j(F). \tag{8.5}$$

It can be shown [45] that the axiomatic properties of Chern classes imply that if L_1 and L_2 are line bundles, then $c_1(L_1 \otimes L_2) = c_1(L_1) + c_1(L_2)$. From this fact, it readily follows that $c_1(L) = 0$ if L is a trivial line bundle, and hence that $c_k(E)$ is zero for any trivial bundle E, by the Whitney formula.

We now specialize to the problem at hand. Let $\mathrm{Gr}(k, n)$ denote the Grassmannian space of k-dimensional subspaces of an n-dimensional complex vector space. Let T be the tautological bundle of dimension k over this Grassmannian, for which the fiber over a subspace V is V itself. Let Q be the quotient bundle over $\mathrm{Gr}(k, n)$ whose fiber over a vector space V is $\mathbb{C}^n/V$. Then the properties of Chern classes imply the following result [40].

Theorem 8.2.2 *The lth Chern class of the quotient bundle, $c_l(Q)$, is equal to the class of the special Schubert variety σ_l.*

Proof Fix a complete flag $F_\bullet$ for the n-dimensional complex vector space. Let $\alpha(1,l)$ be the partition corresponding to the complement of a $1\times l$ rectangle in the $k\times(n-k)$ rectangle. We must show that for any partition $\alpha\subset k\times(n-k)$ of weight $k(n-k)-l$, $c_l(Q)\cup\sigma_\alpha=1$ if $\alpha=\alpha(1,l)$, and $c_l(Q)\cup\sigma_\alpha=0$ otherwise.

Suppose that α has weight $k(n-k)-l$ but $\alpha\neq\alpha(1,l)$. Then $\alpha_k\geq n-k-l+1$, so any $V\in X_\alpha$ satisfies $\dim(V\cap F_{k+l-1})\geq k$. This means that $V\subset F_{k+l-1}$, so that X_α is contained in the smaller Grassmannian $G=\mathrm{Gr}(k,k+l-1)$ of k-dimensional subspaces of F_{k+l-1}. Let $j:G\hookrightarrow G(k,n)$ be the inclusion map. Using the projection formula from topology [40], we have that

$$c_l(Q)\cup\sigma_\alpha=j_*(j^*(c_l(Q))\cup[X_\alpha]),\tag{8.6}$$

where j_* is the Gysin homomorphism on cohomology arising from Poincaré duality. But by the exact sequence of bundles over G,

$$0\to F_{k+l-1}/V\to\mathbb{C}^n/V\to\mathbb{C}^n/F_{k+l-1}\to 0,\tag{8.7}$$

the restriction Q_G of the quotient bundle to G can be written $Q_G=F_{k+l-1}/V\oplus\mathbb{C}^n/F_{k+l-1}$, where the latter bundle in the direct sum is trivial. It follows from the Whitney formula that $c_l(Q_G)=0$, so since $c_l(Q_G)=j^*(c_l(Q))$, we must have that $c_l(Q)\cup\sigma_\alpha=0$ by Equation 8.6.

Now suppose that $\alpha=\alpha(1,l)$. In this case

$$X_\alpha=\{V\in\mathrm{Gr}(n,k)|F_{k-1}\subset V\subset F_{k-l}\},\tag{8.8}$$

which is isomorphic to the l-dimensional projective space $\mathbb{P}=\mathbb{P}(F_{k+l}/F_{k-1})$. Let i denote the natural isomorphism from X_α to $\mathbb{P}$. On $\mathbb{P}$ we have the exact sequence

$$0\to V/F_{k-1}\to F_{k+l}/F_{k-1}\to Q_{\mathbb{P}}\to 0.\tag{8.9}$$

Here V/F_{k-1} is the tautological line bundle, $Q_{\mathbb{P}}$ is the quotient bundle, and F_{k+l}/F_{k-1}

is a trivial bundle. It follows that the total Chern class of $Q_{\mathbb{P}}$ is $c(Q_{\mathbb{P}}) = (1-h)^{-1}$ (where h is the class of the hyperplane). Now the projection formula tells us that

$$\begin{aligned} c_l(Q) \cup \sigma_\alpha &= i_*(i^*(c_l(Q)) \cup [X_\alpha]) \\ &= i_*(i^*(c_l(Q)) \\ &= i_*(c_l(Q_{\mathbb{P}})) \\ &= 1. \end{aligned}$$

□

8.3 The Splitting Principle

We have seen that the Chern classes of the quotient bundle Q correspond to special Schubert classes in Grassmannian cohomology. Since all Schubert classes can be obtained as products of these special Schubert classes, characterizing the action of ϕ^* on the Chern classes of Q will be sufficient to determine the action of ϕ^* on $H^*(\mathrm{Gr}(k,n))$. To do this, we will need the *splitting principle*, an important result from the study of Chern classes of vector bundles. In what follows, let E be any vector bundle over a manifold M, whose dimension we denote by m. We shall have in mind the case where $M = \mathrm{Gr}(k,n)$ and E is the quotient bundle Q defined above (so that $m = n-k$).

Starting with the bundle E over M, let $P(E)$ be the projectivization of E, and let f_1 be the induced map from $P(E)$ to M. Let $f_1^*(E)$ be the pullback bundle:

$$\begin{tikzcd} f_1^*(E) \arrow[r] \arrow[d] & E \arrow[d] \\ P(E) \arrow[r, "f_1"'] & M. \end{tikzcd}$$

Let L_1 be the tautological line bundle of the pullback $f^*(E)$. Then we have an exact sequence

$$0 \to L_1 \to f_1^*(E) \to Q_1 \to 0, \tag{8.10}$$

where E is an $(m-1)$-dimensional bundle over M, so $f_1^*(E)$ is isomorphic to $L_1 \oplus Q_1$. Similarly, let $P(Q_1)$ be the projectivization of Q_1, with f_2 as the map from $P(Q_1)$ to $P(Q)$. If L_2 is the tautological line bundle of $P(Q_1)$, then L_2 gives rise to a quotient Q_2 such that $f_2^*(Q_1)$ is isomorphic to $L_2 \oplus Q_2$. We can thus pull back E to a direct sum of Q_2 and two line bundles:

$$\begin{tikzcd}
 & & f_2^*(L_1) \oplus L_2 \oplus Q_2 \arrow[d] \\
 & L_1 \oplus Q_1 \arrow[d] & P(Q_1) \arrow[dl, "f_2"'] \\
E \arrow[d] & P(E) \arrow[dl, "f_1"'] & \\
M & &
\end{tikzcd} \tag{8.11}$$

Continuing in this way, we obtain bundles $Q_3, \ldots, Q_{m-1}$, and projectivizations $P(Q_2), \ldots, P(Q_{m-2})$, such that the pullback of E by the map from $P(Q_{m-2})$ to M is a direct sum of line bundles. If $f = f_1 \circ f_2 \circ \ldots f_{m-2}$ is the map from $P(Q_{m-2})$ to M, then it can be shown that the induced map on cohomology $f^* : H^*(M) \to H^*(P(Q_{m-2}))$ is injective [44]. We summarize these facts in the following theorem, known as the splitting principle:.

Theorem 8.3.1 (The Splitting Principle) *For any vector bundle E on a manifold M, there exists a manifold N and a continuous $f : N \to M$ such that $f^*(M) \to f^*(N)$ is injective, and pullback bundle $f^*(E)$ is a direct sum of line bundles.*

We now illustrate the splitting principle by using it to derive a result that will be useful to us. Let E be a vector bundle, and let $f : N \to M$ be the map given by Theorem 8.1.1, so that the pullback $f^*(E)$ splits as the direct sum of line bundles $L_1, \ldots, L_n$. Let $x_i = c_1(L_i)$. Then the Whitney sum formula $c_k(E_1 \oplus E_2) = \sum_{i+j=k} c_i(E_1) \cup c_j(E_2)$ implies that

$$c_k(f^*(E)) = c_k(x_1, \ldots, x_n) \tag{8.12}$$

is the kth elementary symmetric polynomial in the first Chern classes of $f^*(E)$. By the functoriality of the Chern classes, it follows that $f^*(c_k(E))$ is the kth elementary symmetric polynomial in $c_1(L_1), \dots, c_1(L_n)$.

Let us revisit the construction of the split manifold of a vector bundle E. $P(E)$ consists of pairs (x, ℓ), where $x \in M$ and ℓ is a line in E_x. Proposition 8.1.1 allows us to consider all the bundles $Q_1, \dots Q_{n-1}$ as subbundles of E. Now $P(Q_1)$ consists of triples (x, ℓ_1, ℓ_2) where ℓ_2 is a line in the linear complement of ℓ_1 in E_p. In general, a point of $P(Q_j)$ over $(x, \ell_1, \dots, \ell_j)$ in $P(Q_{j-1})$ is a $(j+2)$-tuple $(x, \ell_1, \dots, \ell_j, \ell_{j+1})$ where ℓ_{j+1} is a line in the complement of $\ell_1, \dots, \ell_j$. We conclude that the split manifold $P(Q_{m-2})$ is in fact the *flag bundle*:

$$\mathrm{Fl}(E) = \{(x, \ell_1 \subset \langle \ell_1, \ell_2 \rangle \subset \langle \ell_1, \ell_2, \ell_3 \rangle \subset \ldots \subset E_x) | x \in M\}. \tag{8.13}$$

8.4 Representations and Line Bundles

We have seen that the splitting principle allows us to regard the Chern classes of a vector bundle E as (symmetric) polynomials in the first Chern classes of the line bundles of a flag bundle associated to E. Given an m-dimensional vector space V, the space $F\ell(V)$ of all complete flags on V can be identified with $\mathrm{GL}(V)/P$, where P is the group of upper triangular matrices. (This follows because $\mathrm{GL}(V)$ is transitive on the flags and P, the stabilizer of the standard flag $0 \subset \langle e_1 \rangle \subset \langle e_1, e_2 \rangle \subset \cdots \subset \langle e_1, \dots, e_m \rangle = V$, is isomorphic to the stabilizer of any given flag.) We can associate any one-dimensional representation $\chi : P \to \mathbb{C}^*$ to an equivariant line bundle over the flag manifold $F\ell(V)$ as follows:

$$L(\chi) = \mathrm{GL}(V) \times \mathbb{C}/((gp, z) \sim (g, \chi(p)z)) \tag{8.14}$$

for $g \in \mathrm{GL}(V)$, $p \in P$, and $z \in \mathbb{C}$. The projection of $L(\chi)$ onto $F\ell(V)$ is just $(g, z) \xrightarrow{\pi} (gP)$. Under the action of $\mathrm{GL}(V)$ given by $h(gp, z) = (hgp, z)$, the following diagram commutes:

$$\begin{tikzcd} L(\chi) \arrow[r, "GL(V)"] \arrow[d, "\pi"'] & L(\chi) \arrow[d, "\pi"] \\ F\ell(V) \arrow[r, "GL(V)"] & F\ell(V), \end{tikzcd}$$

since $(hgp, z) = (hg, \chi(p)z)$. The line bundle $L(\chi)$ is thus equivariant with respect to the bundle projection.

Conversely, suppose L is an equivariant line bundle over $F\ell(V)$. Then P acts on the fiber over eP, so this fiber is a one-dimensional representation χ of P. Let us show that the line bundle $L(\chi)$ corresponding to this representation is isomorphic to L. Let $y \in L$ lie in the fiber over eP. Then we claim that the map $r : L(\chi) \to L$ given by $r(g, z) = z(g \cdot y)$ is an isomorphism. We have $r(gp, z) = z(gp \cdot y) = z(g\chi(p)y) = r(g, \chi(p)z)$, so this map is well-defined. Since G acts transitively on the fibers of L, and multiplication by z is a surjective map on any given fiber, r is surjective. For injectivity, suppose that $z_1(g_1 \cdot y) = z_2(g_2 \cdot y)$. If $z_1 \neq 0$, then $y = z_1^{-1}z_2g_1^{-1}g_2 \cdot y$, so $g_1^{-1}g_2 \in P$. This means that $g_2 = g_1p$ for some $p \in P$ and $z_1^{-1}z_2\chi(p) = 1$, so $z_2 = z_1\chi(p)^{-1}$. Thus, as elements of $L(\chi)$, $(g_2, z_2) = (g_1p, \chi(p)^{-1}z_1) = (g_1, z_1)$, so r is indeed injective. The correspondence between line bundles and one-dimensional representations of P is therefore a bijection.

We can identify the characters χ with line bundles on a flag manifold $F\ell(V)$ more explicitly. Consider the *tautological filtration* [41]

$$0 = U_0 \subset U_1 \subset U_2 \subset \cdots \subset U_m = F\ell(V) \times V \tag{8.15}$$

of vector bundles over $F\ell(V)$, where $F\ell(V) \times V$ is the product bundle, and U_k is the k-dimensional bundle over $F\ell(V)$ whose fiber over a flag $V_1 \subset \cdots \subset V_m$ is V_k. It follows from the splitting principle that the cohomology ring $H^*(F\ell(V))$ is generated by the first Chern classes of the line bundles $L_i = U_i/U_{i-1}$, setting $x_i = c_1(L_1)$. The identity matrix fixes the standard flag $\{e_1, \ldots, e_m\}$. Therefore, over eP, the fiber of L_i is V_i/V_{i-1}, where $V_i = \langle e_1, \ldots, e_i\rangle$. If $v = \sum_{k=1}^{i} \alpha_i e_i \in V_i$ and $p \in P$, then $p \cdot v = w + p_{ii}e_i$, where $w \in V_{i-1}$ and p_{ii} is the ith diagonal entry of p. We have shown the following.

Theorem 8.4.1 *If L_i is the line bundle over a flag manifold defined as above, then the character χ associated to L_i is the map taking p to p_{ii}.*

Let us adapt this machinery to the problem at hand. Recall that we have two complex vector spaces A and B of dimensions d_A and d_B, respectively, together with a map $\phi : \mathrm{Gr}_k(A) \to \mathrm{Gr}_{kd_B}(A \otimes B)$ given by $\phi(V) = V \otimes B$. We wish to compute the action of the induced map $\phi^* : H^*(\mathrm{Gr}_{kd_B}(A \otimes B)) \to H^*(\mathrm{Gr}_k(A))$.

Let Q_A and Q_{AB} be the quotient bundles of Theorem 8.2.2 over the Grassmannians $\mathrm{Gr}_k(A)$ and $\mathrm{Gr}_{kd_B}(A \otimes B)$ respectively. The Chern classes of these bundles are the classes of the special Schubert varieties in the cohomology rings. By the splitting principle, the associated flag bundles $\mathrm{Fl}(A)$ and $\mathrm{Fl}(A \otimes B)$ have pullbacks which split as a direct sum of line bundles L_i of the respective tautological filtrations. The cohomology of the Grassmannians embed in the cohomology of these pullbacks, so we may determine ϕ^* by its action on the Chern classes of the pullback bundle of $\mathrm{Fl}(A \otimes B)$.

It follows from the definition of pullback bundles that the bundle $\phi^*(L(\chi_i))$ is the set of triples $(gP_A, \phi(g), z) \in \mathrm{GL}(A)/P_A \times \mathrm{GL}(A \otimes B) \times \mathbb{C}$ with the identification $(gP_A, \phi(g \cdot p), z) \sim (gP_A, \phi(g), \chi(\phi(p))z)$. This means that $\phi^*(L(\chi_i)) = L(\phi^*(\chi_i))$. And the pullback of the map induced by ϕ on the characters of the group P_{AB} is readily computed: for a matrix $X \in P_A$, and the character χ_i taking a matrix to its ith diagonal entry, we have $\phi^*(\chi_i)(X) = \chi_i(\phi(X)) = \chi_i(X \otimes I) = \chi_{\lceil i/d_B \rceil}(X)$. So $\phi^*(\chi_i) = \chi_{\lceil i/d_B \rceil}$. Now we can calculate the action of ϕ^* on the Chern classes:

Theorem 8.4.2 $\phi^*(x_i) = c_1(L(\chi_{\lceil i/d_B \rceil}))$.

Proof

$$\begin{aligned}
\phi^*(x_i) &= \phi^*(c_1(L(\chi_i))) && (8.16)\\
&= c_1(\phi^*(L(\chi_i))) && (8.17)\\
&= c_1(L(\phi^*(\chi_i))) && (8.18)\\
&= c_1(L(\chi_{\lceil i/d_B \rceil})). && (8.19)
\end{aligned}$$

□

Chapter 9

Determining the Inequalities

In this chapter we use our knowledge of how ϕ^* behaves to explicitly derive inequalities relating the spectra of ρ_{AB} and of ρ_A. We work out some examples in low dimensions. We also restate how to obtain the inequalities in the language of representation theory. We discuss recent progress in symplectic geometry that shows that the inequalities derived using our method are sufficient. Finally, we prove that if $d_B \geq \frac{1}{2}d_A^2$, then the inequalities simplify greatly.

9.1 Putting It All Together

Let $\rho_A = \mathrm{Tr}_B\,\rho_{AB}$, and let λ, μ, and $\tilde{\lambda}$ denote the spectra of ρ_{AB}, $-\rho_{AB}$, and ρ_A, respectively. Theorem 6.2.2 can be interpreted cohomologically as saying that if

$$\phi^*(\sigma_\pi) \cup \tilde{\sigma}_\nu \neq 0, \tag{9.1}$$

where $\sigma_\pi \in H^*(\mathrm{Gr}(kd_B, d_A d_B))$ and $\tilde{\sigma}_\nu \in H^*(\mathrm{Gr}(k, d_a))$ are Schubert classes, then the spectra μ and $\tilde{\lambda}$ must satisfy the inequalities

$$\sum \nu(i)\tilde{\lambda}_i + \sum \pi(i)\mu_i \leq 0. \tag{9.2}$$

Now $\phi^*(\sigma_\pi)$ is an integer combination of Schubert classes,

$$\phi^*(\sigma_\pi) = \sum_i n_i \tilde{\sigma}_{\pi_i}. \tag{9.3}$$

For each of these classes, $\tilde{\sigma}_{\pi_i} \cup \tilde{\sigma}_\nu \neq 0$ iff ν contains the complement of π_i in the $k \times (n-k)$ rectangle. But if we consider the case where ν is in fact the complement of π_i, then we see that the Inequalities 9.2 are the strongest in this case; for any other $\nu' \supset \nu$, the inequalities determined by ν' are implied by the inequalities determined by ν. So it is sufficient to consider complements of each Schubert class $\tilde{\sigma}_{\pi_i}$ contained in $\phi^*(\sigma_\pi)$, in order to obtain the inequalities relating $-\rho_{AB}$ and ρ_A. Now if μ is the spectrum of $-\rho_{AB}$, then the spectrum λ of ρ_{AB} is given by $\lambda_i = -\mu_{d_A-i+1}$ (since the ordering of the eigenvalues is reversed). Given binary strings $\pi, \hat{\pi} \in \binom{d_A d_B}{k}$ satisfying $\hat{\pi}(i) = \pi(d_A d_B - i + 1)$, so that $\hat{\pi}$ is simply the string π in reverse, the Schubert cell $S_{\hat{\pi}}$ corresponds to the complementary partition to that of S_π. This means that we obtain inequalities

$$\sum \nu(i)\tilde{\lambda}_i \leq \sum \pi(i)\lambda_i \tag{9.4}$$

whenever $\phi^*(\sigma_{\hat{\pi}})$ contains $\sigma_{\hat{\nu}}$ (where $\hat{\nu}$ is the complementary partition to ν) as a summand. It then follows that Inequalities 9.4 are obtained whenever $\phi^*(\sigma_\pi)$ contains σ_ν as a summand.

Theorem 8.2.2 says that the lth Chern class $c_l(Q)$ of the universal quotient bundle Q over the Grassmannian $\mathrm{Gr}(k,n)$ is equal to the special Schubert class $\sigma_l \in H^*(\mathrm{Gr}(k,n))$. And the splitting principle allows us to conclude that

$$c_l(Q) = e_l(x_1, \ldots, x_{n-k}), \tag{9.5}$$

where $x_i = c_1(L_i)$ is the first Chern class of the ith split component of $f^*(Q)$, and e_l is the lth elementary symmetric polynomial. Because the special Schubert classes σ_l generate the cohomology ring, we therefore have a surjective ring homomorphism

$$\begin{aligned} \tilde{\psi} : \qquad \Lambda_{n-k} \qquad &\rightarrow H^*(\mathrm{Gr}(k,n)) \\ e_l(x_1, \ldots x_{n-k}) \quad &\mapsto \sigma_l. \end{aligned}$$

We may compose the map $\tilde{\psi}$ with the involution $\omega : \Lambda_{n-k} \rightarrow \Lambda_{n-k}$, $\omega(e_k) = h_k$, to obtain a map

$$\begin{aligned} \psi : \qquad \Lambda_{n-k} &\rightarrow H^*(\mathrm{Gr}(k,n)) \\ h_l(x_1, \ldots x_{n-k}) &\mapsto \sigma_l. \end{aligned}$$

Now, by the Pieri rule, it follows that for any partition λ, $\psi(s_\lambda(x_1, \ldots, x_{n-k})) = \sigma_\lambda$. Thus, we may determine how ϕ^* acts on $H^*(\mathrm{Gr}(kd_B, d_A d_B))$ by determining how the map $x_i \mapsto x_{\lceil i/d_B \rceil}$ acts on Schur functions.

9.2 Some Observations

In this section we make some observations about the map ϕ^* that will simplify our computations to some degree. First, we note that ϕ^* is particularly easy to calculate on the Newton power sums $p_j = \sum_i x_i^j$:

$$\begin{aligned} \phi^*(p_j(x_1, \ldots, x_{(d_A-k)d_B})) &= \phi^*(\sum_{i=1}^{(d_A-k)d_B} x_i^j) && (9.6) \\ &= \sum_{i=1}^{(d_A-k)d_B} x_{\lceil i/d_B \rceil}^j && (9.7) \\ &= \sum_{i=1}^{d_A-k} d_B x_i^j && (9.8) \\ &= d_B p_j(x_1, \ldots, x_{d_A-k}). && (9.9) \end{aligned}$$

We further note that the total degree of a polynomial in the Chern classes $x_1, \ldots x_{n-k}$ is equal to the weight of the corresponding partition, and ϕ^* maps every monomial in $x_1, \ldots, x_{(d_A-k)d_B}$ to a monomial in $x_1, \ldots, x_{d_A-k}$ of the same total degree, so that $\phi^*(\sigma_\pi)$ is a sum of Schubert classes of the same weight as π.

Applying this observation to the empty partition $\alpha = (0)$, which corresponds to the binary string $\underbrace{11\ldots1}_{k}\underbrace{00\ldots0}_{n-k}$ in $\mathrm{Gr}(k,n)$, we obtain the inequalities

$$\sum_{i=1}^{k}\tilde{\lambda}_i \leq \sum_{i=1}^{d_B k}\lambda_i \tag{9.10}$$

for every $k \in \{1,\ldots,d_A\}$. These are the same inequalities previously derived in Theorem 6.1.2, using only Ky Fan's Maximum Principle. We will call Inequalities 9.10 *basic inequalities*. As we shall see, many of the inequalities that arise from considering the intersections of Schubert classes will not contain additional information; rather, they will be consequences of the basic inequalities. We call such inequalities *redundant inequalities*.

Finally, we argue that it is sufficient to consider inequalities derived from ϕ^* acting on $H^*(\mathrm{Gr}(kd_B, d_Ad_B))$, where $k \leq \frac{d_A}{2}$. To see this, suppose there is an inequality of the form

$$\sum_{i=1}^{d_A}\nu(i)\tilde{\lambda}_i \leq \sum_{i=1}^{d_Ad_B}\pi(i)\lambda_i, \tag{9.11}$$

where the weight of ν is greater than $\frac{d_A}{2}$. We may apply this inequality to the matrices $-\rho_{AB}$ and $-\rho_A$ and use the trace condition to conclude that

$$\sum_{i=1}^{d_A}\nu'(i)\tilde{\lambda}_i \leq \sum_{i=1}^{d_Ad_B}\pi'(i)\lambda_i, \tag{9.12}$$

where $\nu'(i) = 1 - \nu(i)$ for all i, and similarly for π'. If the weight of ν is greater than $\frac{d_A}{2}$, then the weight of ν' is less than $\frac{d_A}{2}$. Thus, the desired inequality is a consequence of an inequality involving fewer than $\frac{d_A}{2}$ eigenvalues. (This argument is not valid unless we know that our method generates all possible valid inequalities. This is indeed the case, but we postpone the discussion for Section 9.5.)

α	s_α	$\pi_\alpha \in H^*(\mathrm{Gr}(2,6))$	$\pi_\alpha \in H^*(\mathrm{Gr}(1,3))$
$\square$	p_1	101000	010
$\square\!\square$	$\frac{1}{2}(p_1^2 + p_2)$	100100	001
$\substack{\square\\\square}$	$\frac{1}{2}(p_1^2 - p_2)$	011000	—

Figure 9.1 Partitions, their Schur polynomials and binary strings

9.3 Examples

We now work out the inequalities for some examples. The case $d_A = 2$ was already solved in Section 6.3, where it was shown that the basic inequalities were the only constraints on the eigenvalues of ρ_A and ρ_{AB}. Thus, the simplest remaining case is $d_A = 3$, $d_B = 2$, which we will now illustrate. We use h_l to refer to the lth complete symmetric function, and p_l to refer to the lth Newton power sum symmetric function. We identify Schur functions with their images as Schubert classes, denoting either by a (Young diagram of a) partition.

As we have argued, we may restrict attention to inequalities involving at most $\frac{d_A}{2}$ eigenvalues; in the case $d_A = 3$, this means that it suffices to consider maps $\phi^* : H^*(\mathrm{Gr}(2,6)) \to H^*(\mathrm{Gr}(1,3))$. The Schubert classes of $H^*(\mathrm{Gr}(1,3))$ correspond to partitions that fit inside a 1×2 rectangle, of which there are only two (excluding the empty partition, for which we obtain the basic inequalities): $\square$ and $\square\!\square$. Because ϕ^* preserves the weight of a partition, we need only consider partitions of weight one and two in $H^*(\mathrm{Gr}(2,6))$: namely, $\square$, $\square\!\square$, and $\substack{\square\\\square}$. Figure 9.1 lists the Schur polynomials and binary strings associated to each of these partitions (the polynomials are readily computed using the Jacobi-Trudi formula).

Using this information, we can calculate ϕ^* on each of the Schubert classes $\square$, $\square\!\square$, and $\substack{\square\\\square} \in H^*(\mathrm{Gr}(2,6))$:

(1) $\phi^*(\square) = \phi^*(p_1) = 2p_1 = 2\square$. This yields the inequality $\tilde{\lambda}_2 \leq \lambda_1 + \lambda_3$.

(2) $\phi^*(\square\!\square) = \phi^*(\frac{1}{2}(p_1^2 + p_2)) = 2p_1^2 + p_2 = 3\square\!\square + \substack{\square\\\square}$. For the $\square\!\square$ term on the right side, we get the inequality $\tilde{\lambda}_3 \leq \lambda_1 + \lambda_4$. The $\substack{\square\\\square}$ term does not yield an inequality because $\substack{\square\\\square} = 0$ in $H^*(\mathrm{Gr}(1,3))$.

(3) $\phi^*(\begin{smallmatrix}\square\\\square\end{smallmatrix}) = \phi^*(\frac{1}{2}(p_1^2 - p_2)) = 2p_1^2 - p_2 = 3\begin{smallmatrix}\square\\\square\end{smallmatrix} + \square\square$. As before, the $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ term does not yield an inequality. The $\square\square$ term yields the inequality $\tilde{\lambda}_3 \leq \lambda_2 + \lambda_3$.

So we have three inequalities, $\tilde{\lambda}_2 \leq \lambda_1 + \lambda_3$, $\tilde{\lambda}_3 \leq \lambda_1 + \lambda_4$, and $\tilde{\lambda}_3 \leq \lambda_2 + \lambda_3$. Let us check these inequalities for redundancy. From the basic inequalities, we have that $\tilde{\lambda}_2 \leq \frac{1}{2}(\tilde{\lambda}_1 + \tilde{\lambda}_2) \leq \frac{1}{2}(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \leq \lambda_1 + \lambda_3$, so the first inequality is redundant. And $\tilde{\lambda}_3 \leq \frac{1}{3}(\tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_3) \leq \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6) \leq \lambda_1 + \lambda_4$, so the second inequality is also redundant. However, the inequality $\tilde{\lambda}_3 \leq \lambda_2 + \lambda_3$ is not redundant (for example, $\lambda = (1, 0, 0, 0, 0, 0)$ and $\tilde{\lambda} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ satisfy the basic inequalities, but not $\tilde{\lambda}_3 \leq \lambda_2 + \lambda_3$).

So $\tilde{\lambda}_3 \leq \lambda_2 + \lambda_3$ is the only new inequality we get involving one eigenvalue of ρ_A. By duality, we also have the inequality $\tilde{\lambda}_2 + \tilde{\lambda}_3 \leq \lambda_1 + \lambda_2 + \lambda_3 + \lambda_6$, or $\tilde{\lambda}_1 \geq \lambda_4 + \lambda_5$. Thus, our complete list of eigenvalue constraints on ρ_{AB} and ρ_A is

$$\tilde{\lambda}_1 \leq \lambda_1 + \lambda_2, \tag{9.13}$$

$$\tilde{\lambda}_3 \geq \lambda_5 + \lambda_6, \tag{9.14}$$

$$\tilde{\lambda}_3 \leq \lambda_2 + \lambda_3, \tag{9.15}$$

$$\tilde{\lambda}_1 \geq \lambda_4 + \lambda_5, \tag{9.16}$$

together with the trace condition $(\tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_3) = (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6)$.

Now we consider the case $d_A = 3$, $d_B = 3$. We have that

$$\phi^*(\square) = 3\square, \tag{9.17}$$

$$\phi^*(\square\square) = 6\square\square + 3\begin{smallmatrix}\square\\\square\end{smallmatrix}, \tag{9.18}$$

$$\phi^*(\begin{smallmatrix}\square\\\square\end{smallmatrix}) = 6\begin{smallmatrix}\square\\\square\end{smallmatrix} + 3\square\square, \tag{9.19}$$

yielding inequalities

$$\tilde{\lambda}_2 \leq \lambda_1 + \lambda_2 + \lambda_4, \tag{9.20}$$

$$\tilde{\lambda}_3 \leq \lambda_1 + \lambda_2 + \lambda_5, \tag{9.21}$$

$$\tilde{\lambda}_3 \leq \lambda_1 + \lambda_3 + \lambda_4. \tag{9.22}$$

It is not hard to check that all of these inequalities are redundant. Thus, our only inequalities for the case $d_A = 3$, $d_b = 3$ are the basic inequalities

$$\tilde{\lambda}_1 \leq \lambda_1 + \lambda_2 + \lambda_3, \tag{9.23}$$

$$\tilde{\lambda}_3 \geq \lambda_7 + \lambda_8 + \lambda_9. \tag{9.24}$$

9.4 Representation Theory Perspective

Given a Schur polynomial s_λ, we have seen how to determine $\phi^*(s_\lambda)$ as follows: write s_λ in terms of Newton power sums, evaluate ϕ^* on each of the power sums, and then express the results in terms of Schur polynomials. While this algorithm is fairly straightforward, the relationship between s_λ and the terms appearing in $\phi^*(s_\lambda)$ is less clear. In this section, we see that we can interpret this relationship from the standpoint of group representation theory. Asking which Schur polynomials appear in $\phi^*(s_\lambda)$ is equivalent to asking which irreducible representations appear in a certain tensor product of representations of the symmetric group.

While we are concerned with the action of ϕ^* on Schur polynomials acting on a fixed number of variables, we will simplify our discussion by working in the ring of symmetric functions. Define a *symmetric function* to be a set of symmetric polynomials $p(x_1, \ldots, x_l)$, one for each positive integer l, such that

$$p(x_1, \ldots, x_l, 0, \ldots, 0) = p(x_1, \ldots, x_1). \tag{9.25}$$

Recall that the Newton power sum symmetric functions are defined as follows. For a nonnegative integer s (which we may also think of as a partition of one part of size s), $p_s(X_1, \ldots, X_k) = X_1^s + \cdots + X_k^s$. For a partition $\lambda = (\lambda_1, \ldots, \lambda_l)$ of length l, define $p_\lambda(X_1, \ldots, X_k) = \prod_{i=1}^{l} p_{\lambda_i}(X_1, \ldots, X_k)$. As we have seen, $\phi^*(p_s) = d_B p_s$, so that $\phi^*(p_\lambda) = d_B^{l(\lambda)} p_\lambda$, where $l(\lambda)$ is the length of the partition λ.

We use the following basic facts about the representation theory of the symmetric group [41, 46]. The irreducible representations of the symmetric group S_n on n letters can be put in one-to-one correspondence with the partitions of n, in a standard way. (And the partitions of n also correspond naturally to the conjugacy classes of S_n.) Furthermore, the Newton power sum symmetric functions p_μ and the Schur polynomials s_λ are related as follows. For any partition μ of n, define

$$z(\mu) = \prod_r r^{m_r}(m_r!), \tag{9.26}$$

where m_r is the number of times r occurs in μ. Now for any partition μ of n,

$$p_\mu = \sum_\lambda \chi_\mu^\lambda s_\lambda; \tag{9.27}$$

and for any partition λ of n,

$$s_\lambda = \sum_\mu \frac{1}{z(\mu)} \chi_\mu^\lambda p_\mu, \tag{9.28}$$

where χ_μ^λ is the character of the representation labelled by λ evaluated on a permutation in the conjugacy class labelled by μ.

Let us now return to the fact that $\phi^*(p_\lambda) = d_B^{l(\lambda)} p_\lambda$. This means that ϕ^* is a class function on S_{d_B} (where $d_B = |\lambda|$), so we wish to find a representation ρ of S_{d_B} such that the character χ^ρ of ρ is equal to ϕ^*. Consider the representation ρ of S_{d_B} on $B^{\otimes d_B}$ that acts by permuting the tensor factors: if $\{e_i\}_{i=1}^{d_B}$ is an orthogonal basis for

B, then for $w \in S_{d_B}$,

$$\rho(w)(e_{i_1} \otimes \cdots \otimes e_{i_{d_B}}) = e_{i_{w(1)}} \otimes \cdots \otimes e_{i_{w(d_B)}}. \tag{9.29}$$

We claim that the character $\chi^\rho = \phi^*$, or in other words, for any $w \in S_{d_B}$, the character of ρ evaluated at w is $d_B^{l(w)}$, where $l(w)$ is the number of cycles in w. To see this, recall that by definition, $\chi^\rho(w) = \mathrm{Tr}(\rho(w))$. So $\chi^\rho(w)$ is the number of elements of the basis $\{e_{i_1} \otimes \cdots \otimes e_{i_{d_B}}\}$ fixed by the map ρ; in other words,

$$\chi^\rho(w) = |\{(i_1, \ldots, i_{d_B}) = (i_{w(1)}, \ldots, i_{w(d_B)})\}|. \tag{9.30}$$

Now, if $(i_1, \ldots, i_{d_B})$ is fixed by w, then for any r_1 and r_2 in the same cycle of w, we must have $i_{r_1} = i_{r_2}$. Conversely, if $(i_1, \ldots, i_{d_B})$ satisfies the property that $i_{r_1} = i_{r_2}$ for any r_1 and r_2 in the same cycle of w, then $(i_1, \ldots, i_{d_B})$ is fixed by w. We conclude that the number of elements in the set $\{(i_1, \ldots, i_{d_B}) = (i_{w(1)}, \ldots, i_{w(d_B)})\}$ is equal to the number of ways to assign a basis element to each cycle of w, which is $d_B^{l(w)}$.

Let V_λ be the irreducible representation of S_{d_B} labelled by λ. Let

$$V_\lambda \otimes \rho = \oplus_\pi (V_\pi)^{\otimes m_\pi} \tag{9.31}$$

be a decomposition of $V_\lambda \otimes \rho$ into irreducible representations (each irrep V_π occurs with multiplicity m_π). Then we have that

$$\chi^\lambda(\mu)\chi^\rho(\mu) = \sum_\pi m_\pi \chi^\pi(m), \tag{9.32}$$

a result we will use in the next calculation.

Now let us calculate $\phi^*(s_\lambda)$:

$$\begin{aligned}
\phi^*(s_\lambda) &= \sum_\mu \frac{1}{z(\mu)}\chi^\lambda(\mu)\phi^*(p_\mu) && (9.33)\\
&= \sum_\mu \frac{1}{z(\mu)}\chi^\lambda(\mu)d_B^{l(\mu)}p_\mu && (9.34)\\
&= \sum_\mu \frac{1}{z(\mu)}\chi^\lambda(\mu)\chi^\rho(\mu)p_\mu && (9.35)\\
&= \sum_\mu \frac{1}{z(\mu)}\sum_\pi m_\pi\chi^\pi(\mu)(p_\mu) && (9.36)\\
&= \sum_\pi\sum_\mu \frac{1}{z(\mu)}\chi^\pi(\mu)(p_\mu) && (9.37)\\
&= \sum_\pi m_\pi s_\pi. && (9.38)
\end{aligned}$$

So the Schur polynomials s_π appearing in $\phi^*(s_\lambda)$ are precisely those corresponding to the representations V_π appearing in $V_\lambda \otimes \rho$.

9.5 Sufficiency

We have described an approach using a variational principle to determine inequalities relating a matrix ρ_{AB} to its partial trace ρ_A, along with some observations for simplifying the list of inequalities. While our method has the advantage of relative straightforwardness and simplicity, our techniques do not (to our knowledge) allow us to demonstrate that the inequalities obtained are in fact sufficient: that is, if λ and $\tilde{\lambda}$ satisfy the inequalities, then there exists matrices ρ_{AB} and $\rho_A = \mathrm{Tr}_B\,\rho_{AB}$ such that λ is the spectrum of ρ_{AB} and $\tilde{\lambda}$ is the spectrum of ρ_A. It turns out that the inequalities obtained from our variational principle approach are indeed sufficient. This follows from recent work in symplectic geometry [47], of which we became aware after deriving the inequalities through our methods. In this section, we will state the main result from [47] and show that it yields inequalities equivalent to the ones we have obtained.

We begin with some background from symplectic geometry [47–49]. Let M be a symplectic manifold with symplectic form ω, and let K be a connected Lie group acting on M. K acts on itself by conjugation, and therefore it also acts on its Lie algebra $\mathfrak{k}$ by conjugation. This is the *adjoint representation* of K on $\mathfrak{k}$. The adjoint representation in turn induces an action on the dual space $\mathfrak{k}^*$, a map $\mathrm{Ad}^* : K \to \mathrm{GL}(\mathfrak{k}^*)$ given by $\langle \mathrm{Ad}^*_k \xi, X\rangle = \langle \xi, \mathrm{Ad}_{k^{-1}} X\rangle$ for $\xi \in \mathfrak{k}^*, X \in \mathfrak{k}$, where $\langle \cdot, \cdot \rangle$ is the natural pairing between $\mathfrak{k}^*$ and $\mathfrak{k}$. Ad^* is known as the *coadjoint representation* of K on $\mathfrak{k}^*$. The *coadjoint orbit* $\mathrm{Orb}(\xi)$ through $\xi \in \mathfrak{k}^*$ is defined by

$$\mathrm{Orb}(\xi) = \{Ad^*_{k^{-1}}(\xi) | k \in K\}. \tag{9.39}$$

It is a fact [48] that a unique symplectic manifold structure can be given to any coadjoint orbit $\mathrm{Orb}(\xi)$ of a Lie group, such that the inclusion map $\mathrm{Orb}(\xi) \hookrightarrow \mathfrak{k}^*$ is a moment map (defined next).

A map $\Phi : M \to \mathfrak{k}^*$ is a *moment map* for the action of K on M if the following two conditions hold.

(1) Let $X \in \mathfrak{k}$, so X induces a vector field on M, generated by the one-parameter subgroup $\{\exp tX | t \in \mathbb{R}\}$. Denote this vector field $X^\#$. Let $\Phi^X : M \to \mathbb{R}$ be given by $\Phi^X(p) = \langle \Phi(p), X\rangle$. Then the condition is that Φ^X is a Hamiltonian function for the vector field $X^\#$:

$$d\Phi^X = i_{X^\#}\omega.$$

(This is equivalent to saying that $X^\#$ is the symplectic gradient of $\langle \Phi, X\rangle$.)

(2) Φ is equivariant with respect to the action of K on M and the coadjoint action Ad^* of K on $\mathfrak{k}^*$: $\Phi \circ k = \mathrm{Ad}^*_k \circ \Phi$, for all $k \in K$.

If an action has a moment map then it is said to be *Hamiltonian* and M is called a *Hamiltonian K-manifold.*

We can express our problem in the language of symplectic geometry. Consider the Lie group $U(A \otimes B)$ of unitary matrices on the space $A \otimes B$. For any vector

$\lambda = (\lambda_1, \ldots, \lambda_{d_A d_B})$ with terms arranged in nonincreasing order, the set $\mathcal{O}_\lambda^{AB}$ of Hermitian matrices on $A \otimes B$ with spectrum λ is a coadjoint orbit of $K = U(A \otimes B)$. Now consider the action of the Lie group $\tilde{K} = U(A)$ of unitary matrices on A, by conjugation on the symplectic manifold $\mathcal{O}_\lambda^{AB}$: for $U \in U(A)$,

$$U : \rho_{AB} \mapsto (U \otimes I_B)\rho_{AB}(U^\dagger \otimes I_B). \tag{9.40}$$

It is not hard to verify that this is a Hamiltonian group action, whose moment map is Tr_B, the partial trace with respect to B. So our problem, then, is to describe the image of the symplectic manifold $\mathcal{O}_\lambda^{AB}$ under the moment map Tr_B.

This formulation is useful because considerable work has been done in the study of the image of moment maps. For instance, the following result is due to Kirwan [48]:

Theorem 9.5.1 *Let M be a compact connected Hamiltonian K-manifold, with moment map Φ. Then the intersection of the image of Φ with the positive Weyl chamber $\mathfrak{t}_+^*$ is a convex polytope.*

In our case, the positive Weyl chamber of $U(A)$ consists of diagonal matrices whose diagonal entries are in nonincreasing order (every matrix in the image of Φ has the same spectrum as one such matrix). Kirwan's theorem thus allows us to conclude that the set of all spectra of matrices obtainable by taking the partial traces of matrices with a fixed spectrum must be a region bounded by a finite set of inequalities.

Interestingly, Horn's problem can also be viewed in this framework. Recall that Horn's problem asks for the possible spectra of $X + Y$, given the spectra of $n \times n$ matrices X and Y. Suppose that λ is the spectrum of X and μ is the spectrum of Y. Now we consider the action of the group $U(n)$ of $n \times n$ unitary matrices on the symplectic manifold $\mathcal{O}_\lambda \times \mathcal{O}_\mu$ by diagonal conjugation:

$$U : (X, Y) \mapsto (UXU^\dagger, UYU^\dagger). \tag{9.41}$$

This is a Hamiltonian group action whose moment map takes two Hermitian matrices to their sum. Thus, Horn's problem can be viewed as the problem of determining the image of this moment map.

The following theorem [47] was motivated by the desire to generalize Klaychko's solution to Horn's problem. Before we state it, we give some notation. Let K be a compact connected Lie group, and let $\tilde{K}$ be a closed connected subgoup. Let f be the inclusion map of $\tilde{K}$ into K, $f_* : \tilde{\mathfrak{k}} \to \mathfrak{k}$ be the embedding of Lie algebras induced by f, and $f^* : \mathfrak{k}^* \to \tilde{\mathfrak{k}}^*$ be the dual projection. Choose maximal tori T of K and $\tilde{T}$ of $\tilde{K}$, and Weyl chambers $\mathfrak{t}^*_+ \subset \mathfrak{t}^*$ and $\tilde{\mathfrak{t}}^*_+ \subset \tilde{\mathfrak{t}}^*$, where $\mathfrak{t}$ and $\tilde{\mathfrak{t}}$ are the Lie algebras of T and $\tilde{T}$, respectively. For $\alpha \in \mathfrak{t}^*_+$, let $\Delta(\mathcal{O}_\alpha) = f^*(\mathcal{O}_\alpha) \cap \tilde{\mathfrak{t}}^*_+$. Let $\mathcal{C}$ be the cone spanned by the simple roots of $\mathfrak{t}^*$. Let W and $\tilde{W}$ be the Weyl groups of K and $\tilde{K}$ respectively. Let ϕ be the embedding of the flag variety $\tilde{K}/\tilde{T}$ into the flag variety K/T which is induced by the map f. We now state the main result from [47]:

Theorem 9.5.2 *Let $(\tilde{\alpha}, \alpha) \in \tilde{\mathfrak{t}}^*_+ \times \mathfrak{t}^*_+$. Then $\tilde{\alpha} \in \Delta(\mathcal{O}_\alpha)$ if and only if*

$$\tilde{w}^{-1}\tilde{\alpha} \in f^*(w^{-1}\alpha - v\mathcal{C}) \tag{9.42}$$

for all triples $(\tilde{w}, w, v) \in \tilde{W} \otimes W \otimes W_{rel}$ such that $\phi^*(v\sigma_{wv})(\tilde{c}_{\tilde{w}}) \neq 0$.

(Here W_{rel} is the *relative Weyl set*, defined in [47]. We shall not be concerned with the details of its description; it is equal to $\{1\}$ for our case.) For any $w \in W$, $f^*(w^{-1}\lambda - \mathcal{C})$ is a polyhedral cone in $\tilde{\mathfrak{t}}^*$, so Equation 9.42 represents a finite number of inequalities. The theorem gives us inequalities whenever the condition $\phi^*(v\sigma_{wv})(\tilde{c}_{\tilde{w}}) \neq 0$ is satisfied, where σ_{wv} is the element of the cohomology of the flag variety labelled by Weyl group element wv, and $\tilde{c}_{\tilde{w}}$ is the element of the homology of the flag variety labelled by $\tilde{w}$. This is equivalent to the condition that $\tilde{\sigma}_{\tilde{w}}$ appears in $\phi^*(\sigma_w)$, remembering that $v = 1$ for us.

We review some facts about the cohomology of flag varieties of a complex vector space V [41]. Fix a flag $F_\bullet$ of V. The cohomology classes σ_w, known as *Schubert classes*, are indexed by elements of S_n, where $n = \dim V$. For $w \in S_n$, σ_w corresponds to the class of the *Schubert variety* X_w, which is the closure of the *Schubert cell*

$$\Omega_w = \{E_\bullet \in F\ell(V) | \dim(E_p \cap F_q) = \#\{i \leq p : w(i) \leq q\} \text{ for } 1 \leq p, q \leq m\}. \quad (9.43)$$

Let us specialize to the case of our problem of finding the spectrum of a partial trace. For this case $f^*(\mathcal{C}) = \tilde{\mathcal{C}}$. If $\tilde{\sigma}_{\tilde{w}}$ appears in $\phi^*(\sigma_w)$, Equation 9.42 tells us that

$$f^*(w^{-1}\alpha) - \tilde{w}^{-1}\tilde{\alpha} \in \tilde{\mathcal{C}} \quad (9.44)$$

for elements of the dual space $\alpha \in \mathfrak{t}^*_+, \tilde{\alpha} \in \tilde{\mathfrak{t}}^*_+$. These functionals α, $\tilde{\alpha}$ act on the spectra λ, $\tilde{\lambda}$; we have

$$(w^{-1}\alpha)(\lambda) = \alpha(w^{-1}(\lambda)) = \alpha(\lambda_{w(1)}, \lambda_{w(2)}, \ldots, \lambda_{w(n)}). \quad (9.45)$$

Identifying $\mathfrak{t}$ and $\tilde{\mathfrak{t}}$ with their dual spaces, we have the conditions that

$$f^*(\lambda_{w(1)}, \lambda_{w(2)}, \ldots, \lambda_{w(d_A d_B)}) - (\lambda_{\tilde{w}(1)}, \lambda_{\tilde{w}(2)}, \ldots, \lambda_{\tilde{w}(d_A)}) \in \tilde{\mathcal{C}} \quad (9.46)$$

whenever $\tilde{\sigma}_{\tilde{w}}$ appears in $\phi^*(\sigma_w)$. But the root cone $\mathcal{C}$ is generated by the simple roots $\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \ldots, \lambda_{d_A - 1} - \lambda_{d_A}$ where $\lambda_i \geq \lambda_{i+1}$; in order words, $\mathcal{C}$ is generated by the set of μ such that

$$\sum_{i=1}^{k} \mu_i \geq 0, \text{ for } k < d_A, \quad (9.47)$$

and

$$\sum_{i=1}^{d_A} \mu_i = 0. \quad (9.48)$$

So our conditions are that

$$(0, 0, \ldots, 0) \prec f^*(\lambda_{w(1)}, \lambda_{w(2)}, \ldots, \lambda_{w(d_A d_B)}) - (\lambda_{\tilde{w}(1)}, \lambda_{\tilde{w}(2)}, \ldots, \lambda_{\tilde{w}(d_A)}), \quad (9.49)$$

or

$$(0,0,\ldots,0) \prec$$
$$(\lambda_{w(1)}+\ldots+\lambda_{w(d_B)}, \lambda_{w(d_B+1)}+\ldots+\lambda_{w(2d_B)}, \ldots, \lambda_{(w((d_A-1)d_B+1)}+\ldots+\lambda_{w(d_Ad_B)})$$
$$-(\lambda_{\tilde{w}(1)}, \lambda_{\tilde{w}(2)}, \ldots, \lambda_{\tilde{w}(d_A)}).$$

This is turn yields $(d_A - 1)$ inequalities:

$$\sum_{i=1}^{d_B} \lambda_{w(i)} \leq \tilde{\lambda}_{\tilde{w}(1)}, \tag{9.50}$$
$$\sum_{i=1}^{2d_B} \lambda_{w(i)} \leq \tilde{\lambda}_{\tilde{w}(1)} + \lambda_{\tilde{w}(2)}, \tag{9.51}$$
$$\vdots \tag{9.52}$$
$$\sum_{i=1}^{(d_A-1)d_B} \leq \sum_{i=1}^{d_A-1} \tilde{\lambda}_{\tilde{w}(i)}. \tag{9.53}$$

These inequalities arise from intersections of Schubert cells of the flag varieties. We show that any such inequality can be obtained as a consequence of an intersection of Grassmannian Schubert varieties. Choose a flag variety $F_\bullet$ of $A \otimes B$ corresponding to the eigenspaces of ρ_{AB} arranged in nonincreasing order of eigenvalues, and a flag variety $\tilde{F}_\bullet$ of A corresponding to the eigenspaces of ρ_A arranged in nonincreasing order of eigenvalues. Now define π_w to be the binary string of length d_Ad_B such that

$$\pi_w(i) = 1 \quad \text{if } w^{-1}(i) \leq kd_B,$$
$$\pi_w(i) = 0 \quad \text{otherwise.}$$

Similarly, define $\tilde{\pi}_{\tilde{w}}$ to be the binary string of length d_A which takes on the value 1 only at those positions i such that $\tilde{w}^{-1}(i) \leq k$.

Now consider any inequality of the form

$$\sum_{i=1}^{kd_B} \lambda_{w(i)} \leq \sum_{i=1}^{k} \tilde{\lambda}_{\tilde{w}(i)}, \tag{9.54}$$

for some permuations w and $\tilde{w}$, arising from the intersection of $\Omega_w(F_\bullet)$ and $\phi(\Omega_{\tilde{w}}(\tilde{F}_\bullet))$. Suppose $E_\bullet \in \Omega_w(F_\bullet)$ and $\tilde{E}_\bullet \in \Omega_{\tilde{w}}(\tilde{F}_\bullet)$, such that $\phi(\tilde{E}_\bullet) = E_\bullet$. Therefore, the subspaces E_{nk} and $\tilde{E}_k$ satisfy $\phi(\tilde{E}_k) = E_{nk}$. Note that $E_{nk} \in \Omega_{\pi_w}(F_\bullet)$, and $\tilde{E}_k \in \Omega_{\pi_{\tilde{w}}}(\tilde{F}_\bullet)$, where $\Omega_{\pi_w}(F_\bullet)$ and $\Omega_{\pi_{\tilde{w}}}(\tilde{F}_\bullet)$ are Grassmannian Schubert cells. Therefore, we have a nonempty intersection $\Omega_{\pi_w}(F_\bullet) \cap \phi(\Omega_{\pi_{\tilde{w}}}(\tilde{F}_\bullet)) \neq \emptyset$, which by Theorem 6.2.2 yields the same inequality

$$\sum_{i=1}^{kd_B} \lambda_{w(i)} \leq \sum_{i=1}^{k} \tilde{\lambda}_{\tilde{w}(i)}. \tag{9.55}$$

Thus, considering only Grassmannian intersections is enough to derive any inequality of Theorem 9.5.2 applied to our problem. So the inequalities derived by the approach we have described in Part II of this thesis are indeed sufficient.

9.6 Saturation

Having determined how to find the inequalities relating ρ_{AB} and ρ_A, we seek methods of simplifying the list of inequalities. It turns out that the inequalities governing the relationship between the spectra of ρ_{AB} and of ρ_A are particularly simple when d_B is large compared to d_A. In this section we will show that if $d_B \geq \frac{1}{2}d_A^2$, then the basic inequalities are sufficient (all other inequalities are redundant). Physically, thinking in terms of a quantum communication protocol where Alice sends d_B qubits to Bob, such a result is plausible because a large amount of communication gives Alice a great deal of freedom in manipulating her portion of the system, so we should not expect there to be much restriction in the states she might end up with.

Suppose that $d_B \geq \frac{1}{2}d_A^2$, and consider an arbitrary inequality resulting from the nonzero cup product $\tilde{\sigma}_\nu \cup \phi^*(\sigma_\pi) \neq 0$. (As discussed in Section 9.1, we may assume that $\tilde{\sigma}_\nu$ is a summand in the expansion of $\phi^*(\sigma_\pi) \neq 0$ as a sum of Schubert classes.) Such an inequality is of the form

$$\sum_{i\in I} \tilde{\lambda}_i \leq \sum_{j\in J} \lambda_j \tag{9.56}$$

where if $|I| = k$, then $|J| = d_B k$. As in Section 9.2, we may assume that $k \leq \frac{d_A}{2}$. Consider the partitions π and ν in the equation $\tilde{\sigma}_\nu \cup \phi^*(\sigma_\pi) \neq 0$ to be binary strings. Let u be the $(0,1)$ vector of length $d_A d_B$, whose ith component is equal to 1 if and only if $\pi(i) = 1$. Similarly, let $\tilde{u}$ be the $(0,1)$ vector of length d_A, whose ith component is equal to 1 if and only if $\nu(i) = 1$. Then Inequality 9.56 can be rewritten as

$$\tilde{\lambda} \cdot \tilde{u} \leq \lambda \cdot u. \tag{9.57}$$

We now prove some facts about this situation, ending with our desired result.

Observation 9.6.1 *The Young diagram corresponding to π can't have more than $(\frac{d_A}{2})^2$ boxes.*

This follows because the Young diagram corresponding to ν must fit in a $k \times (d_A - k)$ rectangle, and so cannot have more than $(\frac{d_A}{2})^2$ boxes; and π must have the same number of boxes in its Young diagram as ν. □

Observation 9.6.2 *If $u \prec u'$, then $\lambda \cdot u \leq \lambda \cdot u'$.*

This follows easily from the fact that λ has its terms arranged in nonincreasing order. □

Claim 9.6.3 *If $j > d_B k + (\frac{d_A}{2})^2$, then $j \notin J$ in Inequality 9.56 (in other words, λ_j is not one of the terms in the right hand sum).*

Proof If $j \in J$, then the Young diagram corresponding to π would have more than $(\frac{d_A}{2})^2$ boxes in its jth row. □

Claim 9.6.4 *The first zero of u can't appear before the $(d_B k - \lfloor (\frac{d_A}{2})^2 \rfloor)$th component. In other words, if $j \leq d_B k - (\frac{d_A}{2})^2$, then $j \in J$ in Inequality 9.56.*

Proof Otherwise, the Young diagram corresponding to π would have more than $(\frac{d_A}{2})^2$ rows. □

Lemma 9.6.5

$$(\underbrace{1,\ldots,1}_{d_Bk-\lfloor(\frac{d_A}{2})^2\rfloor},\underbrace{0,\ldots,0}_{\lfloor(\frac{d_A}{2})^2\rfloor},\underbrace{1,\ldots,1}_{\lfloor(\frac{d_A}{2})^2\rfloor},0,\ldots,0)\prec u. \tag{9.58}$$

Consequently, since $d_B \geq \frac{d_A^2}{2}$,

$$(\underbrace{1,\ldots,1}_{d_Bk-\lfloor\frac{d_B}{2}\rfloor},\underbrace{0,\ldots,0}_{\lfloor\frac{d_B}{2}\rfloor},\underbrace{1,\ldots,1}_{\lfloor\frac{d_B}{2}\rfloor},0,\ldots,0)\prec u. \tag{9.59}$$

Proof This follows from Claims 9.6.3 and 9.6.4. □

Theorem 9.6.6 *If* $d_B \geq \frac{1}{2}d_A^2$*, then Inequality 9.57 is redundant. In other words, the basic inequalities are sufficient to characterize the relationship between the spectrum of* ρ_{AB} *and the spectrum of* ρ_A*.*

Proof It is sufficient to assume that $\tilde{u} \prec (\underbrace{1,\ldots,1}_{k-1},0,1,0,\ldots,0)$ (the only possible $\tilde{u}$ that does not satisfy this condition is $\tilde{u} = (\underbrace{1,\ldots,1}_{k},0,\ldots,0)$, which gives rise to the basic inequalities). Then we have

$$
\begin{aligned}
\tilde{\lambda} \cdot \tilde{u} &\leq \lambda \cdot (\underbrace{1, \ldots, 1}_{k-1}, 0, 1, 0, \ldots, 0) \\
&= \sum_{i=1}^{k-1} \tilde{\lambda}_i + \tilde{\lambda}_{k+1} \\
&\leq \frac{1}{2}\left[\sum_{i=1}^{k-1} \tilde{\lambda}_i + \tilde{\lambda}_k + \tilde{\lambda}_{k+1} + \sum_{i=1}^{k-1} \tilde{\lambda}_i\right] \\
&= (\underbrace{1, \ldots, 1}_{k-1}, \frac{1}{2}, \frac{1}{2}, 0, \ldots, 0) \cdot \tilde{\lambda} \\
&= \frac{1}{2}(\underbrace{1, \ldots, 1}_{k-1}, 0, \ldots, 0) \cdot \tilde{\lambda} + \frac{1}{2}(\underbrace{1, \ldots, 1}_{k+1}, 0, \ldots, 0) \cdot \tilde{\lambda} \\
&\leq \frac{1}{2}(\underbrace{1, \ldots, 1}_{d_B(k-1)}, 0, \ldots, 0) \cdot \lambda + \frac{1}{2}(\underbrace{1, \ldots, 1}_{d_B(k+1)}, 0, \ldots, 0) \cdot \lambda \text{ by the basic inequalities} \\
&= (\underbrace{1, \ldots, 1}_{d_B(k-1)}, \underbrace{\frac{1}{2}, \ldots, \frac{1}{2}}_{2d_B}, 0, \ldots, 0) \cdot \lambda \\
&\leq (\underbrace{1, \ldots, 1}_{d_B k - \lfloor \frac{d_B}{2} \rfloor}, \underbrace{0, \ldots, 0}_{\lfloor \frac{d_B}{2} \rfloor}, \underbrace{1, \ldots, 1}_{\lfloor \frac{d_B}{2} \rfloor}, 0, \ldots, 0) \cdot \lambda.
\end{aligned}
$$

But the right hand side of Inequality 9.57 must be greater than equal to

$$(\underbrace{1, \ldots, 1}_{d_B k - \lfloor \frac{d_B}{2} \rfloor}, \underbrace{0, \ldots, 0}_{\lfloor \frac{d_B}{2} \rfloor}, \underbrace{1, \ldots, 1}_{\lfloor \frac{d_B}{2} \rfloor}, 0, \ldots, 0) \cdot \lambda,$$

by Lemma 9.6.5 and Observation 9.6.2. Thus, we have shown that Inequality 9.57 must hold, assuming only the basic inequalities; so this inequality must be redundant, for an arbitrary inequality arising from $\tilde{\sigma}_\nu \cup \phi^*(\sigma_\pi) \neq 0$. □

We conjecture a stronger result, which we have verified for $d_A = 2, 3$, and 4 (the cases $d_A = 2$ and $d_A = 3$ have been shown explicitly in this thesis).

Conjecture 9.6.7 *If $d_B \geq d_A$, then the basic inequalities are sufficient to characterize the relationship between the spectrum of ρ_{AB} and the spectrum of ρ_A.*

Bibliography

[1] M. A. Nielsen. A partial order on the entangled states. *Phys. Rev. Lett.* **83**, Number 2, 436–439, 1999.

[2] M. A. Nielsen and J. Kempe. Separable states are more disordered globally than locally. *Phys. Rev. Lett.* **86**, 5184 (2001).

[3] M. A. Nielsen. Characterizing mixing and measurement in quantum mechanics. Phys. Rev. A **63**, 022114 (2001).

[4] R. Orus, J. I. Latorre, and M. A. Martin-Delgado. Systematic analysis of majorization in quantum algorithms. arXiv e-print quant-ph/0212094.

[5] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York, 1973.

[6] M. A. Nielsen. Majorization and its applications to quantum information theory. *http://www.theory.caltech.edu/˜mnielsen/info/majorize.html.* 1999.

[7] M. O. Lorenz. Methods of measuring concentration of wealth. *J. Amer. Statist. Assoc.* **9**, 209–219, 1905.

[8] E. C. Pigou. *Wealth and Welfare.* Macmillan, New York, 1912.

[9] H. Dalton. The measurement of the inequality of incomes. *Econom. J.* bf 30, 348—361, 1920.

[10] A. M. Ostrowski. Sur quelqes applications des foctions convexes et concaves au sense de I. Schur. *J. Math Pures Appl.* [9] **31**, 253–292.

[11] G. H. Hardy, J. E. Littlewood, and G. Pólya. Some simple inequalities satisfied by convex functions. *Messenger Math.* **58**, 145–152.

[12] J. H. van Lint and R. M. Wilson. *A Course in Combinatorics.* Cambridge University Press, 1992.

[13] S. Daftuar and M. Klimesh. Mathematical structure of entanglement catalysis. *Phys. Rev. A* **64**, 042314 (2001).

[14] C. H. Bennett and G. Brassard. Quantum cryptography: Public-key distribution and coin-tossing. *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India, 1984*, 175–179.

[15] C. H. Bennett and G. Brassard. Quantum public key distribution. IBM Technical Disclosure Bulletin **28**, 3153–3163, 1985.

[16] C. H. Bennett, G. Brassard, C. Crepeau, R. Josza, A. Peres, and W. K. Wooters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899, 1993.

[17] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen state. *Phys. Rev. Lett.* **69**(20), 2881–2884, 1992.

[18] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2000.

[19] J. Preskill. Physics 229: Advanced mathematical methods of physics — Quantum computation and information. *http://www.theory.caltech.edu/people/preskill/ph229/#lecture*. 1998.

[20] D. Jonathan and M. B. Plenio. Entanglement-assisted local manipulation of pure quantum states. *Phys. Rev. Lett.* **83**, 3566 (1999).

[21] G. Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.* **83**, 1046 (1999).

[22] P. H. Anspach. Two-qubit catalysis in a four-state pure bipartite system. arXiv e-print quant-ph/0102067.

[23] G. Hardy, J. E. Littlewood, and G Pólya. *Inequalities.* Cambridge University Press, 1952.

[24] A. W. Zhou and G. C. Guo. Basic limitations for entanglement catalysis. arXiv e-print quant-ph/0005005.

[25] M. A. Nielsen and G. Vidal. Majorization and the interconversion of the bipartite states. *Quantum Information and Computation*, Vol. 1, No. 1 (2001) 76–93.

[26] R. Bhatia. *Matrix Analysis.* Springer-Verlag, New York, 1997.

[27] W. van Dam and P. Hayden. Rényi-entropic bounds on quantum communication. arXiv e-print quant-ph/020093.

[28] H. Weyl. Das asymtotische Verteilungsgesetz de Eigenwerte lineare parieller Differentialgleichungen. *Math. Ann.* **71** (1912), 441–479.

[29] A. A. Klyachko. Stable bundles, representation theory and Hermitian operators. *Selecta Math.* (1998), 419–445.

[30] A. A. Klyachko. Random walks on symmetric spaces and inequalities for matrix spectra. preprint, 1999.

[31] B. Totaro. Tensor products of semistables are semistable. *Geometry and analysis on complex manifolds.* World Sci. Publ., 1994, 242–250.

[32] A. Knutson and T. Tao. The honeycomb model of $GL_n(\mathbb{C})$ tensor products I: proof of the saturation conjecture. *Journal of the AMS*, 12 (1999), no. 4, 1055–1090.

[33] A. Knutson, T. Tao, and C. Woodward. The honeycomb model of $GL_n(\mathbb{C})$ tensor products II: puzzles determine facets of the Littlewood-Richardson cone. *Journal of the AMS*, to appear.

[34] P. Belkale. *Local systems on $\mathbb{P}^1 \backslash S$ for S a finite set.* Ph.D. thesis, University of Chicago, 1999.

[35] A. Horn. Eigenvalues of sums of Hermitian matrices. *Pacific J. Math.* **12** (1962), 225–241.

[36] B. V. Lidskii. Spectral polyhedron of a sum of two Hermitian matrices. *Functional Analysis and Appl.*, **10** (1982), 76–77.

[37] J. Hersch and B. Zwahlen. Évaluations par défaut pour une summe quelconque de valeurs propers γ_k d'un opérateur $C = A + B$, a l'aide de valuers propres α_i de A et β_j de B. *C. R. Acad. Sc. Paris* **254** (1962), 1559-1561.

[38] S. Johnson. *The Schubert calculus and eigenvalue inequalities for sums of Hermitian matrices.* Ph. D. thesis, University of California, Santa Barbara, 1979.

[39] U. Helmke and J. Rosenthal. Eigenvalue inequalities and Schubert calculus. *Math. Nachr.* **171** (1995), 207–225.

[40] L. Manivel. *Symmetric Functions, Schubert Polynomials and Degeneracy Loci.* Translated by J. R. Swallow. American Mathematical Society, 2001; Société Mathématique de France, 1998.

[41] W. Fulton. *Young Tableaux.* Cambridge University Press, 1997.

[42] P. Griffiths and J. Harris. *Principles of Algebraic Geometry.* John Wiley & Sons, New York, 1978.

[43] H. M. Edwards. *Galois Theory.* Springer-Verlag, New York, 1984.

[44] R. Bott and L. W. Tu. *Differential Forms in Algebraic Topology.* Springer-Verlag, New York, 1982.

[45] A. Hatcher. *Vector Bundles and K-Theory.* Incomplete text, available at http://www.math.cornell.edu/ hatcher/VBKT/VBpage.html.

[46] W. Fulton and J. Harris. *Representation Theory: A First Course.* Springer-Verlag, New York, 1991.

[47] A. Berenstein and R. Sjamaar. Coadjoint orbits, moment maps, and the Hilbert-Mumford criterion. arXiv e-print math.SG/9810125.

[48] A. Knutson. The symplectic and algebraic geometry of Horn's problem. *Linear Algebra and its Applications* **319** (2000), no. 1-3, 61–81.

[49] A. C. da Silva. *Lectures on Symplectic Geometry.* Springer-Verlag, Berlin, 2001.

[50] W. Fulton. Eigenvalues, invariant factors, highest weights, and Schubert calculus. *Bull. Amer. Math. Soc.* **37** (2000), 209–249.