Ground State Properties of Spin-Lattice Models

An Application of the Perron-Frobenius Theorem

Diploma Thesis
of

Rudolf A. Römer

at the
Freie Universität Berlin,
Fachbereich Physik

October 1991

Advisor: Prof. Dr. R. Schrader

Zusammenfassung

In der vorliegenden Arbeit wird die Eindeutigkeit des Grundzustandes von Heisenberg, Hubbard und tJ-Modell mit Hilfe des Perron-Frobenius Theorems untersucht. Halbgruppenmethoden und Verallgemeinerungen der klassischen Perron-Frobenius Theorie werden benutzt, um in Anlehnung an Arbeiten von Faris [F 72] und Glimm-Jaffe [GJ 70], die Eindeutigkeitsbeweise von Lieb-Mattis zu vereinfachen [LM 62a, LM 62b, L 89].

Abstract

This thesis investigates the uniqueness of ground states in the Heisenberg, Hubbard, and the tJ-model using the Perron-Frobenius theorem. Following ideas of Glimm-Jaffe [GJ 70] and Faris [F 72], semigroup methods and generalizations of the classical Perron-Frobenius theory are used to simplify the uniqueness theorems of Lieb-Mattis [LM 62a, LM 62b, L 89].

Ground State Properties of Spin-Lattice Models

An Application of the Perron-Frobenius Theorem

Diploma Thesis
of

Rudolf A. Römer

at the
Freie Universität Berlin,
Fachbereich Physik

October 1991

Advisor: Prof. Dr. R. Schrader

Contents

Chapter 1

Introduction

In this thesis, we investigate the existence and uniqueness of ground states for a number of models of strongly correlated electron systems. Interest in these models has been greatly stimulated by the discovery of high-T_c superconductors in 1986 by J. G. Bednorz and K. A. Müller[1]. Since then the search for an explanation and theoretical description as well as the study of the properties and possible uses of these materials has resulted in numerous publications and conferences [PT91].

The starting point for most theories of strongly correlated electron systems is usually the one-band Hubbard model. This is a model in which band electrons interact via a two-body repulsive Coulomb interaction and no phonons are present. For this reason, the Hubbard model has been associated with magnetism. Superconductivity, on the other hand, has been traditionally interpreted as instability of the ground state because of the effectively attractive electron-phonon interactions of the standard BCS theory[2].

A novel situation has arisen with Anderson's suggestion [A87] that in the new high-T_c materials, superconductivity may originate from purely repulsive interactions. This suggestion was motivated by the observation that superconductivity occurs at an insulator-metal transition and thus seems to originate from doping (that is, extracting or adding charges to) the insulating state.

Two other models are also of interest in this context. The first one is the general (quantum) Heisenberg model which describes the interaction of localized magnetic moments. The second one is a direct outgrowth of detailed studies of superconducting materials and their respective electronic orderings as well as the ideas of Anderson [ZR 88]. It is called the tJ-model and may be viewed as a merger of the localized behavior of the Heisenberg model and the bandlike behavior of the Hubbard model. For all these models, no rigorous solution exists except for some special cases [LW 68]. This is largely due to the fact that neither the

[1] J. G. Bednorz, K. A. Müller, Z. Phys. B **64**, 189, (1986)

[2] J. Bardeen, L. N. Cooper, J. R. Schrieffer, Phys. Rev. **108**, 1175, (1957)

(quantum) Heisenberg [S 85] nor the Hubbard model exhibit *reflection positivity*[3]; this inhibits the usual proof technique of Fröhlich-Simon-Spencer [FSS 76] and Dyson-Lieb-Simon [DLS 78]. For example, currently there exists no proof of a nonzero spontaneous magnetization in the (quantum) Heisenberg model. Only for the classical model and the $S \geq 1$ antiferromagnet does such a proof exist.

It is thus fortunate that there is a (rather old) mathematical theory which may be used to obtain rigorous and non-perturbative information about the ground state–and sometimes about the ordering of *all* electronic energy levels as well–of these models. At the beginning of this century O. Perron and G. Frobenius studied the spectra of nonnegative matrices [P 07, F 08, F 12]. They showed that for these matrices–mapping nonnegative vectors into nonnegative vectors–the spectral radius is an eigenvalue. Assuming an additional property, for which Frobenius coined "unzerlegbar" and which we will denote as *irreducible*, they then proved that the corresponding eigenvector is unique.

Krein-Rutman [KR 48] and various other authors (see the introductory remarks in chapter 3 for a more complete list) have subsequently generalized this concept to operators which leave cones in Banach spaces invariant. For Hilbert spaces and quantum theory, the applications of Perron-Frobenius theory have developed in two directions. For quantum field theory, the starting point is a 1970 publication by Glimm-Jaffe [GJ 70]. In this as well as in a number of subsequent papers [F 72, GJ 87, and references therein], they used Perron-Frobenius theory in conjunction with semigroup methods to establish the existence and uniqueness of physical ground states.

As early as 1962, D. Mattis and E. Lieb applied the– essentially classical–Perron-Frobenius theorem to models of statistical physics [LM 62a, LM 62b]. Besides proving uniqueness of the ground state, they were able even to specify the value of the total spin. This resulted in a proof that the antiferromagnetic Heisenberg model indeed has $S = 0$. In addition, they showed that on a linear chain of atoms ferromagnetism must result explicitly from either spin or velocity dependent forces. This last result is now known as the *Lieb-Mattis theorem.*

The impetus to this thesis was given by a recent publication by E. Lieb [L 89]. He showed that for the Hubbard model, uniqueness of the ground state may be assured when using a nonstandard cone. He called this property *spin-reflection positivity* in analogy to the reflection positivity mentioned above. For an attractive on-site interaction the total spin value equals zero, but for the repulsive case and for the half-filled band, the spin value may be nonzero. Thus the Hubbard model exhibits a provable itinerant-electron ferromagnetism[4].

---

[3]K. Osterwalder, R. Schrader, Comm. Math. Phys. **31**, 83, (1973), **42**, 281, (1975)

[4]Here ferromagnetism is used in the sense that the spin is an extensive quantity. No spatial ordering is implied.

In the present work we use the ideas of Lieb-Mattis and the semigroup methods of Glimm-Jaffe to give slightly alternative proofs of the uniqueness of the ground state for the linear chain of itinerant electrons, the general Heisenberg model, the Hubbard model and the tJ-model. None of these proofs have been given in conjunction with semigroup methods before. The application of Perron-Frobenius arguments for the tJ-model is new and has not previously appeared in the literature.

The thesis is organized as follows:

In chapter 2, we review the existence and uniqueness theorems of the classical Perron-Frobenius theory for positive matrices as established by Perron. Then we generalize these results to nonnegative matrices. Irreducibility of matrices is introduced to insure uniqueness.

In chapter 3, we investigate the possible generalizations of the classical Perron-Frobenius theory to Banach and Hilbert spaces. We first introduce the notion of a cone and study its geometry. Then in section 3.2 we study matrices that leave a cone invariant. In section 3.3 we review a Perron-Frobenius theory for Banach spaces and investigate so-called minihedral cones. An existence theorem is given. Positive operators in Hilbert spaces are studied in section 3.4. Here the important concept of an ergodic operator, closely related to irreducible matrices, is introduced. We show that for ergodic operators a uniqueness theorem holds. At the end of this chapter, we present possible generalizations of the concept of irreducibility.

The idea that enabled Lieb-Mattis to apply the Perron-Frobenius theorem to ground states of spin-lattice models is presented in chapter 4. Semigroups are introduced and three 'perturbation' results for Hamiltonians of the form $H = H_0 - V$ are derived.

In chapter 5 we apply the results of the previous chapters to the Lieb-Mattis model and the Heisenberg model. Canonical transformations are used to rearrange the Hamiltonians in the $H = H_0 - V$ order. The total spin value for the ground state of both models is derived.

Chapter 6 is devoted to an extensive inquiry into the application of the Perron-Frobenius theory for the Hubbard model. After a short review of the model, we derive the equivalent matrix problem. The no-coupling limit is treated first, then the attractive and the repulsive Hubbard models are studied. We show that, whereas the attractive model explicitly exhibits a spin-reflection positivity, the repulsive model does so only after a hole-particle transformation which can be applied using the assumption that the interaction parameter is a constant. We then add a one-body potential so that the resulting repulsive Hamiltonian once again exhibits spin-reflection positivity. The total spin value is determined by the limiting behavior of the strong-coupling Hubbard model.

Finally, in chapter 7, we apply the ideas of the previous chapter to the tJ-model on a linear chain.

Notation For the reader's convenience, a notational guide is provided in the appendix. In general, standard notation is used. *Emphasized* passages (a) indicate that a term not previously encountered is introduced or (b) denote especially important statements. Single quotes ' ' label a slightly misused term such as 'perturbation' in chapter 4. Double quotes enclose a quotation.

Two different citation symbols are used. The brackets [ ] denote a publication that is to be found in the references. Footnotes are sometimes used to give references that are of a more general character and are given for convenience and completeness only.

Chapter 2

Fundamental Concepts and Results of the Theory of Nonnegative Matrices

In this chapter, we will review the fundamental concepts and theorems of the theory of nonnegative matrices as established by O. Perron [P 07] and G. Froebenius [F 08, F 12] at the beginning of the century. Our review will be rather brief since this material can now be found in most text books on matrix analysis. The classical source is still Gantmacher's book on matrix theory [G 59], but somewhat more modern accounts are given by Seneta [S 73] and Horn-Johnson [HJ 85]. We will largely follow the latter in our notation and in the exposition of the material presented.

2.1 Preliminaries

We recall some of the basic definitions and notations of matrix theory to be used throughout the following chapters.

We denote the n-dimensional vector space over the field $\mathbf{F}$, usually the real numbers R or the complex numbers C, by $\mathbf{F}^n$. The set of real-entried (or complex-entried, respectively) *vectors* is thus denoted by R^n (C^n), both interpreted as column vectors. By $M_{m,n}(\mathbf{F})$ we denote the $m \times n$ *matrices* over $\mathbf{F}$. We think of a matrix $A \in M_{m,n}(\mathbf{F})$ primarily as a linear transformation of $\mathbf{F}^n$ into $\mathbf{F}^m$ (with respect to given bases in $\mathbf{F}^n$ and $\mathbf{F}^m$), but is is also useful to think of it as an array of numbers $[a_{ij}]$. If $m = n$, the matrix is said to be *square* and $M_{n,n}(\mathbf{F})$ is abbreviated to $M_n(\mathbf{F})$. In most cases in which $\mathbf{F} = \mathsf{C}$, $M_n(\mathsf{C})$ is further abbreviated to M_n and $M_{m,n}(\mathsf{C})$ to $M_{m,n}$.

The *transpose* of $A = [a_{ij}] \in M_{m,n}(\mathbf{F})$ is the matrix $[a_{ji}] \in M_{n,m}(\mathbf{F})$, denoted

by A^T, and, if $\mathbf{F} \subseteq \mathsf{C}$, the *Hermitian adjoint* is the conjugate transpose $[\bar{a}_{ji}]$ of A, denoted by A^*. Similarly, if $x \in \mathbf{F}^n$, x^T denotes the row vector with the same entries as x, and, if $\mathbf{F} \subseteq \mathsf{C}$, x^* denotes the row vector whose entries are the complex conjugates of those of x. Here the overbar $\bar{\cdot}$ denotes the complex conjugate of a complex scalar or the component-wise complex conjugate of a vector or matrix. The *absolute value* of some vector $x \in \mathsf{C}^n$ (matrix $A \in M_{m,n}(\mathsf{C})$) denotes the vector (matrix) formed by taking the absolute value of each coordinate (entry), i.e. $|x| = [|x_i|]$ ($|A| = [|a_{ij}|]$). The matrix $D = [d_{ij}] \in M_n$ is called *diagonal* if $d_{ij} = 0$ whenever $j \neq i$. By convention, we denote such a matrix by $D = diag(d_1, d_2, \ldots, d_n)$. A *permutation matrix* P is a square matrix in which all entries are 0 or 1; in each row and column of P there is precisely one 1. The matrix $I(A)$ is called the *indicator matrix of* $A \in M_n$ and is given by $I(A)_{ij} = 1$ iff $a_{ij} \neq 0$.

In the next two definitions, we recall the concept of a spectrum $\sigma(A)$ and its accompanying eigenvectors.

Definition (2.1.1) If $A \in M_n$ and $x \in \mathsf{C}^n$, we consider the equation

$$Ax = \lambda x, \; x \neq 0 \tag{2.1.2}$$

where λ is a scalar. If a scalar λ and a nonzero vector x happen to satisfy this equation, then λ is called the *eigenvalue* of A and x is called the *eigenvector* of A, associated with λ.

Definition (2.1.3) The set of all $\lambda \in \mathsf{C}$ that are eigenvalues of $A \in M_n$ is called the *spectrum* of A and is denoted by $\sigma(A)$. The *spectral radius* of A is the nonnegative real number $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$. This is just the radius of the smallest disc centered at the origin in the complex plane that contains all the eigenvalues of A.

The classical Perron-Froebenius theory to be reviewed in this chapter deals basically with the spectra of matrices that leave invariant a special set of vectors. These vectors and matrices are defined next.

Definition (2.1.4) Given the vector $x = [x_i] \in \mathsf{C}^n$ and the matrix $A = [a_{ij}] \in M_n$. We say that $x \geq 0$ ($A \geq 0$) is a *nonnegative vector* (*nonnegative matrix*) if all coordinates x_i (entries a_{ij}) are real and nonnegative. If $x_i > 0$ ($a_{ij} > 0$) for all permissible coordinates (entries), then the vector x (matrix A) is called *positive*.

Note that $\geq$ induces a partial ordering for nonnegative vectors and matrices. Therefore we may order vectors and matrices, i.e. given two matrices $A, B \in M_{m,n}$, $A \geq B$ iff $a_{ij} \geq b_{ij}$ for all permissible i and j indices.

Before we proceed to the main results in the next sections, we will show some well-known facts regarding upper and lower bounds for the spectral radius $\rho(A)$ of a matrix $A \in M_n$.

Lemma (2.1.5) Let $A \in M_n$ and suppose that $A \geq 0$. If the row sums of A are constant, then $\rho(A) = ||A||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|$ (the *maximum row sum matrix norm*). If the column sums of A are constant, then $\rho(A) = ||A||_1 = \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|$ (*the maximum column sum matrix norm*).

Proof (2.1.6) By [HJ 85, theorem 5.6.9], $\rho(A) \leq ||A||_\infty$. Now, if the row sums are constant, $x = [1, \ldots, 1]^T$ is an eigenvector with eigenvalue $||A||_\infty$, and therefore $\rho(A) = ||A||_\infty$. The statement for column sums follows from applying the same argument to A^T, since $\|A^T\|_\infty = \|A\|_1$. ∎

Theorem (2.1.7) Let $A \in M_n$ and suppose $A \geq 0$. Then for any positive vector $x \in \mathbb{C}^n$ we have

$$\min_{1\leq i\leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j \leq \rho(A) \leq \max_{1\leq i\leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j \tag{2.1.8}$$

and

$$\min_{1\leq j\leq n} x_j \sum_{i=1}^n \frac{a_{ij}}{x_i} \leq \rho(A) \leq \max_{1\leq j\leq n} x_j \sum_{i=1}^n \frac{a_{ij}}{x_i} \tag{2.1.9}$$

Proof (2.1.10) We first prove that

$$\min_{1\leq i\leq n} \sum_{j=1}^n a_{ij} \leq \rho(A) \tag{2.1.11}$$

Let $\alpha = \min_{1\leq i\leq n} \sum_{j=1}^n a_{ij}$. We now want to construct a matrix B such that the row sum $\sum_{j=1}^n b_{ij} = \alpha$ for all $i = 1, 2, \ldots, n$. In addition, we require that $A \geq B \geq 0$. The construction is straightforward, i.e. if $\alpha = 0$, we could choose $B = 0$, and if $\alpha > 0$, we could set $b_{ij} = \alpha a_{ij}(\sum_{j=1}^n a_{ij})^{-1}$. By lemma 2.1.5 above, we then have $\rho(B) = \alpha$. Since $A \geq B$, we know by [HJ 85, Corollary 8.1.19] that $\rho(A) \geq \rho(B)$, and this establishes equation (2.1.11).

Now we generalize this result in order to obtain equation (2.1.8). Whenever S is invertible, $\rho(S^{-1}AS) = \rho(A)$ holds such that we may choose $S = diag(x_1, \ldots, x_n)$, and if all $x_i > 0$, then $A \geq 0$ implies $S^{-1}AS \geq 0$. Thus, substituting $A = [a_{ij}] \rightarrow S^{-1}AS = [a_{ij}x_jx_i^{-1}]$, we obtain equation (2.1.8).

In the same way, the upper bound can be established. As before, the column bounds follow from applying the arguments to A^T. ∎

Corollary (2.1.12) Let $A \in M_n$. If $A \geq 0$ and $\sum_{j=1}^n a_{ij} > 0$ for all $i = 1, 2, \ldots, n$, then $\rho(A) > 0$.

$A \geq 0$ alone is not sufficient for $\rho(A) > 0$ since if

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

then A is certainly nonnegative but $\rho(A) = 0$. On the other hand, a zero row sum alone is not sufficient for a zero spectral radius since

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

has spectral radius $\rho(B) = 1$. Note that, in particular, $\rho(A) > 0$ if $A > 0$.

Corollary (2.1.13) Let $A \in M_n$, let $x \in \mathsf{R}^n$, and suppose that $A \geq 0$ and $x > 0$. If $\alpha, \beta \geq 0$ are such that $\alpha x \leq Ax \leq \beta x$, then $\alpha \leq \rho(A) \leq \beta$. If $\alpha x < Ax$, then $\alpha < \rho(A)$; if $Ax < \beta x$, then $\rho(A) < \beta$.

Proof (2.1.14) If $\alpha x \leq Ax$ then $\alpha \leq \min_{1 \leq i \leq n} x_i^{-1} \sum_{j=1}^{n} a_{ij} x_j$. By theorem 2.1.7 we conclude $\alpha \leq \rho(A)$. If $\alpha x < Ax$, then there exists some $\alpha' > \alpha$ such that $\alpha' x \leq Ax$ and thus $\rho(A) \geq \alpha' > \alpha$. Again, the upper bounds are verified similarly. ∎

2.2 The Perron-Frobenius Theorem for Positive Matrices

The theory of nonnegative matrices assumes its simplest and most elegant form for positive matrices, and it is for these matrices that Perron made his discoveries in 1907.

As discussed on page 2, the Perron-Froebenius theory shows the existence and the uniqueness of the eigenvector corresponding to the eigenvalue of $\rho(A)$. The first of the following more technical lemmata shows the existence and the second proves the uniqueness.

Lemma (2.2.1) Let $A \in M_n$ be positive. Assume that $Ax = \lambda x$, $x \neq 0$, $|\lambda| = \rho(A)$. Then $A|x| = \rho(A)|x|$ and $|x| > 0$.

Proof (2.2.2)

$$A|x| = |A||x| \geq |Ax| = |\lambda x| = |\lambda||x| = \rho(A)|x| \tag{2.2.3}$$

Thus if we define

$$y \equiv A|x| - \rho(A)|x|$$

then we know that $y \geq 0$. Since $|x| \geq 0$, $x \neq 0$, we know by the positivity of A that $A|x| > 0$. Corollary 2.1.12 shows that in addition $\rho(A) > 0$. It remains for us to show that $y = 0$. Assuming the contrary, we set $z \equiv A|x| > 0$ and thus obtain

$$0 < Ay = Az - \rho(A)z$$

By corollary 2.1.13 this implies $\rho(A) > \rho(A)$, which is an absurdity. Therefore $y = 0$. This in turn yields the eigenvalue equation with eigenvector $|x| = \frac{A|x|}{\rho(A)} > 0$. ∎

Lemma (2.2.4) Let $A \in M_n$ be a positive matrix. Let x, y be nonzero, positive vectors in C^n and assume that $Ax = \lambda x$, $Ay = \lambda y$, $\lambda \neq 0$. Then there is an $\alpha \in \mathsf{C}$ such that $x = \alpha y$.

Proof (2.2.5) Set $\alpha \equiv \min_{1 \leq i \leq n} x_i y_i^{-1}$ and define $r \equiv x - \alpha y$. Notice that $r \geq 0$ and that there is at least one coordinate of r that is equal to zero, so r is by construction *not positive*. But $Ar = Ax - \alpha Ay = \lambda x - \alpha \lambda y = \lambda r$. Assuming $r \neq 0$, we get $r = \lambda^{-1} Ar > 0$ since $A > 0$. But this is a contradiction, and thus $r = 0$. Therefore $x = \alpha y$. ∎

Next, by summarizing the contents of the above lemmata, we deduce the main result for positive matrices.

Theorem (2.2.6) [Perron] Let $A \in M_n$ and suppose that A is positive. Then $\rho(A) > 0$, $\rho(A)$ is an eigenvalue of A, and there is a positive vector x, unique up to a scalar multiple, such that $Ax = \rho(A)x$.

Proof (2.2.7) By lemma 2.2.1, we know that there exists an eigenvalue λ such that $|\lambda| = \rho(A) > 0$, and an associated eigenvector $|x| > 0$. Suppose now that there exists another vector $y \in \mathsf{C}^n$ which is also an eigenvector of A with eigenvalue $\rho(A)$. Then by lemma 2.2.4, there exists an $\alpha \in \mathsf{C}$ such that $x = \alpha y$, and hence the eigenspace corresponding to the eigenvalue $\rho(A)$ is one-dimensional. ∎

A matrix $A \in M_n$ is said to have the *Perron property* if $\rho(A)$ is an eigenvalue. In this case, $\rho(A)$ is often called the *Perron root* of A. The eigenvector as given in the theorem may be called the *Perron vector*. Since the above theorem applies to A^T as well, the Perron vector of A^T is known as the *left Perron vector*.

2.3 The Perron-Frobenius Theorem for Irreducible Matrices

We now wish to extend the above theorem so that it applies to nonnegative matrices as well. For the existence part, we will show that simply by taking limits, the

theorem can be proved. As to the positivity of the Perron vector and thus the uniqueness, we have to introduce some new characterizations of matrices.

Definition (2.3.1) A matrix $A \in M_n$ is said to be *reducible* if either

(a) $n = 1$ and $A = 0$; or

(b) $n \geq 2$, there is a permutation matrix $P \in M_n$, and there is some integer r with $1 \leq r \leq n-1$ such that

$$P^T A P = \tilde{A} = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \tag{2.3.2}$$

where $B \in M_r$, $D \in M_{n-r}$, $C \in M_{r,n-r}$, and $0 \in M_{n-r,r}$ is a zero matrix.

Remark (2.3.3) Here, B, C and D are allowed to have zero entries. We only insist that we should be able to get an $(n-r) \times r$ block of zero entries in the indicated positions by some sequence of row and column interchanges. If $A > 0$, then A is not reducible, and if A is reducible, it must have at least $(n-1)$ zero entries.

Definition (2.3.4) A matrix $A \in M_n$ is said to be *irreducible* if it is not reducible.

The next theorem allows us to extend the Perron-Frobenius theorem to non-negative matrices.

Theorem (2.3.5) A matrix $A \in M_n$ is irreducible iff

$$(\mathbb{1} + |A|)^{n-1} > 0$$

or, equivalently, if $[\mathbb{1} + I(A)]^{n-1} > 0$.

Proof (2.3.6) The proof can be found in the appendix, section A.1. It is rather elaborate, not so much w.r.t. hard mathematics, but w.r.t. the number of new definitions needed. Some of the concepts will become useful in later sections of this thesis. □

Theorem (2.3.7) [Frobenius] Let $A \in M_n$ and suppose that A is irreducible and nonnegative. Then

(a) $\rho(A) > 0$;

(b) $\rho(A)$ is an eigenvalue of A;

(c) There is a unique positive vector x such that $Ax = \rho(A)x$.

Proof (2.3.8) Assertion (a) follows from corollary 2.1.12 since irreducible matrices can not have a zero row or column.

Assertion (b) follows from a limiting argument. For any $\epsilon > 0$, define $A(\epsilon) \equiv [a_{ij} + \epsilon] > 0$. Then by theorem 2.2.6, there is a normalized Perron vector $x(\epsilon)$ of $A(\epsilon)$ and $x(\epsilon) > 0$. Since the set $\{x(\epsilon) : \epsilon > 0\}$ is contained in the compact ball $\{x : x \in \mathsf{C}^n, \sum_{i=1}^n |x_i| \leq 1\}$, there is a monotone decreasing sequence $\epsilon_1, \epsilon_2, \ldots$ with $\lim_{k\to\infty} \epsilon_k = 0$ such that $\lim_{k\to\infty} x(\epsilon_k) \equiv x$ exists. Since $x(\epsilon_k) > 0$ for all $k = 1, 2, \ldots$, it follows that $x = \lim_{k\to\infty} x(\epsilon_k) \geq 0$. $x = 0$ is impossible since all $x(\epsilon_k)$ are normed for all $k = 1, 2, \ldots$.

Since $A(\epsilon_k) \geq A(\epsilon_{k+1}) \geq \ldots \geq A$ by construction, it follows that $\rho(A(\epsilon_k)) \geq \rho(A(\epsilon_{k+1})) \geq \ldots \geq \rho(A)$ for all $k = 1, 2, \ldots$, so the sequence of real numbers $\{\rho(A(\epsilon_k))\}_{k=1,2,\ldots}$ is a monotone decreasing sequence. Thus, $\rho \equiv \lim_{k\to\infty} \rho(A(\epsilon_k))$ exists and $\rho \geq \rho(A)$. On the other hand, we have

$$\begin{aligned} Ax &= \lim_{k\to\infty} A(\epsilon_k)x(\epsilon_k) \\ &= \lim_{k\to\infty} \rho(A(\epsilon_k))x(\epsilon_k) \\ &= \lim_{k\to\infty} \rho(A(\epsilon_k)) \lim_{k\to\infty} x(\epsilon_k) \\ &= \rho x \end{aligned} \tag{2.3.9}$$

and since $x \neq 0$, ρ must be an eigenvalue of A. But then we have $\rho \leq \rho(A)$ which implies $\rho = \rho(A)$.

We next prove the positivity of the eigenvector x. Since $Ax = \rho(A)x$, it follows that $(\mathbb{1} + A)^{n-1}x = (1 + \rho(A))^{n-1}x$. By the irreducibility of A, we know from theorem 2.3.5 that $(\mathbb{1} + A)^{n-1} > 0$ and thus $x = (1 + \rho(A))^{1-n}(\mathbb{1} + A)^{n-1}x$ must be positive.

Finally, assuming that there exists a $y \in \mathsf{C}^n$ such that $Ay = \rho(A)y$, $y \neq 0$, we apply the above argument again to show that $y = (1 + \rho(A))^{1-n}(\mathbb{1} + A)^{n-1}y > 0$, and thus, by lemma 2.2.4, there is an $\alpha \in \mathsf{C}$ such that $y = \alpha y$. ∎

Aside from irreducible matrices, there exists another class of matrices to which the Perron-Frobenius theorem is applicable. Since the existence of the Perron root is assured by the nonnegativity of the matrix alone, we only need a property analogous to theorem 2.3.5 to prove the strict positivity of the Perron vector, i.e.

Definition (2.3.10) A nonnegative matrix $A \in M_n$ is said to be *primitive* if there exists a positive integer k such that $A^k > 0$.

It should be clear how the arguments of proof 2.3.8 apply to primitive matrices. Therefore, substituting 'nonnegative' with 'primitive', theorem 2.3.7 holds here, also. For a direct proof, see [S 73].

Chapter 3

Generalizations of the Perron-Frobenius Theory

Generalizations of the classical Perron-Frobenius theory as discussed in the previous chapter can be made, in principle, using three different methods:

The most obvious one is to extend the existence and uniqueness theorems to infinite-dimensional spaces. For Banach spaces this was done for the first time in a fundamental paper[1] by Krein-Rutman [KR 48] which was published in 1948. More recent operator theories are provided by Krasnoselskii [K 64] and Schäfer [S 74], among others.

The second possible method of extending the Perron-Frobenius theory attempts to generalize the notion of the invariant positive hyperoctant to cones, left invariant under some operator. This approach has been taken mostly in conjunction with the first one, yielding a wealth of characterizations of cones in different spaces [B 73]. Besides Perron-Frobenius theory, applications range from invariance theory and iterative procedures to operator inequalities [S 87].

Finally, the concept of an irreducible matrix may be redefined to cope with the extensions stated above. Several new characterizations of this kind have been developed over the years by numerous authors [K 64, V 68, SV 70].

In this chapter, after having introduced the concept of cones in a Banach space in section 3.1, we go on to develop Perron-Frobenius theories for Banach and Hilbert spaces (section 3.3 and section 3.4, respectively). Section 3.5 is devoted to the extension of irreducibility. Lastly, in section 3.6, we will study the cone of positive semidefinite matrices as an important example of a matrix cone.

[1]The paper is fundamental in the sense that Krein-Rutman were the first to construct a complete Perron-Frobenius operator theory from scratch. Before them, P. Jentsch had already developed a Perron-Frobenius theory for integrals with a positive kernel, P. Jentsch, *Über Integralgleichungen mit positivem Kern*, Reine Angew. Math. **141**, 235, (1912).

3.1 Basic Definitions

Before we start, we extend the notation developed in section 2.1 to infinite-dimensional spaces.

As usual, we call a complete normed linear space a *Banach space* and denote it by the symbol $\mathcal{X}$. For convenience, we will always assume $\mathcal{X}$ to be *real.* This restriction is customary throughout the literature, and will be commented upon again in section 3.4.1. The finite dimensionality of $\mathcal{X}$, if true, is again indicated by a superscript, i.e. $\mathcal{X}^n$. *Hilbert spaces* (possibly over C), denoted $\mathcal{W}$, are Banach spaces equipped with an positive definite *inner product* $(\cdot,\cdot)$, antilinear in the first and linear in the second entry such that $\|x\|^2 = (x,x)$, where $\|\cdot\|$ is the Banach norm. Vectors of either Banach or Hilbert spaces are typically denoted by $x, y, \ldots$. For cones in Banach and Hilbert spaces we use the symbol $\mathcal{K}$; cone elements are typically denoted by $u, v, \ldots$. The *interior* of a cone $\mathcal{K}$ is $\mathcal{K}^\circ$ and its *boundary* is given by $\partial\mathcal{K}$.

3.1.1 Convex Cones

Definition (3.1.1) A set $\mathcal{S} \subset \mathcal{X}$ is *convex* if $\lambda x + (1-\lambda)y \in \mathcal{S}$ for all $x, y \in \mathcal{K}$ and $0 \leq \lambda \leq 1$.

Definition (3.1.2) [Cone] A set $\mathcal{K} \subset \mathcal{X}$ is a *cone* if $\alpha\mathcal{K} \subseteq \mathcal{K}$ for all $\alpha \geq 0$. $\mathcal{K}$ is a *convex cone* if it is a convex set and a cone. A *proper cone* will then denote a closed convex cone which is also *pointed*, i.e. $\mathcal{K} \cap -\mathcal{K} = \{0\}$. A convex cone will be called *solid* if it contains interior points, i.e. $\mathcal{K}^\circ \neq \emptyset$. A convex cone $\mathcal{K}$ is called *reproducing* if every element $x \in \mathcal{X}$ can be written as $x = u - v$ where $u, v \in \mathcal{K}$[2]. $\mathcal{K}$ is a *full* cone if it is pointed and solid.

For convenience we will use the term *cone* instead of proper cone in most of the following sections. If a cone is explicitly neither pointed nor closed we will retreat to the use of the complete specification.

Example (3.1.3)

(a) Let $C_{[a,b]}$ denote the functions which are continuous on a bounded, closed set $[a,b]$. With the sup-norm, $C_{[a,b]}$ is a Banach space. Then $\mathcal{K} = \{u \in C_{[a,b]} : u(t) \geq 0\}$ is a solid cone since the constant $u_0(t) = 1$ is in $\mathcal{K}$ and is an interior point.

(b) $\mathcal{K} = \{u \in L^p(M, d\mu) : u(t) \geq 0\}$ where $d\mu$ denotes some measure on M. $\mathcal{K}$ is reproducing since any $u \in L^p$ can be written as $u(t) = u_+(t) - u_-(t)$. But

[2]Note that in general u and v are not unique.

$\mathcal{K}$ is not solid. To see this, suppose that there is some $u \in \mathcal{K}^\circ$. This means that given some $\varepsilon > 0$, $u + x \in \mathcal{K}$ for all x such that $\|x\| < \varepsilon$. Let E be a measurable set such that $u(t) > 0$ for all $t \in E$ and $\|u\chi_E\| < \varepsilon/2$, where χ_E is the characteristic function of E. Then let $v \equiv -2u\chi_E$. v is an L^p function and $\|v\| = 2\|u\chi_E\| < 2\varepsilon/2 = \varepsilon$. But $u + v \notin \mathcal{K}$, a contradiction.

(c) $\mathcal{K} = \{u \in C_{[a,b]} : u(t) \geq 0, u \text{ nondecreasing}\}$. $\mathcal{K}$ is a cone, but it is not reproducing since only functions of bounded variation can be written as a difference of $\mathcal{K}$ elements [R 88].

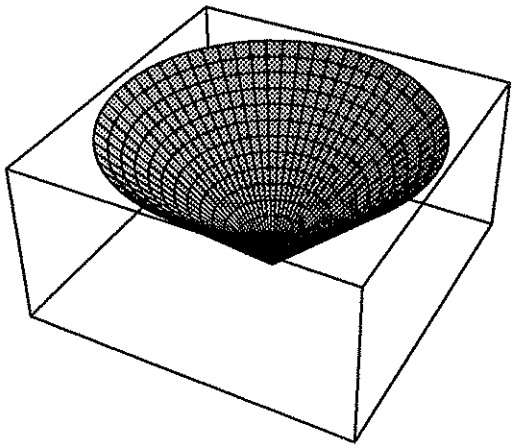

Figure 3.1: A cone in R^3, the so-called "ice-cream-cone".

A general method of constructing cones in Banach spaces was pointed out in [K 64]. In the above examples of cones in some standard spaces of functional analysis, it is evident that reproducing cones are not always solid. The next theorem therefore makes the connection between the two concepts explicit.

Theorem (3.1.4) Every solid cone in $\mathcal{X}$ is reproducing. If the space $\mathcal{X}$ is finite-dimensional, then the converse is also true.

Proof (3.1.5) Assume that $\mathcal{K} \subset \mathcal{X}$ is solid. Let $u \in \mathcal{K}^\circ$ and $x \in \mathcal{X}$ be arbitrary. Then, for sufficiently small $\epsilon > 0$, $w = u + \epsilon x$ is in $\mathcal{K}$; but this means that the element x can be written as $x = \frac{w}{\epsilon} - \frac{u}{\epsilon}$. Thus $\mathcal{K}$ is reproducing.

To show the converse in the finite-dimensional case, assume that $\mathcal{K} \subset \mathcal{X}^n$ is not solid. But then $\mathcal{K} - \mathcal{K}$ is of dimension less than n and thus $\mathcal{K}$ is not reproducing. ∎

3.1.2 The Partial Ordering

As we have stated above, the introduction of cones is useful in many different contexts, particularly in the theory of operator inequalities. The basic tool is the following:

Definition (3.1.6) Let $u, v \in \mathcal{X}$. Let $\mathcal{K} \subset \mathcal{X}$ be a cone. We write $u \succeq v$ if $u - v \in \mathcal{K}$ and $u \succ v$ if $u - v \in \mathcal{K}^\circ$.

In particular $u \succeq 0$ means that $u \in \mathcal{K}$ and $u \succ 0$ stands for $u \in \mathcal{K}^\circ$. If in some cases it is not clear which cone is meant, we write

$$u \overset{\mathcal{K}}{\succeq} 0. \tag{3.1.7}$$

Lemma (3.1.8) $\succeq$ induces a partial ordering in the space $\mathcal{X}$, i.e. $\succeq$ is a reflexive, transitive and antisymmetric relation.

Proof (3.1.9)

Antisymmetry Let $u, v \in \mathcal{K}$. Then

$$u \succeq v \Leftrightarrow u - v \succeq 0$$

and

$$v \succeq u \Leftrightarrow v - u = -(u - v) \succeq 0$$

Since $\mathcal{K}$ is pointed, this implies $u - v \equiv 0$ and thus $u = v$.

Transitivity Let $u, v, w \in \mathcal{K}$ such that $u \succeq v (\Leftrightarrow u - v \in \mathcal{K})$ and $v \succeq w (\Leftrightarrow v - w \in \mathcal{K})$. Then by convexity,

$$u - v + v - w = u - w \succeq 0$$

which is equivalent to $u \succeq w$.

Reflexivity $0 \in \mathcal{K}$ implies that $u \succeq u$ for all $u \in \mathcal{K}$.

∎

An important property of a partial ordering, defined by means of a cone, is that one can pass to the limit in the inequalities, i.e. if $\|x_n - x\| \to 0$, $\|y_n - y\| \to 0$ and $x_n \preceq y_n$ for all n, then $x \preceq y$. Thus such notions as differentiability and fixed points can be defined [K 64].

3.1.3 The Geometry of Cones

Several useful properties of cones can be established by purely geometrical arguments. This will be done here.

Definition (3.1.10) We define a *ray* in a cone $\mathcal{K}$ as the set $\{\lambda x : \lambda \geq 0\}$ for some given $x \in \mathcal{K}$. A ray is an *extreme ray* if $x = ay + bz$, $a, b > 0$ and $y, z \in \mathcal{K}$ imply that both y and z lie on the ray. Elements of an extreme ray will be called *extremal*. A *face* $\mathcal{F}$ of $\mathcal{K}$ is a subset of $\mathcal{K}$, which is a pointed cone such that $x \in \mathcal{F}, y \in \mathcal{K}, x - y \in \mathcal{K}$ implies $y \in \mathcal{F}$. $\mathcal{F} \notin \{\emptyset, \mathcal{K}\}$ is called *proper*. A *subcone* of $\mathcal{K}$ is any cone contained in $\mathcal{K}$, and an *extremal subcone* is a subcone which is generated by some subset of the extremal vectors of $\mathcal{K}$. In this sense, a face $\mathcal{F}$ is nothing but an extremal subcone which is contained in the boundary $\partial\mathcal{K}$ of $\mathcal{K}$. The smallest subspace containing a given face $\mathcal{F}$ will be denoted by $H_{\mathcal{F}}$.

Theorem (3.1.11) Let $\mathcal{K} \subset \mathcal{X}$ be a pointed cone. Then there is no straight line in $\mathcal{X}$ such that it also lies completely in the cone $\mathcal{K}$.

Proof (3.1.12) Let x, y be two arbitrary elements of $\mathcal{X}$. Then the collection

$$z(t) = x + ty, \; t \in]-\infty, \infty[\tag{3.1.13}$$

is a straight line in $\mathcal{X}$. Suppose that $y \neq 0$ and $z(t) \in \mathcal{K}$ for all t. Then for $t \neq 0$, we have

$$\frac{z(t)}{|t|} = \frac{x}{|t|} + \operatorname{sgn}(t)\, y \in \mathcal{K} \text{ for all } |t| < \infty$$

Then taking the limit $|t| \to \infty$, we have that y and $-y$ lie in $\mathcal{K}$ since by definition $\mathcal{K}$ is closed. This is a contradiction to $\mathcal{K}$ being pointed. ∎

It follows from the convexity of the cone that the set of those values t for which the element (3.1.13) belongs to the cone (if such t's exist), either forms a finite closed interval $[t_0, t_1]$, or one of two half-infinite closed intervals: either $[t_0, \infty[$ or $]-\infty, t_0]$. The first case occurs in the case $y \in \mathcal{K}$, the second in the case $-y \in \mathcal{K}$. The next four corollaries can be obtained from these facts.

Corollary (3.1.14) Let $u \in \mathcal{K}$, $x \in \mathcal{X}$. Then $x \preceq t_0 u$ for some t_0 implies that $x \preceq tu$ for all $t > t_0$.

Corollary (3.1.15) Let $u \in \mathcal{K}$. For an element $x \in \mathcal{X}$, suppose that there exists a t such that $x \preceq tu$. Then there is a smallest t_0 for which $x \preceq t_0 u$.

Corollary (3.1.16) Let $u \in \mathcal{K}$. For an element $x \in \mathcal{X}$, suppose that there exists a t such that $x \succeq -tu$. Then there is a smallest t_0 for which $x \succeq -t_0 u$.

The last corollary is an extension of theorem 3.1.11 to unpointed cones.

Corollary (3.1.17) Let $\mathcal{K} \subset \mathcal{X}$ be a pointed cone and let $x, y \in \mathcal{K}$. Assume that the ray $z(t) = x + ty$ lies completely in $\mathcal{K} \cup -\mathcal{K}$ for all $t \in]-\infty, \infty[$. Then x and y are collinear.

Proof (3.1.18) Assume first that $z(t) = x + ty$ lies in $\mathcal{K}$. Then $x + ty \succeq 0$ and thus $x \succeq -ty$. By corollary 3.1.16, there exists some smallest t_+ such that $x + t_+ y \succeq 0$. Now assume that $z(t) = x + ty$ lies in $-\mathcal{K}$ which implies that $x + ty \preceq 0$ and thus $x \preceq -ty$. By corollary 3.1.15–keeping in mind that the corollary has to be changed from dealing with $\mathcal{K}$ to $-\mathcal{K}$– there exist some smallest t_- such that $x + t_- y \preceq 0$. If $z(t)$ is to lie completely in $\mathcal{K} \cup -\mathcal{K}$, then t_+ and t_- should be equal, i.e. $t_0 \equiv t_+ = t_-$. But then the inequalities for t_+ and t_- imply that $0 \preceq x + t_0 y \preceq 0$ and thus $x = -t_0 y$. Therefore x and y are collinear. ∎

Definition (3.1.19) A cone is *generated* by a set of vectors, if any element in the cone can be written as a finite linear combination of these vectors, using only nonnegative coefficients. A *polyhedral cone* is a convex cone generated by a finite set of vectors.

A polyhedral cone is e.g. R^n_+. In figure 3.2 we show another example of a polyhedral cone. A typical cone that is not polyhedral[3] is the "ice-cream-cone" [W 72], i.e.

$$icc = \{(x, y, z) \in \mathsf{R}^3 : x \geq 0, x^2 - y^2 - z^2 \geq 0\} \tag{3.1.20}$$

which is shown in figure 3.1.

Theorem (3.1.21) Let $\mathcal{K} \subset \mathcal{X}^n$ be an arbitrary cone. Then $\mathcal{K}$ is generated by its extremal rays.

Proof by Induction (3.1.22)
$\diamond$: For $n = 1$ the theorem is obvious. For $n = 2$, the cone is either generated by only one vector which is thus the extreme ray, or by two linearly independent vectors which are thus extreme, also. In addition, we note that any interior vector of the cone can be written as the sum of two vectors on the boundary. For suppose $x \in \mathcal{K}^\circ$. Then select a $u \in \mathcal{K}$ linearly independent of x and let $\aleph$ be the plane spanned by u and x. $\tilde{\mathcal{K}} \equiv \aleph \cap \mathcal{K}$ is a cone and $x \in \tilde{\mathcal{K}}$. $\tilde{\mathcal{K}}$ is two-dimensional and thus spanned by its two extreme vectors, say $\tilde{u}, \tilde{v}$, which is turn lie on the boundary of $\mathcal{K}$, i.e. $\tilde{u}, \tilde{v} \in \partial\mathcal{K}$. Thus we only have to show that any $x \in \partial\mathcal{K}$ is generated by extreme rays.
$\hookrightarrow$: So suppose that the theorem has be proved for all $n_0 < n$. Let $x \in \partial\mathcal{K}$ and suppose that x is not extremal. Thus $x = u + v$, and u, v are linear independent of x. If either u or v are in $\mathcal{K}^\circ$, so is x and thus we may assume $u, v \in \partial\mathcal{K}$. Then $x \in \{\alpha u + \beta v : \alpha, \beta \geq 0\} \equiv S$. Denote by $\mathcal{S}$ the largest such cone S such

[3]The statement is true only in dimensions greater than two. In two or fewer dimensions, any cone is polyhedral.

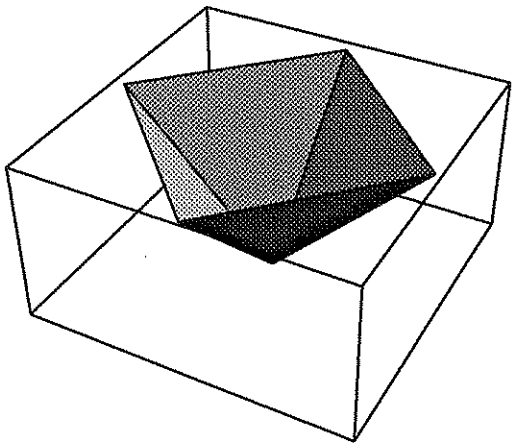

Figure 3.2: A polyhedral cone, generated by 4 vectors in R^3.

that $x \in \mathcal{S} \subset \partial\mathcal{K}$, and by $H_{\mathcal{S}}$ the smallest linear subspace containing $\mathcal{S}$. Since $\dim H_{\mathcal{S}} < n$, $\mathcal{S}$ is generated by extreme rays by the induction.

To complete the proof, we have to show that an extreme ray in $\mathcal{S}$ is an extreme ray in $\mathcal{K}$ as well. Assume the contrary, i.e. y is extremal in $\mathcal{S}$ but not in $\mathcal{K}$. Write $y = u + v$, $u, v \in \mathcal{K}$ and w.l.o.g. we may take $u \notin \mathcal{S}$. Define $\mathcal{S}' \equiv \{w + \alpha u : w \in \mathcal{S}, \alpha \geq 0\}$. Then $w + \alpha u = w + \alpha(y - v) \overset{\mathcal{K}}{\preceq} w + \alpha y \in \mathcal{S} \subset \partial\mathcal{K}$ and thus $\mathcal{S}' \subset \mathcal{S}$. But $\mathcal{S}$ was defined to be the largest such cone. Thus y cannot be extremal in $\mathcal{S}$. ∎

To every element $x \in \partial\mathcal{K}$ there corresponds a particular face whose properties we will examine in the next lemma.

Lemma (3.1.23) Given any $x \in \partial\mathcal{K}$, there exists a face $\mathcal{F}_x$ such that

(a) $x \in \mathcal{F}_x^{\circ}$, relative to the space $H_{\mathcal{F}_x}$,

(b) $\mathcal{F}_x = \partial\mathcal{K} \cap H_{\mathcal{F}_x}$,

(c) $0 \overset{\mathcal{K}}{\preceq} y \overset{\mathcal{K}}{\preceq} x$ implies $y \in \mathcal{F}_x$.

Proof (3.1.24) By theorem 3.1.21, any $x \in \partial\mathcal{K}$ can be written as

$$x = \sum_{i=1}^{m} \alpha_i e_i$$

where $\alpha_i \geq 0$ and e_i is extremal for all $i = 1, \ldots, m$. The cone generated by these x_i satisfies assertion (a). Assertion (b) is also true. Let $\mathcal{F}_x$ be the largest face such that (a) is true. Then $\mathcal{F}_x \subseteq \partial\mathcal{K} \cap H_{\mathcal{F}_x}$. Since $\partial\mathcal{K} \cap H_{\mathcal{F}_x}$ is a cone satisfying (a), we have $\mathcal{F}_x = \partial\mathcal{K} \cap H_{\mathcal{F}_x}$. Finally, $0 \overset{\mathcal{K}}{\preceq} y \overset{\mathcal{K}}{\preceq} x$ implies that $y \in \partial\mathcal{K}$. If $y \notin H_{\mathcal{F}_x}$ then let H' be the subspace spanned by y and $H_{\mathcal{F}_x}$ and $\mathcal{F}' = \partial\mathcal{K} \cap H'$. $\mathcal{F}'$ is a face and $x - y \in \mathcal{F}'$ which in turn implies that x is interior to $\mathcal{F}'$, relative to H'. This is a contradiction to the definition of $\mathcal{F}_x$ and thus $y \in H_{\mathcal{F}_x}$. ∎

3.1.4 Operators Leaving a Cone Invariant

Having introduced the concept of cones, we are now in a position to generalize definition 2.1.4 to operators leaving a cone in a Banach space invariant.

Definition (3.1.25) Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be cones in $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. Denote by $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ the set of operators $A : \mathcal{X}_1 \to \mathcal{X}_2$ for which $A\mathcal{K}_1 \subseteq \mathcal{K}_2$. $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is called the *cone of positivity preserving operators.*

If $\mathcal{K}_1 = \mathcal{K}_2 = \mathcal{K}$, we abbreviate $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ to $\Pi(\mathcal{K})$. As in definition 3.1.6 we introduce an equivalent notation for $A \in \Pi(\mathcal{K})$: We use $A \overset{\mathcal{K}}{\succeq} 0$. If it may not be clear which cone is meant to remain invariant under the action of A, we make this explicit by putting the cone symbol above the relation $\succeq$, i.e.

$$A \overset{\mathcal{K}}{\succeq} 0. \tag{3.1.26}$$

Example (3.1.27) For $\mathcal{K}_1 = \mathsf{R}^n_+$ and $\mathcal{K}_2 = \mathsf{R}^m_+$, $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is the set of all nonnegative $m \times n$ matrices, studied in chapter 2.

Theorem (3.1.28) $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is a proper cone if $\mathcal{K}_1$ is reproducing.

Proof (3.1.29) $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is a cone since if $A, B \succeq 0$, then for all $u \in \mathcal{K}_1$,

$$(A + B)u = \underbrace{Au}_{\succeq 0} + \underbrace{Bu}_{\succeq 0} \succeq 0.$$

Let $\alpha \geq 0$. Then

$$(\alpha A)u = \alpha \underbrace{(Au)}_{\succeq 0} \succeq 0$$

Finally, if both A and $-A$ are in $\Pi(\mathcal{K}_1, \mathcal{K}_2)$, then

$$\left.\begin{array}{l} Au \succeq 0 \\ -Au \succeq 0 \end{array}\right\} \rightsquigarrow Au = 0$$

for all $u \in \mathcal{K}_1$ and thus $A \equiv 0$ since $\mathcal{K}_1$ is reproducing. $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is therefore a (pointed convex) cone. ∎

Note that $\mathcal{K}_1$ had to be reproducing in order to assure that $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is pointed. If we drop this requirement, $\Pi(\mathcal{K}_1, \mathcal{K}_2)$ is still a cone though not pointed. If in addition $\mathcal{K}_1$ and $\mathcal{K}_2$ are full cones, then so is $\Pi(\mathcal{K})$. Furthermore, if $A, B \in \Pi(\mathcal{K})$, then $AB \in \Pi(\mathcal{K})$.

3.2 More about Matrices

Before we proceed to the general theory of operators leaving a cone invariant in either Banach or Hilbert space, we again consider the positive hyperoctant of the classical Perron-Frobenius theory.

Definition (3.2.1) Given a cone $\mathcal{K} \subset \mathsf{R}^n$, we will call $\mathcal{K}$ *fat*, if we can find an orthonormal basis $\{e_i\}_{i=1,\ldots,n}$ such that

(a) $e_i \in \mathcal{K}$ for all $i = 1, \ldots, n$;

(b) for all $u \in \mathcal{K}$ there exist real numbers v_i such that $v = \sum_{i=1}^{n} v_i e_i$ and $v_i \geq 0$.

The basis $\{e_i\}_{i=1,\ldots,n}$ is called the *cone basis*.

Fat cones are full ($x = (1, 1, \ldots, 1) \in \mathcal{K}$) cones and certainly reproducing (by theorem 3.1.4 or, more simply, by the fact that its generators form the basis). In addition, the following lemma is evident.

Lemma (3.2.2) Let $\mathcal{K} \subset \mathsf{R}^n$ be a fat cone. If for some $u \in \mathsf{R}^n$, $(u, v) \geq 0$ for all $v \in \mathcal{K}$, then already $u \in \mathcal{K}$.

Theorem (3.2.3) Let $\mathcal{K} \subset \mathsf{R}^n$ be a fat cone. Then the cone basis $\{e_i\}_{i=1,\ldots,n}$ consists of extreme rays.

Proof (3.2.4) Let e_k be an arbitrary basis vector. Assume that e_k is not an extreme ray, i.e. there exist $x = \{x_i\}, y = \{y_i\} \in \mathcal{K}$ such that $e_k = x + y$ and $x, y \neq 0$ do not both lie on the ray of e_k. But then for all $i = 1, 2, \ldots, n$, we have

$$\begin{aligned} (e_k, e_i) &= (x + y, e_i) \\ &= x_i + y_i \\ &\stackrel{!}{=} \delta_{ik} \end{aligned}$$

and since $x_i, y_i \geq 0$ by the cone property, we have that

$$x = \{0, 0, \ldots, \underbrace{x_k}_{k\text{th position}}, 0, \ldots, 0)$$

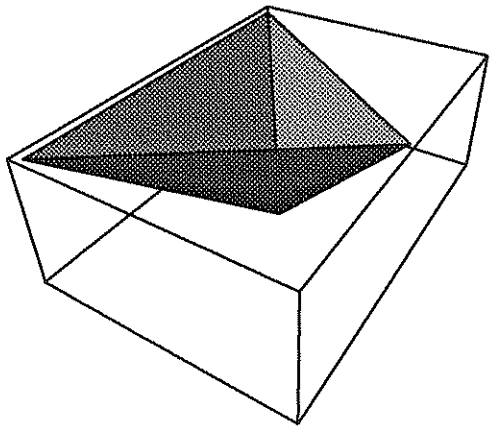

Figure 3.3: A fat cone in R^3.

and the same for y. Thus e_k is an extreme ray. ∎

An important property of fat cones is the mutual orthogonality of their extreme rays. Since this is certainly not true for any *icc*, the next lemma serves to distinguish fat cones from *icc* in dimensions larger than 2.

Lemma (3.2.5) Given some cone $\mathcal{K}$ and some $s \in \mathcal{K}$. Denote by $S^\perp$ the set

$$S^\perp = \{u \in \mathcal{K} : (u, s) = 0 \ \forall \ s \in S\}.$$

Then if $\mathcal{K}$ is a fat cone and x is some element of $\mathcal{K}$, we have that $\{x\}^\perp$ and $\{\{x\}^\perp\}^\perp$ generate $\mathcal{K}$. This is not true if $\mathcal{K} = icc$.

Proof (3.2.6) First, let $\mathcal{K}$ be a fat cone and $\{e_i\}_{i=1,\ldots,n}$ be the cone basis. Let $x = \sum_{i=1}^n x_i e_i$ be in $\mathcal{K}$ and define the index set $\Im(x) = \{i : x_i = 0\}$. Then $\{x\}^\perp = \{y \in \mathcal{K} : y = \sum_{i \in \Im(x)} y_i e_i\}$. Further, $\{\{x\}^\perp\}^\perp = \{z \in \mathcal{K} : z = \sum_{i \in \Im(x)} z_i e_i\}$ and thus $\mathcal{K}$ is generated by $\{x\}^\perp$ and $\{\{x\}^\perp\}^\perp$.

We next show the statement for $\mathcal{K} = icc$. Observe first that for any $a \in icc^\circ$, $\{a\}^\perp = \emptyset$ and thus $\{\{a\}^\perp\}^\perp = icc$. So we may restrict our attention to elements of the boundary. Let $a \in \partial icc$. We now show that there is only one ray in *icc* that is perpendicular to a. Let $x = (a_0, -a_1, -a_2)$. Clearly $x \in \partial icc$ and $(x, a) = 0$. Now suppose that there is some other vector $y \in \partial icc$ such that $y = x + c$ where c is some nonzero difference vector. Then if y is perpendicular to a, so is c. From $y \in \partial icc$ is

follows that $c_0^2 - c_1^2 - c_2^2 = 0$. Thus, after normalization, we obtain $c_0(c_0 + 2a_0) = 0$. The solutions to this equation are either $c_0 = -2a_0$, which implies that $y \in -icc$, or $c_0 = 0$ and thus $c = 0$, which gives uniqueness of x. Therefore $\{a\}^\perp = \{x\}$ and $\{\{a\}^\perp\}^\perp = \{a\}$ and thus icc is not spanned by the vectors of these two sets in more than two dimensions. ∎

Theorem (3.2.7) Let $\mathcal{K} \subset \mathsf{R}^n$ be a fat cone. Then the following holds:

(a) $\Pi(\mathcal{K}) \simeq M_n(\mathsf{R}_+)$.

(b) $A \succeq 0$ implies $A^* \succeq 0$.

(c) Suppose that U is unitary and positivity preserving. Then $U^{-1} \in \Pi(\mathcal{K})$.

(d) U is necessarily a permutation matrix.

Proof (3.2.8)

to (a) Given the cone basis $\{e_i\}_{i=1,\ldots,n}$ and some $A \in \Pi(\mathcal{K})$. Let $[a_{ij}]$ be the representation of A w.r.t. the cone basis and let $u \succeq 0$. Then

$$(Au)_i = \sum_{j=1}^{n} a_{ij} u_j \geq 0.$$

Since this must hold for all $u \succeq 0$, the matrix is nonnegative, i.e. $a_{ij} \geq 0$ for all $i, j = 1, \ldots, n$.

to (b) Assume $A \in \Pi(\mathcal{K})$. Then for any $u, v \in \mathcal{K}$, $0 \leq (u, Av) = (A^*u, v)$. By lemma 3.2.2, $A^* \in \Pi(\mathcal{K})$.

to (c) U is unitary; thus $U^{-1} = U^*$, and this in turn is in $\Pi(\mathcal{K})$ by (b).

to (d) $U_{ij} \geq 0$ for all $i, j = 1, \ldots, n$ by (a). In addition, since U is unitary, we have

$$UU^* = \mathbb{1} \Leftrightarrow \sum_{j=1}^{n} U_{ij} U_{kj} = \delta_{ik}$$

Therefore $U_{ij}U_{kj} = 0 \ \forall \ i \neq k$, which is equivalent to saying that each column has only one nonzero element. By the transpose relation $U^*U = \mathbb{1}$, we get $U_{ji}U_{jk} = 0 \ \forall \ i \neq k$, which is equivalent to saying that each row has only one nonzero element. Again by unitarity, these elements can only be equal to one, and thus U is a permutation matrix.

∎

An important corollary of the above theorem is the following:

Corollary (3.2.9) Let $\mathcal{K} \subset \mathsf{R}^n$ be a fat cone. Then the orthogonal cone basis of $\mathcal{K}$ is essentially unique (i.e. up to a permutation of the basis vectors).

Proof (3.2.10) Let $\{e_i\}$ and $\{f_i\}$ define two cone bases of $\mathcal{K}$. Let U be some unitary transformation such that $Ue_i = f_i$ for all $i = 1, \ldots, n$. U is in $\Pi(\mathcal{K})$ and thus by theorem 3.2.7, U is a permutation of the basis vectors. ∎

More can be said about irreducible matrices before we go on to generalize this concept in a later section. We are especially interested in the question of whether irreducibility of A already implies irreducibility of the exponential $\exp(-tA)$. As it turns out, this is indeed the case.

Lemma (3.2.11) Let $A \in M_n$ be reducible. Then A^m is reducible for all $m \in \mathsf{N}$.

Proof (3.2.12) Let $\tilde{A}$ be given as in definition 2.3.1 and let P the permutation used to reduce A. We then show that the size of the zero matrix in $\tilde{A}$ remains invariant for $\tilde{A}^2$, i.e. let $i \in \mathsf{N}[n-r, n]$ and $k \in \mathsf{N}[1, r]$. Then

$$\begin{aligned}(\tilde{A})_{ik} &= \sum_{j=1}^{r} \underbrace{\tilde{A}_{ij}}_{0} \tilde{A}_{jk} + \sum_{j=n-r}^{n} \tilde{A}_{ij} \underbrace{\tilde{A}_{jk}}_{0} \\ &= 0 \end{aligned} \tag{3.2.13}$$

Thus

$$P^T A^2 P = P^T A P P^T A P = \tilde{A}^2$$

is reducible, which by repetition of the argument implies reducibility for all $m \in \mathsf{N}$. ∎

Corollary (3.2.14) Let $A \in M_n$ be reducible. Then $\exp(-tA)$ and $\exp(-itA)$ are reducible for all $t \in \mathsf{R}$.

Proof (3.2.15) By the series expansion of the exponential and the fact that sums of A^i have the same zero matrix in the lower left corner if A is reducible by the above theorem. ∎

The logical negation of the next theorem then proves the above proposition, i.e. A irreducible implies $\exp(-tA)$ irreducible for some $t > 0$.

Theorem (3.2.16) Let $A \in M_n$. Suppose that $\exp(-tA)$ is reducible for all $t > 0$. Then A is already reducible.

Proof (3.2.17) Let

$$f(A, t) \equiv \exp(-tA)$$

be reducible for all $t > 0$. Then $\mathbb{1} - f(A,t)$ can be reduced by P if P is the permutation matrix that reduces $f(A,t)$. For all $t > 0$, the term

$$\frac{\mathbb{1} - f(A,t)}{-t}$$

can be reduced by P, too and thus

$$\lim_{t \to 0} \frac{\mathbb{1} - f(A,t)}{-t} = A$$

is reducible. ∎

3.3 Positive Operators and Cones in a Banach Space

The particular Perron-Frobenius theory for Banach spaces as presented in this section reviews some facts given in Krein-Rutman [KR 48] and Krasnoselskii [K 64] important to the course of this thesis. We remark that Krein-Rutman uses the term *linear semi-group* for what we call a convex precone; Krasnoselskii uses *conjugate* for dual and both use *cone* for what we call a full cone.

3.3.1 Minihedral Cones

Various types of cones in Banach spaces have been studied in the literature. Among the more important ones are *regular*, *normal* and *minihedral* cones. The first two types are useful primarily for establishing seminorms, bounds and limits of sequences [K 64]; the last one is of more interest to us and will be studied in the following.

Definition (3.3.1) Let $\mathcal{K} \subset \mathcal{X}$. We write $z = \sup(x,y)$ if there is a z such that $z \succeq x$, $z \succeq y$ and $z \preceq \xi$ holds whenever $\xi \succeq x$, $\xi \succeq y$. In an analogous way, we define $\inf(x,y) = -\sup(-x,-y)$.

We remark that the suprenum as defined above is unique for each pair x, y. For suppose that there exists some additional $z' = \sup(x,y)$. This then implies that $z \preceq z'$ and $z' \preceq z$ which gives $z = z'$.

Lemma (3.3.2)

(a) Let $z = \sup(x,y)$ and $u \in \mathcal{X}$. Then $z + u = \sup(x+u, y+u)$. The same holds for the infimum.

(b) Given two elements $x, y \in \mathcal{X}$ such that $\sup(x,y)$ and $\inf(x,y)$ exists, then $\sup(x,y) + \inf(x,y) = x + y$.

Proof (3.3.3) The first part of the lemma follows directly from the definition of the suprenum (infimum). The second part follows from the first one by the following identities

$$\sup(x,y) = x + \sup(0, y-x)$$

$$\inf(x,y) = y + \inf(x-y, 0) = y - \sup(y-x, 0)$$

∎

Definition (3.3.4) A cone $\mathcal{K} \subset \mathcal{X}$ will be called *minihedral* if, for each $x, y \in \mathcal{K}$, $z = \sup(x,y)$ exists.

Remark (3.3.5) Given an $x \in \mathcal{X}$. Define $\mathcal{K}_x = \{x' \in \mathcal{X} : x' = u + x, u \in \mathcal{K}\}$. Thus $\mathcal{K}_x$ is defined from $\mathcal{K}$ by a translation. Minihedrality of $\mathcal{K}$ then means that for all $x, y \in \mathcal{K}$, there exists a $\sup(x,y) \in \mathcal{K}$ such that

$$\mathcal{K}_x \bigcap \mathcal{K}_y = \mathcal{K}_{\sup(x,y)}$$

The cones 1 and 2 presented in example 3.1.3 are minihedral. In addition, the cone of all elements $x \in l^p = \{a : |a|_p \equiv (\sum_{i=1}^{\infty} |a_n|^p)^{\frac{1}{p}} < \infty\}$, with nonnegative coordinates is minihedral, i.e. if $x = \{x_i\}_0^\infty$ and $y = \{y_i\}_0^\infty$, then $\sup(x,y) = \{\max(x_i, y_i)\}_0^\infty$.

Polyhedral cones are not necessarily minihedral, but the converse is true by theorem 3.3.13 below. See figure 3.4 for an example of a polyhedral cone which is not minihedral.

Minihedral cones are important in the study of *Banach lattices* [S 74]. Explicitly, an ordered pair $(\mathcal{X}, \succeq)$ is called a *lattice*, if for all $x, y \in \mathcal{X}$, $\sup(x,y)$ and $\inf(x,y)$ exist. The importance of minihedral cones in the study of Banach lattices is evidenced by the next lemma.

Lemma (3.3.6) Given a minihedral cone $\mathcal{K} \subset \mathcal{X}$. $\mathcal{K}$ is reproducing iff for all $x, y \in \mathcal{X}$ (and not only $\mathcal{K}$) there exists a $z = \sup(x,y)$.

Proof (3.3.7) Assume $\mathcal{K} \subset \mathcal{X}$ minihedral and reproducing. Let $x, y \in \mathcal{X}$ and $u_1, u_2, v_1, v_2 \in \mathcal{K}$ such that $x = u_1 - v_1, y = u_2 - v_2$. Then $x + v_1 + v_2, y + v_1 + v_2 \in \mathcal{K}$, and, consequently, there exists

$$w \equiv \sup(x + v_1 + v_2, y + v_1 + v_2).$$

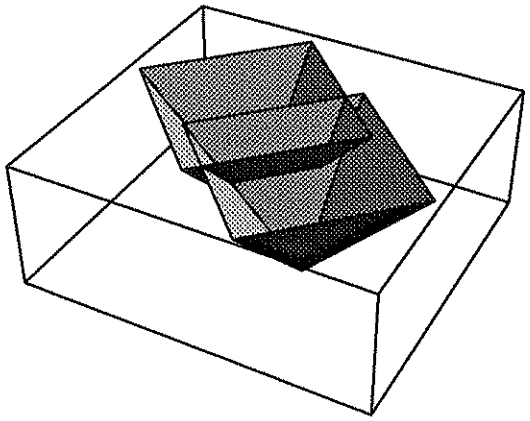

Figure 3.4: The intersection of these two polyhedral cones is not of the same form as the original cone. Therefore this cone is not minihedral.

But then, by the computation rules applied before,

$$w - v_1 - v_2 = \sup(x, y).$$

To show the converse, assume that for each $x, y \in \mathcal{X}$ there exists $z = \sup(x, y)$. Then for each x there exists $x_+ \equiv \sup(x, 0)$. Then define $x_- \equiv x_+ - x = \sup(x, 0) - x = \sup(-x, 0)$. By the definition of $\sup(\cdot, \cdot)$, $x_+, -x_- \in \mathcal{K}$ and x can be written as $x = x_+ - x_-$. ∎

Let $x \in \mathcal{X}$ and assume that x_+ and x_- exist. Then the decomposition $x = x_+ - x_-$ is called the *minimal decomposition* among all decompositions of the form $x = x' - x'', x', x'' \succeq 0$. The decomposition is called minimal in the sense that $x' \succeq x_+$, $x'' \succeq x_-$ for all such possible decompositions.

An important property of Banach spaces is the concept of the *dual space* $\mathcal{X}^*$, i.e. the set of all linear continuous functionals on $\mathcal{X}$. It is therefore natural to extend this concept to cones as well.

Definition (3.3.8) The (not necessarily pointed) cone of positive functionals on the cone $\mathcal{K} \subset \mathcal{X}$ is called the *dual cone* and is denoted by $\mathcal{K}^*$. If $\mathcal{K} \equiv \mathcal{K}^*$, then $\mathcal{K}$ is called *self-dual.*

Let $\mathcal{K}$ be a cone in $\mathcal{X}^n$. Then Krein-Rutman [KR 48] have shown that $\mathcal{K}^*$ exists

if $\mathcal{K}$ is reproducing and $\mathcal{K}$ is pointed iff $\mathcal{K}^*$ is solid. This implies that only pointed cones can be self-dual.

Lemma (3.3.9) Fat cones in R^n are self-dual.

Proof (3.3.10) See lemma 3.2.2. ∎

Theorem (3.3.11) If $\mathcal{K}$ is a solid minihedral cone, then the dual cone $\mathcal{K}^*$ is minihedral.

Proof (3.3.12) The proof can be found in Krein-Rutman [KR 48, page 223]. The principal idea goes back to F. Riesz [R 40]. □

Theorem (3.3.13) Let $\mathcal{X}^n$ be an n-dimensional linear space and $\mathcal{K} \subset \mathcal{X}^n$ be a reproducing cone[4]. Then $\mathcal{K}$ is minihedral iff there exists in $\mathcal{X}^n$ a (not necessarily orthogonal) basis $\{e_i\}_{i=1,\ldots,n}$ such that the set of vectors $x = \sum_{i=1}^{n} a_i e_i, a_i \geq 0$ for all i coincides with $\mathcal{K}$.

Proof (3.3.14) see [KR 48, page 225ff] □

3.3.2 The Krein-Rutman Theorem

The version of the Krein-Rutman theorem that we will state here is called the "finite-dimensional" Krein-Rutman in the literature [B 73]. The full theorem for operators in a Banach space $\mathcal{X}$ may be proved by a direct method as in [KR 48] or by applying the Schauder fixed-point principle as in [K 64]. Another more constructive proof which uses instead the Jordan canonical form, was given by G. Birkhoff [B 67]. We will not prove the theorem, but rather outline how the fixed-point principle may be applied.

Theorem (3.3.15) [(finite-dimensional) Krein-Rutman] Let $\mathcal{K} \in \mathcal{X}^n$ be a full cone. Assume that there exists a linear operator A on $\mathcal{X}^n$ such that $A\mathcal{K} \subseteq \mathcal{K}$. Then $\mathcal{K}$ contains an eigenvector w whose eigenvalue is the spectral radius of A, i.e.

$$Aw = \rho(A)w \tag{3.3.16}$$

Proof (3.3.17) Recall that by the well-known Brouwer fixed-point principle (the finite-dimensional version of the Schauder fixed-point principle [RS1 80]), any linear operator f that leaves invariant a compact and convex subset $\mathcal{S}$ of $\mathcal{X}^n$ and which is in addition continuous, has at least one fixed point $x \in \mathcal{S}$. The fixed point x is a solution of the equation $fx = x$.

[4] In finite dimensions this implies solidity.

Let $\mathcal{S}$ be the intersection of $\mathcal{K}$ and the unit sphere in $\mathcal{X}^n$. If $Ax = 0$ for some $x \in \mathcal{K}$ the theorem is shown for $\lambda = 0$. Thus in the following, assume that $Ax \neq 0$. Then we define the operator $f : \mathcal{S} \to \mathcal{S}$, $x \mapsto Ax/\|Ax\|$ which is continuous. Then by the Brouwer fixed-point theorem, there exists a $u \in \mathcal{S}$ such that $fu = u$ which implies that $Au = \lambda u$ and $\lambda = \|Au\| > 0$. □

3.4 Positive Operators and Cones in a Hilbert Space

A Perron-Frobenius theory applicable to quantum mechanics and quantum field theory in Hilbert spaces was first developed by Glimm-Jaffe [GJ 70]. Extensions and improvements are due to Gross [G 72], Faris [F 72] and Faris-Simon [FS 75]. In this section, we introduce the basic definitions as put forward in Faris [F 72].

3.4.1 Real Hilbert Spaces

In quantum mechanics one deals usually with a complex Hilbert space. However, many problems can be reduced to problems in a real Hilbert space. For suppose that $\mathcal{W}$ is a complex Hilbert space and let $T : \mathcal{W} \to \mathcal{W}$ be a *conjugation*. That is, T is antilinear, $T^2 = 1$ and $(Tu, Tv) = (u, v)^*$. Then we call an element $x \in \mathcal{W}$ such that $Tu = u$ *real*. The set of all real elements forms a real Hilbert space $\mathcal{H}$.

Now if $A : \mathcal{W} \to \mathcal{W}$ is a linear operator such that $AT = TA$, then A leaves the real space invariant. Such an operator is called *real* with respect to T. Many questions concerning the spectrum of $A : \mathcal{W} \to \mathcal{W}$ may be reduced to questions about the restricted operator $A : \mathcal{H} \to \mathcal{H}$ acting in the real Hilbert space. For instance, if $Au = \lambda u$ with λ real and $u \in \mathcal{W}$, then $Au_r = \lambda u_r$ and $Au_i = \lambda u_i$, where $u = u_r + iu_i$ and $u_r, u_i \in \mathcal{H}$. Thus *the multiplicity of the eigenvalue λ may be computed in the real space.*

Example (3.4.1) Let $\mathcal{W} = L^2(M, \mu)$ with inner product $(u, v) = \int_M u^* v d\mu$. Then $Tu = u^*$ defines a conjugation and $\mathcal{H}$ is the space of real functions in L^2. This is the typical example used in nonrelativistic quantum mechanics of particles [F 72]. Here the conjugation can be given a physical meaning. The Hilbert space of a system of n particles is given by $\mathcal{W} = L^2(\mathbb{R}^{3n}, dx)$. The conjugation is given by $Tu(x) = u(x)^*$. Let us now go into the momentum representation by the isomorphism $F : \mathcal{W} \to L^2(\mathbb{R}^{3n}, dk)$ (the Fourier transform). Then the formula $FTF^{-1}g(k) = g(-k)^*$, $g \in L^2(\mathbb{R}^{3n}, dk)$ shows that T reverses the momenta, i.e. T is the *time reversal conjugation*. Thus operators invariant under time reversal are real with respect to T.

3.4.2 Hilbert Cones and Ergodicity

Thus far most of the cones considered have been solid. But already the example 3.1.3(b) shows that there are important cones that do not have a nonempty interior. Therefore, we are interested in a Perron-Frobenius theory that includes cones which are not necessarily solid. This is done by the next definition of a cone, which includes fat cones and *icc*.

Definition (3.4.2) [Hilbert cone] Let $\mathcal{H}$ be a real Hilbert space with scalar product $(\cdot,\cdot)$. A *Hilbert cone* $\mathcal{K} \subset \mathcal{H}$ is a cone such that

(a) $u, v \in \mathcal{K}$ implies $(u, v) \geq 0$,

(b) For all $w \in \mathcal{H}$, there exist $u, v \in \mathcal{K}$ with $w = u - v$ such that $(u, v) = 0$.

Note that this condition implies that a Hilbert cone is reproducing.

Lemma (3.4.3) The above decomposition $w = u - v$ is unique for all $w \in \mathcal{H}$, $w \neq 0$.

Proof (3.4.4) Given a $w \in \mathcal{H}$, we assume that there exist $u, v \succeq 0$ and $f, g \succeq 0$ such that $u \neq f$, $v \neq g$ and $w = u - v = f - g$ with $(u, v) = (f, g) = 0$. Then there exist $c, d \in \mathcal{H}$ such that $u = f + c$ and $v = g + d$. It follows that $u - v = f + c - g - d$ and thus $c = d$. Now we have

$$\underbrace{(u,g)}_{\geq 0} = \underbrace{(f,g)}_{=0} + (c,g) \quad \rightsquigarrow \quad (c,g) = (u,g) \geq 0$$

$$\underbrace{(v,f)}_{\geq 0} = \underbrace{(g,f)}_{=0} + (c,f) \quad \rightsquigarrow \quad (c,f) = (v,f) \geq 0$$

and consequently

$$\begin{aligned} (u,v) &= 0 \\ &= (f+c, g+c) \\ &= \underbrace{(f,g)}_{=0} + \underbrace{(f,c)}_{\geq 0} + \underbrace{(c,g)}_{\geq 0} + \underbrace{(c,c)}_{\geq 0} \end{aligned}$$

Thus $c \equiv 0$ and therefore $u = f$, $v = g$. ∎

Lemma (3.4.5) Let $\mathcal{K}$ be a Hilbert cone. If there exists a $w \in \mathcal{H}$ such that $(w, f) \geq 0$ for all $f \in \mathcal{K}$, then already $w \in \mathcal{K}$.

Proof (3.4.6) Let $w \in \mathcal{H}$ and $u, v \in \mathcal{K}$ such that $w = u - v$, $(u, v) = 0$. Assume $(w, f) \geq 0$ for all $f \in \mathcal{K}$. Select $f = v$. Then

$$\begin{aligned} 0 \leq (w,v) &= (u,v) - (v,v) \\ &= 0 - \|v\|^2 \\ &\leq 0 \end{aligned}$$

This implies $v \equiv 0$ and thus $w \equiv u \in \mathcal{K}$. We remark that this lemma together with the definition of a Hilbert cone assures that Hilbert cones are self-dual. ∎

The introduction of a Hilbert space, and consequently the existence of an inner product, enables us to introduce some new characterizations of elements of $\Pi(\mathcal{K})$.

Definition (3.4.7) Let $\mathcal{H}$ be a real Hilbert space and $\mathcal{K} \in \mathcal{H}$ be a Hilbert cone. A vector $w \succeq 0$ is *strictly positive*, if whenever $u \succeq 0$, $u \neq 0$, then $(u, w) > 0$.

Definition (3.4.8) In a straightforward translation of definition 3.1.25, we call a linear operator $A : \mathcal{H} \to \mathcal{H}$ *positivity preserving* if $u \succeq 0$ implies $Au \succeq 0$. In addition, we define the following two new concepts: A is *positivity improving*, if for all $u \succeq 0$, $u \neq 0$, Au is strictly positive. A is *ergodic* if A is positivity preserving and in addition if for all $u, v \succeq 0$, $u, v \neq 0$, there exists an integer $n \geq 0$ such that $(u, A^n v) > 0$.

As before, we will usually write $A \succeq 0$ or $A \succ 0$, when A is a positivity preserving or positivity improving operator, respectively. Note that if A is positivity preserving and $u, u', v, v' \in \mathcal{K}$, $u \succeq u', v \succeq v'$ implies $(u, Av) \geq (u', Av')$. In appendix A.2 we show that the above definition of an ergodic operator is indeed compatible with the more common one in the von Neumann or Birkhoff-Khinchin theorems [CFS 82, RS4 78].

Remark (3.4.9) When proving uniqueness of the Perron vector in chapter 2, we needed either irreducibility or primitivity of the operator A in question. The connection between these two concepts and ergodicity is given as follows:

Let $\mathcal{K} \in \mathcal{H}$ be a fat cone and let $\{e_i\}$ be the cone basis. Then a sufficient condition for A to be ergodic is that for all e_i, e_j in the cone basis there exists some $n \in \mathsf{N}$ such that

$$(e_i, A^n e_j) > 0 \tag{3.4.10}$$

In section A.1, we show that equation (3.4.10) is equivalent to $A \in M_n(\mathsf{R}^+)$ being *connected* which in turn is equivalent to irreducibility. Primitivity of A merely means that the exponent n may be chosen to be independent of e_i, e_j. *Therefore, for a fat cone, ergodicity is equivalent to irreducibility.*

3.4.3 The Faris Theorem

The version of the Perron-Frobenius theorem that we will need most during the application to spin-lattice models is due to Faris [F 72].

But before continuing, we must redefine the notion of the spectrum $\sigma(A)$ since A is now an operator on a possibly infinite-dimensional Hilbert space.

Definition (3.4.11) Let A be a bounded operator on $\mathcal{W}$. Then $\lambda \in \mathsf{C}$ is said to be in the *resolvent set* $\varrho(A)$ of A if $\lambda \mathbb{1} - A$ is a bijection with a bounded inverse.

If otherwise $\lambda \notin \varrho(A)$, then λ is said to be in the *spectrum* $\sigma(A)$ of A. In addition, the *spectral radius* of A is defined as $\rho(A) \equiv \sup_{\lambda \in \sigma(A)} |\lambda|$

Thus by this definition, $\rho(A)$ need not be an eigenvalue of A as in the finite-dimensional case of definition 2.1.3. If A is a self-adjoint operator, then by [RS1 80, theorem VI.6, VI.8], $\sigma(A) \subset \mathsf{R}$ and $\rho(A) = \|A\|$, where $\|A\|$ is the usual *operator norm* $\|A\| = \sup_{\|x\|=1} \|Ax\|$. Since we will always consider self-adjoint operators, we will therefore use $\rho(A)$ instead of $\|A\|$ is the next theorem and the following chapters.

Theorem (3.4.12) [Faris] Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded, positive and self-adjoint operator on the real Hilbert space $\mathcal{H}$. Assume that $\rho(A)$ is an eigenvalue of A. Let $\mathcal{K} \subset \mathcal{H}$ be a Hilbert cone. Assume that A is positivity preserving and ergodic. Then $\rho(A)$ is a simple eigenvalue and the eigenspace is spanned by a strictly positive vector.

Proof (3.4.13) We first show the strict positivity of the eigenvector. For convenience, let $\lambda = \rho(A)$. We will assume $\lambda \neq 0$ throughout the proof.

Consider $w \in \mathcal{H}$, $w \neq 0$ with $Aw = \lambda w$. By the Hilbert cone property 3.4.2 we can write $w = u - v$ with $u, v \succeq 0$ and $(u, v) = 0$. Set $z = u + v$. Then $z \succeq 0$ by the cone property and

$$\begin{aligned}(z,z) &= (u,u) + \underbrace{(u,v)}_{0} + \underbrace{(v,u)}_{0} + (v,v) \\ &= (u,u) - \overbrace{(u,v)} - \overbrace{(v,u)} + (v,v) \\ &= (w,w).\end{aligned}$$

By hypothesis A^n is positivity preserving so that $(w, A^n w) \leq (z, A^n z)$. Therefore

$$\begin{aligned}\lambda^n(w,w) &= (w, A^n w) \\ &\leq (z, A^n z) \qquad (3.4.14) \\ &= \rho(A^n)(z,z) \\ &= \rho(A)^n(z,z) \\ &= \lambda^n(z,z) \\ &= \lambda^n(w,w) \qquad (3.4.15)\end{aligned}$$

This leads to the identity

$$(w, A^n w) = (z, A^n z), \qquad (3.4.16)$$

which in its explicit form is given by

$$\begin{aligned}&(u, A^n u) - (u, A^n v) - (v, A^n u) + (v, A^n v) \\ &= (u, A^n u) + (u, A^n v) + (v, A^n u) + (v, A^n v)\end{aligned}$$

We therefore get

$$(u, A^n v) + (v, A^n u) = 0. \tag{3.4.17}$$

Since A is ergodic and $u, v \succeq 0$ by assumption, it follows immediately that either $u = 0$ or $v = 0$. Thus we have shown that $w \succeq 0$ or $-w \succeq 0$, and we may choose $w \succeq 0$.

We next show that λ can have multiplicity of at most one. We will give two reasons. The first is based on the pointedness of Hilbert cones. The second–the one used by Faris–again uses the ergodicity of A.

The geometrical argument goes as follows. Let $w, x \neq 0$ with $Aw = \lambda w$ and $Ax = \lambda x$. Then $w, x \in \mathcal{K} \cup -\mathcal{K}$ by the above; by the linearity of the eigenvalue equation, $w + tx$ is an eigenstate for all $t \in \mathsf{R}$ corresponding to the eigenvalue λ, also. But by corollary 3.1.17, this implies that w, x are collinear if the ray is to lie completely in $\mathcal{K} \cup -\mathcal{K}$.

Faris' argument again uses the ergodicity of A. For all $u \succeq 0$, $u \neq 0$ there exists an $m \in \mathsf{N}$ such that

$$(u, A^m w) = \lambda^m (u, w) > 0 \tag{3.4.18}$$

and therefore w is strictly positive (note that λ is positive by construction). Assume that the eigenspace is two-dimensional, i.e. spanned by w, x and $(x, w) = 0$. But then the above arguments hold and thus w, x are strictly positive and $(w, x) > 0$, which is a contradiction. ∎

Remark (3.4.19) Under the hypothesis of the theorem, uniqueness of the eigenvalue and strict positivity of the eigenvector are sufficient to ensure ergodicity for A. Assume that w is the unique and strictly positive ground state of A. Let $u, v \in \mathcal{K}$ be nonzero vectors such that $(u, A^n v) = 0$ for all $n \geq 0$. Then for all $t \in \mathsf{R}$, $(u, \exp t(A - \rho(A))v) = 0$. By the spectral theorem of bounded operators [RS1 80], the exponential reduces to the projector onto the ground state w, i.e. $\lim_{t \to +\infty} (u, \exp t(A - \rho(A))v) = (u, w)(w, v) = 0$ which is a contradiction since w is strictly positive. Thus there must be some n such that $(u, A^n v) > 0$, and this is ergodicity.

3.5 Generalizing Irreducibility

Before generalizing the notion of irreducibility, we give an alternative definition which emphasizes the geometric nature of this concept [G 59]. If $e_1, e_2, \ldots, e_n$ are the unit coordinate vectors in $\mathcal{X}^n$, then a *coordinate subspace* is a subspace spanned by any subset of $\{e_1, \ldots, e_n\}$. An *irreducible matrix* is a matrix which has no invariant coordinate subspace of dimension less then n. Since the positive

hyperoctant is generated by the vectors $e_1, \ldots, e_n$, a nonnegative irreducible matrix maps the positive hyperoctant onto itself and leaves no face invariant.

After replacing the hyperoctant by an arbitrary solid cone $\mathcal{K}$, the above definition leads to the following generalization.

Definition (3.5.1) A matrix $A \in \Pi(\mathcal{K})$ is *$\mathcal{K}$-irreducible* if A leaves no face of $\mathcal{K}$ invariant. Otherwise A is called *$\mathcal{K}$-reducible.*

Theorem (3.5.2) $A \in \Pi(\mathcal{K})$. Then A is $\mathcal{K}$-irreducible iff no eigenvector of A lies on the boundary of $\mathcal{K}$.

Proof (3.5.3) Suppose that $A \succeq 0$ is $\mathcal{K}$-reducible. Then there exists at least one face $\mathcal{F}$ of $\mathcal{K}$ which is left invariant by A. So A, restricted to $H_{\mathcal{F}}$, leaves the solid cone $\mathcal{F}$ invariant. By the Krein-Rutman theorem 3.3.15 there exists a vector $x \in \mathcal{F}$ such that $A_{H_{\mathcal{F}}} x = \lambda x$. But x is also an eigenvector of A, operating on the whole space, and x is in $\partial\mathcal{K}$. Conversely, suppose that x is an eigenvector of A in $\partial\mathcal{K}$ and let $\mathcal{F}_x$ be the face associated with x as defined in lemma 3.1.23. Then, given any $y \in \mathcal{F}_x$, there exists an $\alpha > 0$ such that $\alpha x \overset{\mathcal{F}_x}{\succeq} y$ and hence $\alpha x \overset{\mathcal{K}}{\succeq} y$. Since $A \in \Pi(\mathcal{K})$ by hypothesis, this implies

$$Ay \overset{\mathcal{K}}{\preceq} A(\alpha x) \overset{\mathcal{K}}{\preceq} \lambda \alpha x$$

which lies in $\mathcal{F}_x$. Thus by lemma 3.1.23(c), $Ay \in \mathcal{F}_x$. Therefore $\mathcal{F}_x$ is invariant under A and thus A is $\mathcal{K}$-reducible. ∎

Theorem (3.5.4) If $A \in \Pi(\mathcal{K})$ has two eigenvectors in $\mathcal{K}^\circ$, then A also has an eigenvector on the boundary of $\mathcal{K}$. Furthermore. the corresponding eigenvalues are all equal.

Proof (3.5.5) Let $x_1, x_2 \in \mathcal{K}^\circ$ be two linear independent eigenvectors of A with corresponding eigenvalues λ_1 and λ_2. Let t_0 be the smallest possible t as given by corollary 3.1.15 such that $x_1 \preceq t_0 x_2$ holds (where we assume that $\lambda_1 \geq \lambda_2 \geq 0$). Then $x_3 \equiv t_0 x_2 - x_1$ lies in $\partial\mathcal{K}$. If $\lambda_1 \neq 0$, then

$$\frac{Ax_3}{\lambda_1} = t_0 \underbrace{\frac{\lambda_2}{\lambda_1}}_{\leq 1} x_2 - x_1 \in \mathcal{K}$$

Therefore $\lambda_1 = \lambda_2$. If $\lambda_1 = 0$, then $\lambda_2 = 0$ and thus $Ax_3 = 0$. Therefore x_3 is an eigenvector of A on the boundary $\partial\mathcal{K}$ for both cases. ∎

The above lemma and theorem can be combined to yield the next theorem.

Theorem (3.5.6) $A \in \Pi(\mathcal{K})$ is $\mathcal{K}$-irreducible iff A has exactly one eigenvector in $\mathcal{K}$, and this eigenvector is in $\mathcal{K}^\circ$.

The Perron-Frobenius theorem for $\mathcal{K}$-irreducible operators then reads:

Theorem (3.5.7) [Vandergraft] If $A \in \Pi(\mathcal{K})$ is $\mathcal{K}$-irreducible, then

(a) $\rho(A)$ is a simple eigenvalue,

(b) there is an eigenvector corresponding to $\rho(A)$ in $\mathcal{K}^\circ$, and no other eigenvector lies in $\mathcal{K}$.

Proof (3.5.8) Assertion (a) is the existence statement and as such is implied by the Krein-Rutman theorem 3.3.15. The uniqueness follows as in the Faris theorem 3.4.12, again using the fact that a degenerate eigenvalue would imply an at least two-dimensional eigenspace and by the above theorems on $\mathcal{K}$-irreducibility this is impossible. ∎

Another generalization of irreducibility is given by

Definition (3.5.9) Let $A \in \Pi(\mathcal{K})$. Then A is called *$\mathcal{K}$-positive* if $A(\mathcal{K}-\{0\}) \subseteq \mathcal{K}^\circ$.

We note that this generalization strongly mimicks the definition 2.3.10 of a primitive matrix. Again a Perron-Frobenius type theorem can be given.

Theorem (3.5.10) If $A \in \Pi(\mathcal{K})$ is $\mathcal{K}$-positive, then the assertions of theorem 3.5.7 hold.

There exist several other Perron-Frobenius theories. A well-known Perron-Frobenius theory builds on the notion of a u_0-positive operator and is developed in Krasnoselskii [K 64] and Schäfer [S 74].

Definition (3.5.11) An operator $A \overset{\mathcal{K}}{\succeq} 0$ is called *u_0-positive* if for some nonzero $u_0 \in \mathcal{K}$ and any nonzero $x \in \mathcal{K}$ there exist real numbers $\alpha(x), \beta(x) > 0$ and an integer $k(x) > 0$ such that

$$\alpha(x)u_0 \preceq A^k x \preceq \beta(x)u_0 \tag{3.5.12}$$

Another Perron-Frobenius theory is developed in Schneider-Vidyasagar [SV 70] for classes of operators of the following type:

Definition (3.5.13) An operator $A \overset{\mathcal{K}}{\succeq} 0$ is called *quasi-interior* if for some $\lambda > \rho(A)$,

$$A(\lambda \mathbb{1} - A)^{-1} \succ 0 \tag{3.5.14}$$

The following connections exist between the various concepts of irreducibility.

(a) $\mathcal{K}$-irreducibility $\Leftrightarrow$ quasi-interior

(b) $\mathcal{K}$-positivity $\rightsquigarrow$ $\mathcal{K}$-irreducibility

(c) u_0-positivity $\rightsquigarrow$ $\mathcal{K}$-irreducibility

Further approaches to a Perron-Frobenius theory for $\Pi(\mathcal{K})$ can be found in [B 73] and the references therein.

3.6 The Cone of Positive Semidefinite Matrices

3.6.1 Characterizations

An important example of a natural cone in M_n is given by the following class of matrices.

Definition (3.6.1) An $n \times n$ matrix Hermitian matrix A is said to be *positive definite* if

$$x^*Ax > 0 \tag{3.6.2}$$

for all nonzero $x \in \mathsf{C}^n$. If the strict inequality is weakened to $x^*Ax \geq 0$, then A is called *positive semidefinite.* We remark that the assumption of Hermiticity is not a necessity but customary.

Positive semidefinite matrices play an important role in many fields of mathematics [T 66]. The next two theorems recall some well-known facts of positive semidefinite matrices which we will later use extensively. Their proofs can be found e.g. in Horn-Johnson [HJ 85].

Theorem (3.6.3) A Hermitian matrix $A \in M_n$ is positive (semi)definite iff all of its eigenvalues are positive (nonnegative). If follows that the trace and the determinant of A are positive (nonnegative). Also, at least one diagonal element of a positive semidefinite matrix is nonzero.

Theorem (3.6.4) A matrix $A \in M_n$ is positive definite iff there is a nonsingular matrix $B \in M_n$ such that $A = B^*B$. If A is positive semidefinite, then there exists for all integers $k \geq 0$ a unique positive semidefinite matrix C such that $C^k = A$.

Since any nonnegative (nontrivial) linear combination of positive (semi)definite matrices of the same size is again positive (semi)definite, the set of all positive (semi)definite matrices in M_n forms a cone, denoted **PSD**, the *cone of positive (semi)definite $n \times n$ matrices.* **PSD** is a proper cone and since the unit matrix $\mathbb{1}$ is positive definite, **PSD** is solid.

Definition (3.6.5) Let $\mathcal{V}$ be the *real space of $n \times n$ Hermitian matrices*, isomorphic to R^{n^2}. Let the *inner product* be defined by

$$\langle A, B \rangle \equiv \mathrm{trace}(A^*B) \tag{3.6.6}$$

If $A = [a_{ij}], B = [b_{ij}] \in M_n$ then the *Hadamard product* of A and B is the matrix

$$A \circ B \equiv [a_{ij} b_{ij}] \tag{3.6.7}$$

Theorem (3.6.8) [Schur product theorem] If $A, B \in M_n$ are positive (semi)definite, then $A \circ B$ is also positive (semi)definite.

Proof (3.6.9) See Horn-Johnson [HJ 85, Theorem 7.5.3] □

Theorem (3.6.10) Given the Hilbert space $\mathcal{V}$ and the cone **PSD**, **PSD** is self-dual, i.e. **PSD** = **PSD***.

Proof (3.6.11) [BB 73]
PSD $\subseteq$ **PSD***: We want to show that given $A, B \in$ **PSD**, then $\langle A, B \rangle \geq 0$. Thus let $A, B \in$ **PSD**. $A \circ B \in$ **PSD** by theorem 3.6.8 and thus

$$((A \circ B)x, x) \geq 0 \; \forall x \in \mathsf{C}^n$$

Let $e = (1, \ldots, 1)$. Then

$$\langle A, B \rangle = \sum_{ij}^{n} a_{ij} b_{ij} = ((A \circ B)e, e) \geq 0$$

PSD* $\subseteq$ **PSD**: Let $A \in$ **PSD***. Then

$$(Ax, x) = \langle A, xx^* \rangle \geq 0$$

since $xx^* \in$ **PSD** and thus A is positive semidefinite. ∎

Theorem (3.6.12) If $A \in M_n$ is a positive semidefinite matrix of rank k, then A may be written in the form

$$A = v_1 v_1^* + v_2 v_2^* + \ldots + v_k v_k^*, \tag{3.6.13}$$

where each $v_i \in \mathsf{C}^n$ and the set $\{v_i\}_{i=1,\ldots,k}$ is an orthogonal set of nonzero vectors.

Proof (3.6.14) By the spectral theorem, we write $A = U \Lambda U^*$, where U is unitary and $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$. Then let v_i be $\lambda_i^{\frac{1}{2}}$ times the ith column of U. ∎

Lemma (3.6.15) A ray in **PSD** is an extreme ray, iff A has rank 1.

Proof (3.6.16) Let $0 \neq x \in \mathsf{C}^n$. Then xx^* is in **PSD** and of rank 1. We now want to show that xx^* is an extreme ray. So suppose that there exist $A, B \in$ **PSD** and a real number $\alpha \in]0, 1[$ such that

$$xx^* = \alpha A + (1 - \alpha) B \tag{3.6.17}$$

Now let us choose an orthonormal set $\{v_i\}_{i,1,\ldots,n}$ such that $x \equiv v_1$. Thus $x^* v_k = (x, v_k) = 0$ for all $k \geq 2$ and

$$v_k^* x x^* v_k = |x^* v_k|^2 = 0 = \alpha \underbrace{v_k^* A v_k}_{\geq 0} + (1-\alpha) \underbrace{v_k^* B v_k^*}_{\geq 0}$$

Thus $v_k^* A v_k = v_k^* B v_k = 0$ for all $k \geq 2$. By the positive semidefiniteness of A and B this implies that $Av_k = Bv_k = 0$ for all $k \geq 0$ and thus A, B have rank 1. In addition, by theorem 3.6.12 we know that $xx^* = aA = bB$. Thus xx^* is an extreme ray.

To show the converse, assume that rank$A \geq 2$. Again by theorem 3.6.12, we can write

$$A = v_1 v_1^* + v_2 v_2^* + \ldots$$

where $\{v_i\}_{i=1,2,\ldots}$ is an orthogonal set of nonzero vectors in C^n. Then if A is an extreme ray, this would imply that

$$\mathbf{PSD} \ni C = v_2 v_2^* + \ldots = \lambda v_1 v_1^*$$

$\lambda \neq 0$. Therefore

$$Cv_1 = \lambda (v_1 v_1^*) v_1 = \lambda |v_1|^2 v_1 \neq 0$$

which is a contradiction since $Cv_1 = 0$ by the orthogonality of the set $\{v_i\}$. Thus A cannot be an extreme ray. ∎

The assertions of theorem 3.6.12 and lemma 3.6.15 can be combined to yield the following theorem.

Theorem (3.6.18) The cone of positive semidefinite matrices is generated by the positive semidefinite matrices of rank 1. The boundary of **PSD** consists of the positive semidefinite matrices, whereas interior of **PSD**, i.e. **PSD**$^\circ$ consists of the positive definite matrices.

Theorem (3.6.19) **PSD** $\subset \mathcal{V}$ is a Hilbert cone.

Proof (3.6.20) Since **PSD** is self-dual by theorem 3.6.10, and reproducing by the solidicity, it remains to show that **PSD** is reproducing such that the decomposition is indeed orthogonal. Let W be an arbitrary Hermitian operator in $\mathcal{V}$. Then the *modulus* of W is defined as the positive square root of W^2, i.e.

$$|W| \equiv \sqrt{(W^2)} \tag{3.6.21}$$

With the help of the modulus, we subsequently define the orthogonal decomposition,

$$W_\pm \equiv \frac{(|W| \pm W)}{2} \tag{3.6.22}$$

By this definition, $W = W_+ - W_-$ and $|W| = W_+ + W_-$. We first show that the decomposition is indeed orthogonal, that is $\langle W_+, W_- \rangle = 0$.

$$\begin{aligned} W_+W_- &= \frac{1}{4}((|W| + W)(|W| - W)) \\ &= \frac{1}{4}\left(|W|^2 - W^2 - |W|W + W|W|\right) \\ &= \frac{1}{4}(-|W|W + W|W|) \end{aligned}$$

which is indeed zero, since W commutes with $|W|$ [HJ 85, theorem 7.2.6]. Finally, it remains to show that W_+, W_- are positive semidefinite. Define W_n by

$$W_n = n\frac{W_+^2}{(\mathbb{1} + nW_+^2)} \tag{3.6.23}$$

Then by the orthogonal decomposition, we have $W_n|W| = W_nW_+$. In addition,

$$\begin{aligned} \|W_n|W| - W_+\|^2 &= \|n\frac{W_+^3}{(\mathbb{1} + nW_+^2)} - W_+\|^2 \\ &= \|\frac{W_+}{(\mathbb{1} + nW_+^2)}\|^2 \\ &= \|\frac{W_+^2}{(\mathbb{1} + nW_+^2)^2}\| \\ &\leq \|\frac{W_+^2}{(\mathbb{1} + nW_+^2)}\| \, \|\frac{1}{(\mathbb{1} + nW_+^2)}\| \\ &= \frac{1}{n}\|\mathbb{1} - \frac{1}{(\mathbb{1} + nW_+^2)}\| \, \|\frac{1}{(\mathbb{1} + nW_+^2)}\| \end{aligned}$$

Now since $\sigma((\mathbb{1} + nW_+^2)) = [1, \infty[$, and hence $\sigma((\mathbb{1} + nW_+^2)^{-1}) = [0, 1]$, we have that

$$\|W_n|W| - W_+\| \leq \sqrt{\frac{1}{n}}$$

Thus $W_n|W|$ converges uniformly to W_+. But $W_n|W|$ can be written as

$$W_n|W| = \left(|W|^{\frac{1}{4}}|W_+|^{\frac{1}{2}}\left(\frac{\mathbb{1}}{n} + W_+^2\right)^{-\frac{1}{2}}|W_+|^{\frac{1}{2}}|W|^{\frac{1}{4}}\right)^{\frac{1}{2}}$$

which is an element of **PSD** by theorem 3.6.4. Since **PSD** is closed, $W_+ \in$ **PSD**. ∎

Theorem (3.6.24) **PSD** is in general not a fat cone.

Proof (3.6.25) Let $\mathcal{K}$ be a fat cone in a space $\mathcal{H}^n$. Then this implies that $\mathcal{K}$ is generated by n orthogonal vectors. Here we will show that the cone of positive semidefinite matrices is generated by infinitely many vectors, thus it can not be a fat cone. We recall the well-known analogy of the lightcone in four-dimensional Minkowski space $\mathcal{M}^4$ with the cone of positive semidefinite matrices in M_2. Let A be a positive semidefinite matrix and let $x = (x_0, x_i)$ be an element of $\mathcal{M}^4$. Then

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix} \tag{3.6.26}$$

where x lies on the forward lightcone, that is $x_0 \geq 0$ and $(x,x)_{\mathcal{M}} = x_0^2 - \sum_{i=1}^3 x_i^2$. As in lemma 3.2.5, the forward lightcone is a four-dimensional *icc* and thus not a fat cone. Now, adding an arbitrary number of dimensions such that we deal with the positive semidefinite matrices in M_n, we have to *add* generating vectors. Thus the number of generating vectors increases and will never be finite. ∎

3.6.2 Some Operators Leaving PSD Invariant

This section is devoted to the study of operators[5] on $\mathcal{V}$ that leave **PSD** invariant. By the isomorphism of $\mathcal{V}$ to R^{n^2}, elements of $\mathcal{V}$ will then be replaced by n^2-dimensional column vectors for which a suitable ordering has to be applied. Operators **A** on $\mathcal{V}$ will be elements of $M_{n^2,n^2}(\mathsf{R})$.

Results on the structure of Π(**PSD**) are rare. A. Berman states in [B 73], that "the general question of characterizing Π(**PSD**) seems to be a very difficult one". Most results that can be found in the literature are due to O. Taussky [T 66, T 67, T 72]. Unfortunately, we were not able to uncover more recent results. Indeed, it seems that little work has been devoted to that question since 1973 [S 91].

The next lemma is concerned with products of two or three positive semidefinite matrices.

Lemma (3.6.27) If $A = S_1 S_2$ and S_i is Hermitian and positive semidefinite, then A has real and positive roots and a full set of eigenvectors. If $A = S_1 S_2 S_3$ and $S_i > 0$ is Hermitian as well as A, then A is also positive semidefinite.

Proof (3.6.28) We note that the second assertion is due to E. Wigner[6]. Various other proofs are reviewed in the paper by Taussky [T 66]. □

Taussky further reviewed a lemma on so-called "generalized automorphs". This lemma is important since it provides the basic ingredient needed to apply the finite-dimensional Krein-Rutman theorem 3.3.15.

[5]For notational clarity, such operators on $\mathcal{V}$ will be typeset as boldface characters, i.e. **T**.

[6]E.P. Wigner, *On weakly positive matrices*, Canadian J. Math. **15**, 313, (1963)

Lemma (3.6.29) Let $T \in M_n(\mathsf{R})$. Then the transformation $\mathbf{T} : A \mapsto TAT^*$ transforms a positive semidefinite matrix A into a positive semidefinite matrix. If in addition, T is nonsingular, then a positive definite A is transformed into a positive definite matrix.

Proof (3.6.30) $(x, Ax) \geq 0$ for all x implies that $(T^*x, AT^*x) = (x, TAT^*x) \geq 0$ for all x and thus the first assertion is proved.

Assume now that A is positive definite and that T is nonsingular, that is, $Tx = 0$ implies $x = 0$. Then $(T^*x, AT^*x) = (x, TAT^*x) = 0$ only if $x = 0$ and thus TAT^* is positive definite. ∎

Remark (3.6.31) $\mathbf{T}$ is certainly not $\mathcal{K}$-irredicible or even $\mathcal{K}$-positive, thus the uniqueness of the Perron vector is not assured by the above generalized Perron-Frobenius methods. Taussky further examined the characteristic roots of $\mathbf{T}$ and established conditions for the strict positivity of some characteristic vectors. If $t_1, \ldots, t_n$ denote the characteristic roots of a *real* matrix T, then $\mathbf{T}$ has characteristic roots $t_i t_k, i, k = 1, \ldots, n$. If x_i is a characteristic vector of t_i then $t_i t_k$ has characteristic vector $x_i x_k^* + x_k x_i^*$. If $X = x_i x_k^* + x_k x_i^*$ is real, then $i \neq k$ implies that X is not positive definite. For $i = k$, X is positive definite iff T is similar to the product of an orthogonal matrix and a scalar matrix. If $x_i \in \mathsf{C}^n$ and $x_k = \bar{x}_i$, then X is not positive definite unless $n = 2$.

In the next theorem, we establish a straightforward generalization of Taussky's result on generalized automorphs. We will later make extensive use of this result.

Theorem (3.6.32) Let $x \in \mathbf{PSD}$. Suppose that there exists an operator $\mathbf{S}$ on $\mathcal{V}$ such that

$$A \mapsto \sum_{ij} S_{ij} T_i A T_j^* \tag{3.6.33}$$

where S is a positive semidefinite matrix in M_m and $\{T_i\}_{i=1,\ldots,m}$ a set of matrices in $M_n(\mathsf{R})$. Then the mapping is positivity preserving w.r.t. **PSD**.

Proof (3.6.34) The proof again utilizes the unitary diagonalization of a positive semidefinite matrix. Let $S = U \Lambda_S U^*$ where $\Lambda_S = diag(\lambda_1, \ldots, \lambda_m)$.

$$\begin{aligned}
\sum_{ij}^{m} S_{ij} T_i A T_j^* &= \sum_{ij}^{m} (U \Lambda_S U^*)_{ij} T_i A T_j^* \\
&= \sum_{ijlk}^{m} U_{il} \Lambda_{lk} U_{kj}^* T_i A T_j^* \\
&= \sum_{ijl}^{m} U_{il} \lambda_l U_{lj}^* T_i A T_j^*
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{ijl}^{m} U_{il}\lambda_l U_{lj}^* T_i A T_j^* \\
&= \sum_{l}^{m} \lambda_l \left(\sum_{i}^{m} U_{il} T_i \right) A \left(\sum_{j}^{m} U_{lj}^* T_j^* \right) \\
&= \sum_{l} \lambda_l K_l A K_l^*
\end{aligned}
$$

which is indeed positive semidefinite since all $\lambda_l \geq 0$ for S positive semidefinite and by lemma 3.6.29. ∎

At the end of this section, we present another interesting property of the generalized automorph $\mathbf{T}$.

Theorem (3.6.35) [Stein] Let $T \in M_n$. Then $T^m \to 0$ iff there exists a matrix $A \in \mathbf{PSD}^\circ$ such that

$$A \succ TAT^* \tag{3.6.36}$$

Proof (3.6.37) The theorem is due to P. Stein. An alternative proof is given by Taussky. See [T 66] for further details. □

Chapter 4

Ground States and Semigroups

To the best of our knowledge, the first time Perron-Frobenius theory was applied to establish uniqueness of ground states for a physical model was in a paper of Elliot Lieb and Daniel Mattis [LM 62a] on a certain one-dimensional spin-lattice model. There, as well as in a subsequent paper of the same authors [LM 62b], essentially the classical Perron-Frobenius theory as presented in chapter 2 and section 3.2 is used, i.e. in the context of fat cones and connected matrices.

In 1970, Glimm-Jaffe used the ideas of section 3.4 to establish existence and uniqueness theorems for some models of quantum field theory [GJ 70]. There, for the first time, semigroups of operators were used instead of the operators themselves. An abstract version of their result was given by Faris [F 72], the fundamentals of which are partially recounted in section 3.4. As stated on page 28 a number of authors improved this treatment and a summary of results can be found in [RS4 78, GJ 87].

In this chapter, we present the two different approaches as put forth in the literature cited above. In section 4.1 we briefly summarize the ideas that enabled Lieb-Mattis to apply Perron-Frobenius theory to ground states of Hamiltonians. In section 4.3, after having introduced semigroups in section 4.2, we subsequently develop the perturbation theory of semigroups along the lines of Faris.

4.1 The Idea of Lieb-Mattis

Lieb-Mattis show the uniqueness of the ground state for a certain class of Hamiltonians, assuming its existence. Their basic idea is the following:

Let $\mathcal{H}$ be a real, finite-dimensional Hilbert space as defined in section 3.4.1. Let H be a Hamiltonian on $\mathcal{H}$ and assume that H can be written is the form

$$H = H_0 - V \tag{4.1.1}$$

where H_0 is assumed to be diagonal w.r.t. a given cone basis $\{e_i\}_{i=1,\ldots,n}$ in $\mathcal{H}$. Let $\mathcal{K} = \{e_i\}^+$ denote the fat cone in $\mathcal{H}$, i.e. the cone spanned by nonnegative linear combinations of the basis vectors e_i. Assume further that H_0 and V are both positivity preserving w.r.t. the fat cone $\mathcal{K}$.

Denote by $w = \sum_{i=1}^n w_i e_i$ the ground state of the Hamiltonian H. Then if the ground state energy is E_0, all other energy expectation values are larger, i.e.

$$E_M = (w, Hw) \leq (|w|, H|w|) \tag{4.1.2}$$

where $|w| = \sum_{i=1}^n |w_i| e_i$. Thus we have

$$(w, H_0 w) - (w, Vw) \leq (|w|, H_0|w|) - (|w|, V|w|)$$

and since H_0 is diagonal by hypothesis, it follows that

$$(w, H_0 w) = e_0(w, w) = e_0(|w|, |w|) = (|w|, H_0|w|)$$

and thus we are led to the inequality

$$(w, Vw) \geq (|w|, V|w|) \tag{4.1.3}$$

which in turn implies equality in (4.1.2). Therefore, if w is a ground state, then so is $|w| \in \mathcal{K}$. Additional assumptions with regard to V, equivalent to ergodicity, then assure that $w = |w|$ and w is strictly positive, thus implying uniqueness.

The essential tool that enabled us to carry out the above argument is the minus sign in equation (4.1.1). At first glance this sign seems to be a nuisance since the overall positivity of H is lost, but with its help, we are able to redo the Perron-Frobenius argument as presented in equation (2.2.3) (see also equation (3.4.15)) by circumventing the need of the spectral radius $\rho(\cdot)$ yet still producing the desired inequality. The minus sign even penetrates into the semigroup arguments presented in the next section.

4.2 Ergodic Semigroups

Definition (4.2.1) A family of bounded operators $\{T(t) : 0 \leq t \leq \infty\}$ on a Banach space $\mathcal{X}$ is called a *strongly continuous semigroup* if

(a) $T(0) = \mathbb{1}$,

(b) $T(s)T(t) = T(s+t)\ \forall\ s, t \in \mathsf{R}^+$,

(c) For each $\psi \in \mathcal{X}, t \mapsto T(t)\psi$ is continuous.

Such semigroups arise in the theory of partial differential equations (e.g. the heat equation $du/dt = -Tu$) and in quantum theory. Accounts devoted to their study are [P 53, K 66, RS2 78].

From the above definition, it is easy to see that semigroups are somewhat connected to exponentials. Let $T(t)$ be a strongly continuous semigroup on a Banach space $\mathcal{X}$. As is the case with unitary groups on Hilbert spaces[1], we obtain the *infinitesimal generator* of $T(t)$ by differentiation. Set

$$A_t = \frac{\mathbb{1} - T(t)}{t}. \tag{4.2.2}$$

Then the generator A of the semigroup $T(t)$ is given as

$$Ax = \text{s–}\lim_{t \downarrow 0} A_t x \tag{4.2.3}$$

for all $x \in \mathcal{X}$ such that the limit exists. We write

$$T(t) = \exp(-tA) \tag{4.2.4}$$

A fundamental question of the theory of semigroups in Banach spaces is whether, given a semigroup $T(t)$, does its generator A exist and if it does what is its domain $D(A)$. On the other hand, given an operator A on $\mathcal{X}$, what is the semigroup $T(t) = T(t, A)$ thus generated. Since in our case, we will always deal with bounded operators defined on the whole of $\mathcal{X}$, the first problem do not exits. As to the second problem, assume that an operator A is given. Then we can simply define the exponential on $\mathcal{X}$ by its Taylor series,

$$\exp(-tA) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} A^n \tag{4.2.5}$$

as we already did implicitly whenever we used the exponential during one of the last chapters. The group property

$$\exp(-sA)\exp(-tA) = \exp(-(t+s)A) \tag{4.2.6}$$

follows from (4.2.5) and A is indeed the generator, since by differentiation,

$$\frac{d}{dt}\exp(-tA) = -A\exp(-tA). \tag{4.2.7}$$

The applicatibility of semigroups to ground states via the Perron-Frobenius theory is due to the following simple fact: Let H be a self-adjoint and semibounded

[1]See M. Stone, *Linear Transformations in Hilbert space III*, Proc. Nat. Acad. Sci. U.S.A. **16**, 172, (1930)

operator. Then if $E_0 = \inf \sigma(H)$, $\sup \sigma(\exp\{-tH\}) = \exp(-tE_0)$ for all $t > 0$ by the spectral theorem.

In the next definition, we introduce a characterization of semigroups that will allow us to apply Perron-Frobenius theory to semigroups.

Definition (4.2.8) Let $\mathcal{K} \in \mathcal{H}$ be a Hilbert cone. Let H be a bounded, self-adjoint operator acting in $\mathcal{H}$, generating the semigroup $\exp(-tH)$. Assume that $\exp(-tH)$ is positivity preserving for all $t \geq 0$. Then $\exp(-tH)$, $t \geq 0$ is called an *ergodic semigroup*, if for all $u, v \in \mathcal{K}$, $u, v \neq 0$, there exists a $t \geq 0$ such that

$$(u, \exp(-tH)v) > 0 \tag{4.2.9}$$

Theorem (4.2.10) [Perron-Frobenius for semigroups] Let $\mathcal{K} \in \mathcal{H}$ be a Hilbert cone. Let H be a self-adjoint operator acting on $\mathcal{H}$. Assume that H is bounded below by an eigenvalue E_0. Assume further that $\exp(-tH)$ is positivity preserving for all $t \geq 0$. Then if $\exp(-tH)$ is an ergodic semigroup, E_0 is an eigenvalue of multiplicity one and the corresponding eigenspace can be chosen strictly positive.

Proof (4.2.11) Set $A = \exp(-tH)$. Then $0 \leq \sigma(A) \leq \rho(A) = \exp(-tE_0)$ The eigenvector w of H corresponding to E_0, is also an eigenvector of $\exp(-tH)$ corresponding to $\exp(-tE_0)$. A is positivity preserving by hypothesis and positivity improving for a fixed t. Alternatively we could define $A' \equiv \exp(-H)$ and then $(A')^t$ is ergodic if we generalize ergodicity to non-integral exponents t. Since the Perron-Froebenius argument holds for a positivity improving A as well as an ergodic A', the proof is completed. ∎

4.3 Perturbation Results

In this section, we develop some perturbation theorems for semigroups. The term 'perturbation' used here is not to be taken too technical, since we do not talk about the standard Rayleigh-Schrödinger expansion in powers of V. Indeed, there is no requirement as to the smallness of V, as has normally to be assured to yield convergence of the expansion. By perturbation, we merely wish to indicate, that the Hamiltonian of the system in question can be written in the form of equation (4.1.1).

The following theorem is an abstract version of Glimm-Jaffe [GJ 70] as in [F 72].

Theorem (4.3.1) [Perturbation Result 1] Let $\mathcal{K} \in \mathcal{H}$ be a Hilbert cone. Let H_0 and V be self-adjoint and bounded operators in $\mathcal{H}$ such that $H = H_0 - V$ is also bounded in $\mathcal{H}$. Assume that $\exp(-tH_0)$ and $\exp(tV)$ are positivity preserving for

all $t \geq 0$. Assume finally that $u, v \in \mathcal{K}$, $(u, v) = 0$ implies $(u, \exp(+tH_0)v) = 0$ for all $t \geq 0$. Then if $\exp(tV)$ is an ergodic semigroup, $\exp(-tH)$ is positivity preserving and generates an ergodic semigroup.

Proof (4.3.2) By the Trotter product formula [RS1 80],

$$\exp(-t(H_0 - V)) = \text{s}-\lim_{n\to\infty} \left\{\exp(-\frac{t}{n}H_0)\exp(\frac{t}{n}V)\right\}^n \tag{4.3.3}$$

Both $\exp(-\frac{t}{n}H_0)$ and $\exp(\frac{t}{n}V)$ are positivity preserving by hypothesis and thus by 3.1.28, $\exp(-tH)$ is positivity preserving.

Let $v \in \mathcal{K}$, $v \neq 0$. Define

$$M(v) = \{u \in \mathcal{K} | (u, \exp(-tH)v) = 0 \ \forall\, t > 0\}. \tag{4.3.4}$$

In order to prove the ergodicity of $\exp(-tH)$ it suffices to show that $M(v) = \{0\}$. We first observe that $M(v)$ is a closed cone and $\exp(-tH)$ leaves $M(v)$ invariant. The same is true for $\exp(tH_0)$. For suppose that $u, v \succeq 0$ and $(u, \exp(-tH)v) = 0$, then by assumption $(\exp(tH_0)u, \exp(tH)v) = 0$. But then again by the Trotter product formula we have

$$\exp(-t(H - H_0)) = \text{s}-\lim_{n\to\infty} \{\exp(-\frac{t}{n}H)\exp(+\frac{t}{n}H_0)\}^n$$

and thus $\exp(-t(H - H_0))$ leaves $M(v)$ invariant. But since $H - H_0 = -V$, this implies that $\exp(tV)$ leaves $M(v)$ invariant. Now given any $u \in M(v)$ then $\exp(tV)u \in M(v)$ and thus for all $t \geq 0$

$$(\exp(tV)u, v) = 0$$

But by construction we know that $\exp(tV)$ is an ergodic semigroup and that $v \neq 0$. Thus we have shown that $u \equiv 0$ and therefore $M(v) = \{0\}$. ∎

In the next theorem, we replace the final assumption of theorem 4.3.1 — namely that $\exp(tH_0)$ transforms orthogonal elements of $\mathcal{K}$ into orthogonal ones — with a mild positivity improving quality of $\exp(-tH_0)$.

Theorem (4.3.5) [Perturbation Result 2] Let $\mathcal{K} \in \mathcal{H}$ be a Hilbert cone. Let H_0 and V be self-adjoint and bounded operators in $\mathcal{H}$ such that $H = H_0 - V$ is also bounded in $\mathcal{H}$. Assume that $\exp(-tH_0)$ and $\exp(tV)$ are positivity preserving. Assume finally that $\exp(-tH_0)$ is positivity improving in the sense that

$$(u, \exp(-tH_0)v) \geq (u, v) \tag{4.3.6}$$

for all $t \geq 0, u, v \in \mathcal{K}$. Then if $\exp(tV)$ is an ergodic semigroup, $\exp(-tH)$ is positivity preserving and generates an ergodic semigroup.

Proof (4.3.7) Again as in theorem 4.3.1, by the Trotter product formula, $\exp(-tH)$ is positivity preserving. Let $u, v \in \mathcal{K}$, then

$$\begin{aligned}
&(u, \exp(-t(H_0 - V))v) \\
&= \lim_{n\to\infty} (u, (\exp(-t\frac{H_0}{n}) \exp(+t\frac{V}{n}))^n v) \\
&= \lim_{n\to\infty} (\exp(+t\frac{V}{n}) \exp(-t\frac{H_0}{n})u, (\exp(-t\frac{H_0}{n}) \exp(+t\frac{V}{n}))^{n-1} v) \\
&\geq \lim_{n\to\infty} (\exp(+t\frac{V}{n})u, (\ldots)^{n-1} v) \\
&\geq \lim_{n\to\infty} (\exp(+2t\frac{V}{n}u, (\ldots)^{n-2} v) \\
&\vdots \\
&\geq \lim_{n\to\infty} (\exp(+tV)u, v) \\
&= (\exp(+tV)u, v) \\
&> 0
\end{aligned}$$

where the last inequality follows by the ergodicity of $\exp(tV)$. Thus $\exp(-tH)$ is an ergodic semigroup. ∎

Note that if $-H_0$ is positivity preserving, than this is sufficient for equation (4.3.6) to hold.

Faris proved an additional perturbation result starting at the operator-valued function

$$R(\lambda) \equiv (H + c)^{-1}. \tag{4.3.8}$$

Although we will not use it in the course of the present thesis, we still include it here, since it may serve as a starting point in future investigations. The connection with the above theorems is immediate by

Proposition (4.3.9) Let $\mathcal{K} \in \mathcal{H}$ be a Hilbert cone. Let H be a bounded self-adjoint operator on $\mathcal{H}$. Assume that $\exp(-tH)$ is positivity preserving for all $t \geq 0$. Let c be a real number such that $-c < \inf \sigma(H)$. Then $\exp(-tH)$ is an ergodic semigroup iff $(H + c)^{-1}$ is positivity preserving.

Proof (4.3.10)

$$(u, (H + c)^{-1} v) = \int_0^\infty \exp(-tc)(u, \exp(-tH)v)dt \tag{4.3.11}$$

by the formal Laplace transformation [RS2 78, page 237]. ∎

Theorem (4.3.12) [Perturbation Result 3] Let $\mathcal{K} \subset \mathcal{H}$ and H_0, V be given as in theorem 4.3.1 and $H = H_0 - V$. Assume that $(H_0 + c)^{-1}$ is positivity preserving for

all $c \in \mathsf{R}$, such that $c + \inf \sigma(H_0) > 0$ and that V is positivity preserving. Assume also that $(H_0 + c)^{-1}$ is ergodic. Then $(H + c)^{-1}$ is positivity preserving and ergodic for all such c.

Proof (4.3.13) We first show that $(H + c)^{-1}$ is positivity preserving. Consider real numbers c such that $-c < \inf \sigma(H_0)$. Then we have the following expression

$$\begin{aligned} \|V(H_0 + c)^{-1}\| &\leq \|V\|\|(H_0 + c)^{-1}\| \\ &= \|V\|(\inf \sigma(H_0) + c)^{-1} \\ &< 1 \end{aligned}$$

for large enough c. Hence we have the convergent series expansion

$$\begin{aligned} (H + c)^{-1} &= (H_0 - V + c)^{-1} \\ &= (H_0 + c)^{-1}(1 - V(H_0 + c)^{-1})^{-1} \\ &= (H_0 + c)^{-1} \sum_{n=0}^{\infty} (V(H_0 + c)^{-1})^n \end{aligned} \tag{4.3.14}$$

for all such c. It follows from this series representation and the hypothesis $(H_0 + c)^{-1}, V \succeq 0$ that $(H + c)^{-1}$ is positivity preserving for all such c.

Let $E_0 = \inf \sigma(H)$. We now know that there exists a $-c < E_0$ with $(H + c)^{-1}$ positivity preserving. Consider a $d \in \mathsf{R}$ with $d \leq c$ and $-d < E_0$. Then

$$\begin{aligned} \|(c - d)(H + c)^{-1}\| &= (c - d)(E_0 + c)^{-1} \\ &= \frac{1 - d/c}{1 + E_0/c} \\ &< 1 \end{aligned}$$

So

$$(H + d)^{-1} = (H + c)^{-1} \sum_{n=0}^{\infty} [(c - d)(H + c)^{-1}]^n \tag{4.3.15}$$

is also positivity preserving and the first half of the proof is finished. We next show the ergodicity of H for all c sufficiently large. We may write

$$(H + c)^{-1} = (H_0 + c)^{-1} + T \tag{4.3.16}$$

where $T = (H_0 + c)^{-1}(V)(H + c)^{-1}$ is positivity preserving by the above arguments. If $(H_0 + c)^{-1}$ is ergodic, then there exists an $n \in \mathsf{N}$, such that for all $u, v \in \mathcal{K}$ we have

$$(u, (H_0 + c)^{-n} v) > 0 \tag{4.3.17}$$

Thus it immediately follows from above that

$$(u, (H + c)^{-n} v) = (u, [(H_0 + c)^{-1} + T]^n v) > 0 \tag{4.3.18}$$

and so $(H+c)^{-1}$ is ergodic for all such c. Finally, we show that $(H+d)^{-1}$ is ergodic for all d such that $c \geq d > -E_0$. But this follows from the expansion (4.3.15) and the ergodicity of $(H + c)^{-1}$. ∎

In closing, we remark that all the above theorems can be proved for a more general setting. The usual restrictions on the operators are then of the following type: Assume H_0 bounded below and V relatively bounded w.r.t. H_0 with relative bound less than 1.

Chapter 5

Spin-Lattice Models

In 1962, E. Lieb and D. Mattis studied the ordering of electronic energy levels in two models. In the first paper, they considered a one-dimensional chain of electrons under the dynamics of one-electron hopping [LM 62a]. The second paper deals with the ground state of the antiferromagnetic Heisenberg model in arbitrary dimensions [LM 62b]. Both papers use the idea presented in section 4.1 to prove the uniqueness of the ground state. In this chapter, after having introduced the notation we will use in the context of spin-lattice models, we will give a slightly alternative proof of the theorems of Lieb-Mattis by applying semigroup methods.

5.1 Preliminaries

In the subsequent chapters, we will study the ground state properties of various models of strongly correlated electron systems on a lattice of arbitrary dimension d, i.e.

Definition (5.1.1) Given a finite set $\Lambda = \{x, \ldots, z\}$ of elements of Z^d. We call Λ a *lattice* and each $x = \{x_1, \ldots, x_d\} \in \mathsf{Z}^d$ a *lattice site*. We will furthermore denote the *number of lattice sites* in the lattice by $|\Lambda|$.

Constructing a Hilbert space for the lattice Λ is done in two steps. First we will define a Hilbert space $\mathcal{H}(x)$ for each lattice site x independently. Essentially two different lattice site Hilbert spaces will be encountered in the following sections. For the Lieb-Mattis model, which we will study in this chapter, and the Hubbard model, to be studied in chapter 6, the Hilbert space at some lattice site x is given by the Fock space

$$\mathcal{H}(x) = \bigoplus_{n=0}^{2} \mathcal{H}_n(x) \tag{5.1.2}$$

where $\mathcal{H}_n(x)$ denotes the n-particle Hilbert space at site x. $\mathcal{H}_0(x)$ $(= \mathsf{C})$ is spanned by the no-particle (vacuum) state $|0\rangle$ and $\mathcal{H}_1(x)$ $(= \mathsf{C}^2)$ is the 1-particle space spanned by $|\uparrow\rangle$ and $|\downarrow\rangle$. $\mathcal{H}_2(x)$ is constructed from the 1-particle space by the antisymmetric tensor product, i.e. $\mathcal{H}_2(x) = \mathcal{H}_1(x) \otimes^a \mathcal{H}_1(x)$ $(= \mathsf{C})$ and spanned by $|\uparrow\downarrow\rangle$. On this fermionic Fock space $\mathcal{H}(x)$, we define the *creation operator* $c^\dagger_{x\sigma}$ such that[1]

$$\begin{aligned} c^\dagger_{x\sigma}|0\rangle &= |\sigma\rangle \\ c^\dagger_{x\sigma}|\sigma'\rangle &= (1-\delta_{\sigma\sigma'})|\sigma\sigma'\rangle \\ c^\dagger_{x\sigma}|\sigma\sigma'\rangle &= 0 \end{aligned} \tag{5.1.3}$$

where σ denotes the spin index and is as such either $\uparrow$ or $\downarrow$. The *annihilation operator* $c_{x\sigma}$ is correspondingly defined as

$$\begin{aligned} c_{x\sigma}|0\rangle &= 0 \\ c_{x\sigma}|\sigma'\rangle &= \delta_{\sigma\sigma'}|0\rangle \\ c_{x\sigma}|\sigma\sigma'\rangle &= |\sigma'\rangle \end{aligned} \tag{5.1.4}$$

$c^\dagger_{x\sigma}, c_{x\sigma}$ satisfy the usual anticommutation relations[2], i.e.

$$\{c_{x\sigma}, c^\dagger_{y\sigma'}\} = \delta_{xy}\delta_{\sigma\sigma'}, \tag{5.1.5}$$

$$\{c_{x\sigma}, c_{y\sigma'}\} = \{c^\dagger_{x\sigma}, c^\dagger_{y\sigma'}\} = 0 \tag{5.1.6}$$

With the help of the creation and annihilation operators, we next define the *number operators*,

$$n_x = n_{x\uparrow} + n_{x\downarrow} = c^\dagger_{x\uparrow}c_{x\uparrow} + c^\dagger_{x\downarrow}c_{x\downarrow}. \tag{5.1.7}$$

The Hilbert space for the Heisenberg model, which we study in section 5.3, has another form. At each site x we assume that there exists a particle with spin value $s(x)$. We denote the *spin operator* at x by $\sigma(x) = (\sigma_1(x), \sigma_2(x), \sigma_3(x))$ and the components of $\sigma(x)$ obey the commutation relation

$$[\sigma_1(x), \sigma_2(x)] = i\sigma_3(x) \tag{5.1.8}$$

The remaining relations follow by cyclic permutation of the indices. $\sigma(x)^2$ and $\sigma_3(x)$ may be diagonalized simultaneously, $s(x)[s(x)+1]$ and $m(x)$ being the eigenvalues, respectively and $m(x)$ lies in the interval $-s(x) \leq m(x) \leq s(x)$. Then $\mathcal{H}(x) = \mathsf{C}^{2s(x)+1}$.

[1]Instead of *, we will from now on denote the Hermitian adjoint of an operator by the symbol †, as is customary in our context.

[2]We use natural units such that $\hbar = 1$ from now on.

Next we construct the Hilbert space for the whole lattice by using either the symmetric or the antisymmetric tensor product of these lattice site spaces, i.e.

$$\mathcal{H}(\Lambda) \equiv \bigotimes_{x\in\Lambda}^{s} \mathcal{H}(x) \quad \text{or} \quad \mathcal{H}(\Lambda) \equiv \bigotimes_{x\in\Lambda}^{a} \mathcal{H}(x) \tag{5.1.9}$$

On this Hilbert space, we introduce operators in the following way.

Definition (5.1.10) Let $\alpha(x)$ be an operator on $\mathcal{H}(x)$, then $A(x)$ is an operator on $\mathcal{H}(\Lambda)$ defined by

$$A(x) = \underbrace{\mathbb{1} \otimes \cdots \otimes \overbrace{\alpha(x)}^{\text{site } x} \otimes \cdots \otimes \mathbb{1}}_{|\Lambda|\text{th tensor product}} \tag{5.1.11}$$

where the tensor product $\otimes$ may be symmetric or antisymmetric. In addition, we use the notation $A(\Lambda)$ to indicate that $A(\Lambda)$ operates on every lattice site of Λ. A typical such operator is e.g.

$$A(\Lambda) = \sum_{x\in\Lambda} A(x) \tag{5.1.12}$$

For the operators $c^{\dagger}_{x\sigma}, c_{x\sigma}$ as well as the number operators $n_{x\uparrow}, n_{x\downarrow}$, we will *not* distinguish their corresponding $\alpha(x)$ and $A(x)$. As is customary in the literature, both will be denoted by $c^{\dagger}_{x\sigma}, c_{x\sigma}$. But for the spin operators, we make the following definition:

Definition (5.1.13) Let $\sigma(x) = (\sigma_1(x), \sigma_2(x), \sigma_3(x))$ be the spin operator on the space $\mathcal{H}(x)$. Following our above introduced notation, we then construct spin operators on $\mathcal{H}(\Lambda)$ such that

$$\mathbf{S}(x) = \mathbb{1} \otimes^{s} \cdots \otimes^{s} \sigma(x) \otimes^{s} \cdots \otimes^{s} \mathbb{1} \tag{5.1.14}$$

and

$$\mathbf{S}(\Lambda) = \sum_{x\in\Lambda} \mathbf{S}(x), \tag{5.1.15}$$

The eigenvalues are denoted by $S(S+1)$ and M for $\mathbf{S}(\Lambda)^2$ and $S_3(\Lambda)$.

That these operators on $\mathcal{H}(\Lambda)$ indeed obey the spin commutation relations as did the components $\sigma_1(x), \sigma_2(x), \sigma_3(x)$ of the spin operator on $\mathcal{H}(x)$ can be seen by the next calculation.

$$\begin{aligned} &[S_1(\Lambda), S_2(\Lambda)] \\ &= [\sum_{x\in\Lambda} S_1(x), \sum_{y\in\Lambda} S_2(y)] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{x,y\in\Lambda} [1 \otimes^s \ldots \otimes^s \sigma_1(x) \otimes^s \ldots \otimes^s 1, 1 \otimes^s \ldots \otimes^s \sigma_2(y) \otimes^s \ldots \otimes^s 1] \\
&= \sum_{x,y\in\Lambda} 1 \otimes^s \ldots \otimes^s [\sigma_1, \sigma_2](x) \otimes^s \ldots \otimes^s 1 \; \delta(x,y) \\
&= \sum_{x,y\in\Lambda} 1 \otimes^s \ldots \otimes^s i\sigma_3(x) \otimes^s \ldots \otimes^s 1 \; \delta(x,y) \\
&= \sum_{x\in\Lambda} 1 \otimes^s \ldots \otimes^s i\sigma_3(x) \otimes^s \ldots \otimes^s 1 \\
&= iS_3(\Lambda) \qquad (5.1.16)
\end{aligned}
$$

The remaining relations follow again by cyclic permutation. In analogy to the Fermi creation and annihilation operators, we define *spin creation* and *annihilation* operators for the whole lattice such that

$$S_\pm(\Lambda) = \frac{1}{\sqrt{2}}(S_x(\Lambda) \pm i\, S_y(\Lambda)) \qquad (5.1.17)$$

and we may write

$$
\begin{aligned}
\mathbf{S}(x)\cdot\mathbf{S}(y) &= S_1(x)S_1(y) + S_2(x)S_2(y) + S_3(x)S_3(y) \\
&= S_+(x)S_-(y) + S_-(x)S_+(y) + S_3(x)S_3(y) \qquad (5.1.18)
\end{aligned}
$$

Using the Fermi operator representation of above, we may even construct spin operators in the Fock space such that

$$S_3(\Lambda) = \frac{1}{2}\sum_{x\in\Lambda}(n_{x\uparrow} - n_{x\downarrow}) \qquad (5.1.19)$$

$$S_+(\Lambda) = (S_-(\Lambda))^\dagger = \frac{1}{\sqrt{2}}\sum_{x\in\Lambda} c^\dagger_{x\uparrow}c_{x\downarrow} \qquad (5.1.20)$$

It can be shown that these spin operators indeed fulfill the spin commutation relations as in equation (5.1.16).

5.2 The Lieb-Mattis Model

Although the model under investigation by Lieb-Mattis in [LM 62a], which we will henceforth call the *Lieb-Mattis model*, is rather simple, it serves well to introduce the principal points of the application of semigroups.

Let Λ denote a one-dimensional chain consisting of discrete atoms labeled $x = 1, 2, \ldots, |\Lambda|$ separated by a distance l such that only nearest-neighbor overlap of the electron wave functions is important. We assume that each atom in this

chain has only one valence state (capable of double occupancy, however, because of spin degeneracy). Therefore the Hilbert space of the Lieb-Mattis model is the antisymmetric tensor product of the Fock space $\mathcal{H}(x)$ as discussed in the last section. The dynamics in this model is such that only *one-electron hopping*[3] is present. As a minor generalization of the original account, the corresponding real, symmetric *hopping matrix* $t(x,y)$ need not be constant; we require only that the sign of $t(x,y)$ is the same for all $x, y \in \Lambda$.

Aside from the hopping matrix, we assume that the Hamiltonian is diagonal and that the energy in a given state is calculable by specifying which atoms have empty, single occupied or double occupied valence states. This leads directly to the Hamiltonian

$$H_{LM} = \sum_{\substack{(x,y) \\ \sigma=\uparrow,\downarrow}} t(x,y) c^\dagger_{x\sigma} c_{y\sigma} + H_0(\ldots, n_z, \ldots) \tag{5.2.1}$$

where $H_0(\ldots, n_z, \ldots)$ is an arbitrary function of the number operators.

We first note that the energy spectrum of H_{LM} is invariant under an overall sign change of $t(x,y)$ [LW 68]. The proof assumes bipartiteness of Λ, that is

Definition (5.2.2) We call a lattice Λ *bipartite*, if the sites of Λ can be divided into two disjoint sets Λ_A, Λ_B such that $t(x,y) = 0$ whenever $x, y \in \Lambda_A$ or $x, y \in \Lambda_B$[4]. We note that by this condition, which is more an assertion about the hopping matrix $t(x,y)$ than the lattice Λ, we have $t(x,x) = 0$ for all $x \in \Lambda$.

Now, we may apply the unitary BCT transformation (see appendix A.4.1) which changes the overall sign of the hopping part of H_{LM} but leaves invariant the number operators $n_{x\sigma}$ and thus H_0. We therefore assume that $t(x,y) \leq 0$ for all $x, y \in \Lambda$ from now on. The Hamiltonian of the Lieb-Mattis model then reads as

$$H_{LM} = H_0(\ldots, n_z, \ldots) - \sum_{\substack{(x,y) \\ \sigma=\uparrow,\downarrow}} t(x,y) c^\dagger_{x\sigma} c_{y\sigma} \tag{5.2.3}$$

and we have the desired form of equation (4.1.1).

Let $N = \sum_{x\in\Lambda} n_x$ be the number operator[5]. N commutes with H_{LM} (charge conservation) and thus we may restrict ourselves to a subspace with a fixed and

[3]One-electron hops are those where one electron hops from its present site to another site, previously not occupied by an electron of the same spin direction. *Double hops* denote the exchange of two electrons at different sites and are related to spin exchange as we will show in chapter 7.

[4]See appendix A.1 for an equivalent definition of bipartiteness in the context of matrix theory.

[5]We do not distinguish the number operator N from its eigenvalue N by using different symbols. Since we will always work in subspaces with a fixed number of particles, this will not cause any problems.

even number of electrons, i.e. $N = 2n$. Necessarily, $N \leq |\Lambda|$, since each site can accommodate at most two electrons.

The Hamiltonian H_{LM} commutes with $\mathbf{S}(\Lambda)^2$ and $S_3(\Lambda)$. Thus let ψ denote some simultaneous eigenstate of H_{LM} and $S_3(\Lambda)$ with eigenvalues $E(\psi)$ and $M(\psi)$. Then $S_-(\Lambda)^M\psi \neq 0$ and $S_+(\Lambda)^{-M}\psi \neq 0$ and both states are $S_3(\Lambda)$, H_{LM} eigenstates and belong to $M = 0$, $E(\psi)$, respectively. Therefore the ground state in the $M = 0$ subspace is the ground state of the Hamiltonian.

In the $M = 0$ subspace, there are n electrons with spin up and n electrons with spin down. Thus it is spanned by

$$p = \begin{pmatrix} |\Lambda| \\ n \end{pmatrix}^2 \tag{5.2.4}$$

linear independent *pure states.* Each such pure state can be created by applying the corresponding creation operators to the no-particle (vacuum) state $|0\rangle$[6], i.e.

$$c^\dagger_{x_1\sigma_1} c^\dagger_{x_2\sigma_2} \cdots c^\dagger_{x_N\sigma_N} |0\rangle \tag{5.2.5}$$

Pursuing the path of Perron-Frobenius theory as outlined in the preceding chapters, we should now start to look for a suitable cone in the $M = 0$ subspace. Unfortunately, we immediately run into difficulties, since in any cone we might consider, *the ordering of the Fermi creation operators is extremely important.* To see this, suppose that we have selected an arbitrary but fixed ordering for the product of the creation operators for different configurations. We might for example choose the natural order as given by the linear ordering of the lattice sites. Then applying a creation or annihilation operator, located at some lattice site, to this state, we have to commute it through the product of creation operators to preserve this ordering. Since we have no means of knowing how many such commutations have to be performed until the operators are in their standard ordering again and each such commutation introduces a minus sign, it is clear that preserving positivity will be hard to insure.

It is thus fortunate that there exists a unitary transformation –the Jordan-Wigner transformation [JW 28]–which transforms the Fermi operators $c^\dagger_{x\sigma}, c_{x\sigma}$ into *Pauli operators* $b^+_{x\sigma}, b_{x\sigma}$. It is the advantage of Pauli operators that they commute at different lattice sites and anticommute only for $b^+_{x\sigma}, b_{x\sigma}$. In the appendix we show that the hopping form of the Hamiltonian of the Lieb-Mattis model as well as H_0 remain invariant under the Jordan-Wigner transformation if we have restricted ourselves to nearest-neighbor interactions. Thus, in terms of the Pauli operators,

[6]Note that this vacuum state $|0\rangle$ is actually the tensor product of the vacuum states for each lattice site x, i.e. $|0\rangle = |0\rangle \otimes^a \ldots \otimes^a |0\rangle$.

equation (5.2.3) becomes

$$H = H_0(\ldots, n_z, \ldots) - \sum_{\substack{(x,y)\\ \sigma=\uparrow,\downarrow}} t(x,y) b^+_{x\sigma} b_{y\sigma} \tag{5.2.6}$$

and the p distinct states, labeled by the multi-index $a = \{x_1\sigma_1, x_2\sigma_2, \ldots, x_N\sigma_N\}$ are

$$\phi_a = \text{const} \cdot b^+_{x_1\sigma_1} \cdots b^+_{x_N\sigma_N} |0\rangle \tag{5.2.7}$$

Since we are now dealing with Pauli operators, the sign of the constant in the above formula may be chosen as positive.

We may now apply the classical Perron-Froebenius arguments of chapter 2. The Hamiltonian (5.2.6) consists of two $p \times p$ matrices and as our cone, we select the fat cone spanned by the positive hyperoctant in the $M = 0$ subspace and denote it by $\mathcal{K}_{LM}$.

Lemma (5.2.8) The semigroup generated by H_0 is positivity preserving w.r.t. $\mathcal{K}_{LM}$.

Proof (5.2.9) H_0 is diagonal in the chosen basis. Thus let e_a be the eigenvalue of H_0 corresponding to the basis vector ϕ_a. Then for any $u = \sum_a^p u_a \phi_a \in \mathcal{K}_{LM}$, we have

$$H_0 u = H_0 \sum_a^p u_a \phi_a = \sum_a^p u_a H_0 \phi_a = \sum_a^p u_a e_a \phi_a$$

and thus

$$\exp(-tH_0)u = \sum_a^p u_a \exp(-tH_0)\phi_a = \sum_a^p u_a \underbrace{\exp(-te_a)}_{>0} \phi_a \in \mathcal{K}_{LM}$$

∎

The semigroup generated by the hopping part of the Hamiltonian is positivity preserving, since the $b^+_{x\sigma} b_{y\sigma}$ part does not change the overall sign if applied to a given configuration ϕ_a as can be seen by the next calculation.

$$\begin{aligned}
&(b^+_{x\sigma} b_{y\sigma})\ b^+_{x_1\sigma_1} \cdots b^+_{x_N\sigma_N} |0\rangle \\
&= b^+_{x\sigma} \sum_i b^+_{x_1\sigma_1} \cdots [b_{y\sigma}, b^+_{x_i\sigma_{x_i}}] b^+_{x_{i+1}\sigma_{x_{i+1}}} \cdots b^+_{x_N\sigma_N} |0\rangle \\
&= b^+_{x\sigma} \sum_i b^+_{x_1\sigma_1} \cdots \delta_{yx_i}\delta_{\sigma\sigma_{x_i}} \cdots b^+_{x_N\sigma_N} |0\rangle \\
&= \sum_{ij} \delta_{yx_i}\delta_{\sigma\sigma_{x_i}} b^+_{x_1\sigma_1} \cdots (b^+_{x\sigma}) b^+_{x_j\sigma_{x_j}} \cdots b^+_{x_N\sigma_N} |0\rangle \\
&= \sum_{ij} \underbrace{\delta_{yx_i}\delta_{\sigma\sigma_{x_i}} (\mathbb{1} - \delta_{xx_j}\delta_{\sigma\sigma_{x_j}})}_{\geq 0} b^+_{x_1\sigma_1} \cdots b^+_{x_N\sigma_N} |0\rangle
\end{aligned} \tag{5.2.10}$$

where we used the fact that the application of an annihilation operator to the vacuum state yields zero.

Before we may now apply the perturbation result of theorem 4.3.1, we have to show that the hopping part of the Hamiltonian– denoted by T from now on– generates an ergodic semigroup. We are therefore looking for a characterization of the hopping matrix that ensures that given two arbitrary configurations ϕ_a, ϕ_b, we have

$$(\phi_b, T^n \phi_a) > 0 \tag{5.2.11}$$

for some $n \in \mathsf{N}$. So let ϕ_a, ϕ_b be two configurations which have the same configuration at each lattice site except at site x, where $x \in a, x \notin b$, and at site y, where $y \notin a, y \in b$. Then there exist an integer n such that $(\phi_b, T^n \phi_a) > 0$ if and only if there exists a sequence of hopping matrix elements such that

$$t(x, x_2) \cdot t(x_2, x_3) \cdot \ldots \cdot t(x_{n-1}, y) > 0 \tag{5.2.12}$$

By remark 3.4.9, we know that the existence of such a sequence for all such pairs of states a, b is indeed equivalent to ergodicity of T. By theorem 3.2.16 this implies that the semigroup is also ergodic.

Before we may go on, we have to ask ourselves whether the assumption of ergodicity in spin-lattice models is generic. Thus, suppose that the hopping matrix is not connected. Then this means that there exists at least some pure state a, i.e. some configuration of electrons on Λ, such that $(\phi_b, T^n \phi_a) = 0$ for all $b \in \Im[1, p]$ and all $n \in \mathsf{N}$. This in turn implies that there exists at least one lattice site x such that *there is no hopping to and from* x. Generalizing this argument to more than one such state a, it means that the lattice separates into noninteracting sublattices. Excluding this case, we may thus in general assume connectivity of the hopping matrix.

We summarize the assertions obtained about the hopping part of the Lieb-Mattis Hamiltonian in the next theorem.

Theorem (5.2.13) If the hopping matrix $t(x, y)$ is connected, then $\exp(tT)$ is an ergodic semigroup w.r.t. $\mathcal{K}_{LM}$.

Applying theorem 4.3.1 and theorem 4.2.10, we then know that the Hamiltonian of the Lieb-Mattis model has an unique and strictly positive ground state ϕ_0 in the $M = 0$ subspace.

The above argument holds even for the case that H_0 is the zero operator, because its semigroup then reduces the unit matrix which is clearly positivity preserving w.r.t. the cone considered. But in this case, it is well-known that among the ground states, there is some that belongs to $S = 0$. The $H_0 = 0$ ground state is nonnegative, too and therefore not orthogonal to ϕ_0. H_{LM} and the special Hamiltonian $H_{H_0=0}$ both commute with $\mathbf{S}(x)^2$ and thus the ground state belongs to $S = 0$ in the $M = 0$ subspace and for an even number of electrons.

The same argument applies in some M-subspace, that is the ground state of H_{LM} belongs to $S = M$. In addition, an M-subspace includes all $S \geq M$ states and thus the ground states are all ordered such that we conclude

Theorem (5.2.14) [Lieb-Mattis]

$$E(S+1) > E(S) \tag{5.2.15}$$

5.3 The Heisenberg Model

The general Heisenberg Hamiltonian for a system of interacting spins on a lattice Λ of arbitrary dimension is given by

$$H = \sum_{x,y \in \Lambda} J(x,y)\, \mathbf{S}(x) \cdot \mathbf{S}(y). \tag{5.3.1}$$

where the real and symmetric $J(x,y)$ is called the *(spin-)exchange matrix.* The corresponding Hilbert space, as discussed in section 5.1, is given by $\mathcal{H}(\Lambda) = \bigotimes_{x\in\Lambda}^{s} \mathbb{C}^{2s(x)+1}$. The Heisenberg model is a reasonable candidate for describing magnetic systems, where the magnetic moments are localized.

For $J(x,y) \leq 0$, the model favors a parallel alignment of the individual lattice spins such that the model exhibits *ferromagnetism.* Assuming $J(x,y) \geq 0$ for all sites x, y, then the models describes either *ferrimagnetism* (when the individual spins are of unequal magnitude) or *antiferromagnetism* (when the spins are of equal magnitude) *if we restrict ourselves to nearest-neighbor interaction only.* Otherwise, the model may be *frustrated* by a negative exchange matrix. That is, e.g. for the linear chain of three elements $\{x, y, z\}$, the nearest-neighbor spins x, y and y, z want to align antiparallel as do the next-nearest-neighbor spins x, z which is impossible. Finally, if none of the above cases applies, that is, $J(x,y)$ is of arbitrary magnitude and sign for each $x, y \in \Lambda$, then the Heisenberg model describes a so-called *spin-glass.*

In this section, we will prove uniqueness of the ground state of the antiferromagnetic (ferrimagnetic) and the ferromagnetic Hamiltonian. We start with the antiferromagnetic case.

5.3.1 The Antiferromagnetic Heisenberg

We are interested in proving uniqueness of the ground state for a lattice of arbitrary dimension. We are therefore looking for a characterization of the $J(x,y)$ matrix

that circumvents the above mentioned problem of frustration. The solution is a modification of the bipartiteness introduced in section 5.2.

Definition (5.3.2) [Bipartite Condition] Given a collection of sites Λ, then Λ is said to be *bipartite* if the sites of Λ can be divided into two disjoint sets Λ_A and Λ_B, such that for all $x_A, y_A \in \Lambda_A$ and all $x_B, y_B \in \Lambda_B$

$$J(x_A, y_B) \geq 0, \quad J(x_B, y_A) \geq 0. \tag{5.3.3}$$

and (in contrast to the previous definition),

$$J(x_A, y_A) \leq 0, \quad J(x_B, y_B) \leq 0 \tag{5.3.4}$$

Correspondingly, we will define spin operators for the disjoint sublattices in the manner given in definition 5.1.10, i.e.

$$\mathbf{S}(\Lambda_A) = \sum_{x \in \Lambda_A} \mathbf{S}(x), \tag{5.3.5}$$

and

$$\mathbf{S}(\Lambda_B) = \sum_{x \in \Lambda_B} \mathbf{S}(x). \tag{5.3.6}$$

As desired, the above requirement gives a tendency for antiparallel nearest neighbor alignment and parallel next-nearest neighbor alignment for the Hamiltonian (5.3.1). Henceforth, we will thus assume that the lattice Λ is bipartite.

For the application of the Perron-Frobenius theorem, we would like to write our Hamiltonian in the form of equation (4.1.1). By the bipartite condition, we may now apply the bipartite canonical transformation again, but this time instead of Fermi creation and annihilation operators, we are dealing with spin creation and annihilation operators. Thus let

$$\mathbf{S}(x_A) = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}(x_A) \rightarrow \begin{pmatrix} -S_1 \\ -S_2 \\ S_3 \end{pmatrix}(x_A) \tag{5.3.7}$$

and

$$\mathbf{S}(x_B) \rightarrow \mathbf{S}(x_B). \tag{5.3.8}$$

This transformation does not change any of the dynamics[7], but rather shows the intrinsic symmetry properties introduced by the bipartite condition in a more obvious way.

[7]The transformation rotates the spins on the Λ_A sublattice in the S_1, S_2-plane by 180° and is therefore given by the unitary matrix $U = \exp(i\pi \sum_{x \in \Lambda_A} S_3(x))$. This is similar to the usual BCT.

Theorem (5.3.9) The antiferromagnetic Heisenberg Hamiltonian on a bipartite lattice can be written as

$$\begin{aligned} H_{AF} &= H_0 - V \\ &= \sum_{x,y\in\Lambda} J(x,y)S_3(x)S_3(y) - \sum_{x,y\in\Lambda} |J(x,y)|\{S_+(x)S_-(y) + h.c.\} \end{aligned}$$

Proof (5.3.10) By the help of equation (5.1.18), we have

$$\begin{aligned} H(\Lambda) &= \sum_{x,y\in\Lambda} J(x,y)\ \mathbf{S}(x)\cdot\mathbf{S}(y) \\ &= \sum_{x,y\in\Lambda} J(x,y)\{S_+(x)S_-(y) + S_-(x)S_+(y) + S_3(x)S_3(y)\} \\ &= \sum_{x,y\in\Lambda} J(x,y)S_3(x)S_3(y) \\ &\quad - \sum_{x,y\in\Lambda} |J(x,y)|\{S_+(x)S_-(y) + h.c.\}, \end{aligned}$$

since

$$\begin{aligned} J(x_A,y_A)S_+(x_A)S_-(x_A) &\to -|J(x_A,y_A)|S_+(x_A)S_-(x_A) \\ J(x_B,y_B)S_+(x_B)S_-(x_B) &\to -|J(x_B,y_B)|S_+(x_B)S_-(x_B) \end{aligned}$$

and

$$\begin{aligned} J(x_A,y_B)S_+(x_A)S_-(x_B) &\to -|J(x_A,y_B)|S_+(x_A)S_-(x_B) \\ J(x_B,y_A)S_+(x_B)S_-(x_A) &\to -|J(x_B,y_A)|S_+(x_B)S_-(x_A) \end{aligned}$$

∎

Thus, using the modified bipartite condition and after performing the canonical transformation (5.3.7), we have succeeded in rewriting the antiferromagnetic Heisenberg model in the form of equation (4.1.1).

As in the last section, the total spin $\mathbf{S}(\Lambda)$ is a conserved quantity and by $S(S+1)$ and M we denote the eigenvalues of $\mathbf{S}(\Lambda)^2$ and $S_3(\Lambda)$, respectively. We again restrict our attention to some M-subspace and construct a basis for this M-subspace. Assume that we denote an eigenfunction of $\sigma_3(x)$ by $|m(x)\rangle$. Then we choose the basis set to consist of all distinct eigenfunctions ϕ_a of $S_3(\Lambda)$, compatible with an overall eigenvalue M, i.e.

$$\phi_a = \bigotimes_{x\in\Lambda}^{s} |m(x)\rangle \tag{5.3.11}$$

and as in the Lieb-Mattis model, the subscript a labels the finite number of possible such configurations, which we will again denote by p. We then have for some ϕ_a

$$\begin{aligned} S_3(\Lambda)\phi_a &= \sum_{x\in\Lambda} S_3(x) \bigotimes_{x\in\Lambda}^{s} |m(x)\rangle \\ &= \sum_{x\in\Lambda} m(x) \bigotimes_{x\in\Lambda}^{s} |m(x)\rangle \\ &= \left(\sum_{x\in\Lambda} m(x)\right) \phi_a \\ &= M\phi_a, \end{aligned}$$

Definition (5.3.12) Constructing our Hilbert cone for the application of the Perron-Frobenius arguments, we choose the phase of each basis vector ϕ_a such that

$$\phi_a = C \bigotimes_{x\in\Lambda}^{s} S_+^{s(x)+m(x)}(x)|-s(x)\rangle, \tag{5.3.13}$$

where $C \geq 0$ and $s(x)[s(x)+1]$ is the eigenvalue for $\mathbf{S}(x)^2$ on each lattice site. Our actual cone, denoted by $\mathcal{K}_{HBG}$, is then given by all possible linear combinations of those ϕ_a, such that the superposition coefficients, that is *the fat cone of the positive hyperoctant*, is positive again.

We note that this can be done because of the commutation relations of the spin operators. Indeed, by the well-known connection between spin operators and the Pauli operators $b_{x\sigma}^+, b_{x\sigma}$ [B 88], it is clear that in contrast to the Lieb-Mattis model encountered in the last section, we need not employ any special transformation at this point.

Again, H_0 generates a positivity preserving semigroup since it is diagonal in the basis chosen. In order to show that the spin exchange part V of theorem 5.3.9 is positivity preserving, we will show that $V\phi \succeq 0$ holds for an arbitrary basis vector $\phi_a \in \mathcal{K}_{HBG}$.

Theorem (5.3.14) The spin-exchange part of the bipartite Heisenberg model of theorem 5.3.9 is positivity preserving and generates an ergodic semigroup w.r.t. the cone $\mathcal{K}_{HBG}$.

Proof (5.3.15) Let $\phi_a \in \mathcal{K}_{HBG}$ be some basis vector as defined above. Then

$$V\phi_a = +\sum_{x,y\in\Lambda} |J(x,y)|\{S_+(x)S_-(y) + h.c.\}\ \phi_a$$

$$= + \sum_{x,y\in\Lambda} |J(x,y)|\{S_+(x)S_-(y) + h.c.\} \times$$
$$C \bigotimes_{z\in\Lambda}^{s} S_+^{s(z)+m(z)}(z)|-s(z)\rangle$$

One can easily see that the only terms that might violate the positivity preserving property of V are the ones originating from the equal-place commutation relations[8] which give a negative sign for the commutator, i.e.

$$\begin{aligned} S_-S_+ &= [S_-, S_+] + S_+S_- \\ &= -S_3 + S_+S_- \end{aligned} \tag{5.3.16}$$

Using the operator identity of lemma A.3.1, we now calculate these factors.

$$\begin{aligned} S_-S_+^{s+m}|-s\rangle &= \sum_{k=0}^{s+m-1} S_+^k[S_-, S_+]S_+^{s+m-1-k}|-s\rangle \\ &\quad + S_+^{s+m}\underbrace{S_-|-s\rangle}_{0} \\ &= \sum_{k=0}^{s+m-1} S_+^k(-S_3)S_+^{s+m-1-k}|-s\rangle \\ &= \sum_{k=0}^{s+m-1} S_+^k(-S_3)c|m-1-k\rangle \\ &= \sum_{k=0}^{s+m-1} S_+^k(-m+1+k)S_+^{s+m-1-k}|-s\rangle \\ &= \left\{\sum_{k=0}^{s+m-1}(-m+1+k)\right\} S_+^{s+m-1}|-s\rangle \\ &= \frac{1}{2}(s-m+1)(s+m)S_+^{s+m-1}|-s\rangle \\ &= \frac{1}{2}\{s(s+1)-m(m-1)\}S_+^{s+m-1}|-s\rangle \end{aligned}$$

Since $-s \le m \le +s$, it follows that $\{s(s+1)-m(m-1)\} \ge 0$ and thus

$$S_-S_+^{s+m}\mathcal{K}_{HBG} \subset \mathcal{K}_{HBG} \tag{5.3.17}$$

and thus V is positivity preserving.

[8]We drop the lattice site labels and set $\hbar = 1$ as before.

Assuming that the spin-exchange matrix $J(x,y)$ is connected, we know that for arbitrary $u, v \in \mathcal{K}_{HBG}$, $u, v \neq 0$, there exist an $n \in \mathbb{N}$ such that $(u, V^n v) > 0$. Then the generated semigroup may be called ergodic, since by the series expansion, we again have

$$(u, \exp(tV)v) = (u,v) + t(u,Vv) + \frac{t^2}{2}(u,V^2v) + \ldots > 0, \tag{5.3.18}$$

∎

We summarize these results to yield our major theorem about the antiferromagnetic Heisenberg model.

Theorem (5.3.19) Given the Hamiltonian of theorem 5.3.9, let $\mathcal{K}_{HBG}$ be the fat cone spanned by the basis vectors of definition 5.3.12 in a given M-subspace. Assume that $J(x,y)$ is bipartite and connected. Then H_{HBG} has a unique ground state in this M-subspace.

Proof (5.3.20) We apply theorem 4.3.1 to show that H_{AF} generates an ergodic semigroup which by theorem 4.2.10 implies that the corresponding ground state is unique. ∎

As was the case in the Lieb-Mattis model, we now try to specify the spin of the ground state. As it will turn out, the total spin is $S = 0$ for an antiferromagnet and some nonzero constant for a ferrimagnet. The mechanics of the proof is again as on page 57. We compare the ground state of some easy solvable system, which in addition is a special case of the Heisenberg Hamiltonian, with the one for the antiferromagnet (ferrimagnet). Since they are not orthogonal, their total spin value must be the same.

Theorem (5.3.21) Let us define the special Hamiltonian H' by choosing the $J(x,y)$ matrix elements such that

$$J(x_A, y_A) = J(x_B, y_B) = 0, \tag{5.3.22}$$

$$J(x_A, y_B) = J > 0. \tag{5.3.23}$$

Then the lowest energy belonging to each spin is given by

$$E(S) = J\{S(S+1) - S_A(S_A+1) - S_B(S_B+1)\} \tag{5.3.24}$$

where $S_A = \sum_{x \in \Lambda_A} s(x)$ and $S_B = \sum_{x \in \Lambda_B} s(x)$ are the maximum possible spin values on the bipartite sublattices. Therefore the ground state belongs to $S = \mathcal{S} \equiv |S_A - S_B|$.

Proof (5.3.25) The above model has infinite-range forces such that the Hamiltonian reduces to the interaction of the overall spins on the bipartite lattice, i.e.

$$\begin{aligned} H' &= \sum_{x,y\in\Lambda} J(x,y)\,\mathbf{S}(x)\cdot\mathbf{S}(y) \\ &= 2\,J\sum_{x\in\Lambda_A}\mathbf{S}(x)\cdot\sum_{x\in\Lambda_B}\mathbf{S}(x) \\ &= 2J\;\mathbf{S}(\Lambda_A)\cdot\mathbf{S}(\Lambda_B) \\ &= J\{(\underbrace{\mathbf{S}(\Lambda_A)+\mathbf{S}(\Lambda_B)}_{\mathbf{S}(\Lambda)})^2-\mathbf{S}(\Lambda_A)^2-\mathbf{S}(\Lambda_B)^2\}. \end{aligned}$$

The eigenvalues are thus given by

$$E(S)=J\{S(S+1)-s(\Lambda_A)(s(\Lambda_A)+1)-s(\Lambda_B)(s(\Lambda_B)+1)\}.$$

Correspondingly, the lowest eigenvalue of H' for each $S\geq\mathcal{S}$ value is given by

$$E(S)=J\{S(S+1)-S_A(S_A+1)-S_B(S_B+1)\} \tag{5.3.26}$$

and the ground state belongs to $S=\mathcal{S}$. ∎

Theorem (5.3.27) The ground state of the bipartite isotropic Heisenberg model belongs to total spin $S=\mathcal{S}$.

Proof (5.3.28) We now go into the $M=0$ subspace. Since theorem 5.3.19 applies for the special Hamiltonian 5.3.21 too, the ground state w' of H' is nonnegative (although not necessarily positive). Therefore $(w',w)>0$ and thus w and w' are not orthogonal. $\mathbf{S}^2$ and H,H' commute, so every ground state is an eigenstate of $\mathbf{S}^2$, i.e.

$$S'(S'+1)(w',w)=(\mathbf{S}^2w',w)=(w',\mathbf{S}^2w)=S(S+1)(w',w) \tag{5.3.29}$$

and thus $S=S'$. Since by theorem 5.3.21 the ground state of H' has a spin value of $S'=\mathcal{S}$ in the given $M=0$ subspace, so does H. Therefore the spin value of H is $S=\mathcal{S}$, too. ∎

Remark (5.3.30) For the antiferromagnetic case we have $\mathcal{S}=0$, and thus the ground state belongs to a total spin value $S=0$.

5.3.2 The Ferromagnetic Heisenberg

The ferromagnetic Heisenberg Hamiltonian is given by

$$H_F(\Lambda) = -\sum_{x,y\in\Lambda} J(x,y)\ \mathbf{S}(x)\cdot\mathbf{S}(y) \tag{5.3.31}$$

and $J(x,y) \geq 0$ for all $x, y \in \Lambda$, favoring parallel alignment of spins throughout the lattice Λ. Using relation (5.1.18), the Hamiltonian may be easily written in a form compatible with (4.1.1).

$$\begin{aligned} H_F(\Lambda) &= -\sum_{x,y\in\Lambda} J(x,y)\ \mathbf{S}(x)\cdot\mathbf{S}(y) \\ &= -\sum_{x,y\in\Lambda} J(x,y)\{S_+(x)S_-(y) + S_-(x)S_+(y) + S_3(x)S_3(y)\} \\ &= -\sum_{x,y\in\Lambda} J(x,y)S_3(x)S_3(y) \\ &\quad -\sum_{x,y\in\Lambda} J(x,y)\{S_+(x)S_-(y) + h.c.\} \\ &= H_0 - V \end{aligned}$$

The only difference from the Hamiltonian of 5.3.9 is that the sign of the diagonal part is changed. This does not alter the positivity preserving quality of the semigroup generated by H_0, and thus we may again conclude

Theorem (5.3.32) The ground state of the ferromagnetic connected Hamiltonian H_F is unique in each M-subspace, and thus up to the usual $2S+1$ degeneracy.

Chapter 6

The Hubbard Model

As pointed out in the introduction, this work has its origin in a 1989 publication of E. Lieb in *Physical Review Letters* [L 89]. In this paper, Lieb applied the ideas put forth in the preceding chapters. But as we will see now, the application of the Perron-Frobenius arguments to the Hubbard model is considerably more subtle than in the case of the Lieb-Mattis or the Heisenberg models considered in the last chapter. The problems that arise are largely due to the special structure of the cone for the Hubbard model.

6.1 Introduction

The Heisenberg model introduced in section 5.3 deals with localized permanent magnetic moments and thus is a good model for magnetic insulators only. In order to describe successfully magnetic metals such as Fe, Co, Ni, where the band electrons are responsible for both conductance and magnetism, we need another model.

Such a model was proposed by J. Hubbard in a series of papers[1] in the beginning of the sixties. The *Hubbard model* merges the *bandlike* behavior of the Lieb-Mattis model and the *localized* behavior of the Heisenberg model, since it is a model of itinerant electrons which interact via Coulomb forces. The rather large set of bound and continuum electron levels of each lattice ion is reduced to a *single localized orbital s-level* and the Coulomb interaction is assumed to be *screened*, such that only "on-site" interaction terms have to be included explicitly. The second-quantized Hamiltonian on the finite lattice Λ is given by

$$H = \sum_{\substack{x,y\in\Lambda \\ \sigma=\uparrow,\downarrow}} t(x,y)c^{\dagger}_{x\sigma}c_{y\sigma} + \sum_{x\in\Lambda} U(x)n_{x\uparrow}n_{x\downarrow} \tag{6.1.1}$$

[1] J. Hubbard, Proc. Roy. Soc. **A276**, 238, (1963), **A277**, 237, (1964), **A281**, 401, (1964)

where we use the notation of chapter 5, and $U(x)$ denotes the Coulomb interaction parameter at site x. This model Hamiltonian is a somewhat generalized version of the standard Hubbard model, in which the hopping is usually restricted to nearest-neighbors only and the Coulomb parameter $U(x)$ is assumed to be independent of the lattice site and positive, $U \geq 0$.

The Hilbert space at each lattice site is the same as in the case of the Lieb-Mattis model, i.e. it is the tensor product of four states per site: $|0\rangle$ is again the no-particle state (at site x), $|\uparrow\rangle$ and $|\downarrow\rangle$ represent a spin up, spin down electron, respectively, and $|\uparrow\downarrow\rangle$ is the up-down pair.

The Hubbard Hamiltonian obeys charge conservation, that is the total charge $Q = e\sum_{x\in\Lambda} n_x = eN$ is constant. In addition, the model is invariant under a $SU(2)$ rotation of the quantization axis. The last assertion is clearly true for the hopping part of equation (6.1.1). W.r.t. the on-site potential term, we note that with the help of the relations (5.1.19, 5.1.20), we may calculate that

$$\begin{aligned}
\mathbf{S}(x)^2 &= S_3(x)^2 + (S_+(x)S_-(x) + h.c.) \\
&= \frac{1}{4}(n_{x\uparrow} - n_{x\downarrow})^2 + \frac{1}{2}(c^\dagger_{x\uparrow}c_{x\downarrow}c^\dagger_{x\downarrow}c_{x\uparrow} + h.c.) \\
&= \frac{1}{4}(n^2_{x\uparrow} + n^2_{x\downarrow}) - \frac{1}{2}n_{x\uparrow}n_{x\downarrow} + \frac{1}{2}(n_{x\uparrow}(1 - n_{x\downarrow}) + h.c.) \\
&= \frac{3}{4}n_x - \frac{3}{2}n_{x\uparrow}n_{x\downarrow}
\end{aligned}$$

and we have the following identity

$$\sum_{x\in\Lambda} U(x)n_{x\uparrow}n_{x\downarrow} = -\frac{2}{3}\sum_{x\in\Lambda} U(x)\mathbf{S}(x)^2 + \frac{1}{6}\sum_{x\in\Lambda} U(x)n_x \qquad (6.1.2)$$

It is clear that the right-hand side of this equation is spin invariant.

For $U(x) \geq 0$, the Coulomb interaction is *repulsive* because singly occupied lattice sites are favored over doubly occupied ones. Two special cases, dealing with the relative strength of the hopping matrix $t(x,y)$ and the on-site repulsion $U(x)$, may be studied.

For $U \gg t$ and $N = |\Lambda|$ (half-filled band), we consider the *strong-coupling limit* $U \to \infty$. Doubly occupied states are not only minorized but forbidden in this limit[2]. Thus only $|\uparrow\rangle$ and $|\downarrow\rangle$ states are present and they are all degenerate eigenstates of the interaction part of the Hamiltonian[3]. We will now include the effects of the hopping part of H as in a perturbation expansion in terms of $\frac{t}{U}$ in order to derive the effective Hamiltonian in this limit. We first note that the hopping part of H creates pairs of electrons at some sites and thus its first-order matrix elements

[2]This property is called *hard-core repulsion.*

[3]Note that we restrict ourselves to $U = \text{const}$ in this argument.

are zero. We therefore have to extend the expansion to second-order where virtual states of paired electrons are possible. These states are larger than the ground states by an energy U. The formation of the virtual states may occur in two ways, and so the strength of the resulting perturbation is equal to $2t^2/U$, t^2 because the matrix elements in second-order are squared. In addition, since the only allowed transitions are of the spin exchange type, a reasonable candidate for the effective Hamiltonian is the antiferromagnetic (quantum) Heisenberg Hamiltonian,

$$H_{U \gg t} = \frac{2t^2}{U} \sum_{x,y} \mathbf{S}(x) \cdot \mathbf{S}(y) \tag{6.1.3}$$

A rigorous calculation of the above hand-waving arguments can be found in [E 79].

Apart from half-filling, the Hilbert space includes *holes* as well as singly occupied sites. Thus the hopping part of the Hamiltonian may move electrons from occupied sites to these holes, and charge is transported. The effective Hamiltonian in this case is the *tJ-model* which we will study in the next chapter.

In the *weak-coupling limit* $U \ll t$, the Coulomb interaction is treated as a perturbation of the otherwise *free-electron gas*. Such a situation is usually called a *Fermi liquid* and describes a metallic solid.

Exact solutions and rigorous results about the Hubbard model are rare and mostly treat only special cases. For the one-dimensional nearest-neighbor chain, E. Lieb and F. Y. Wu have solved the Hubbard model explicitly [LW 68]. We already know by theorem 5.2.14 that the ground state in this case is antiferromagnetic. Furthermore, for $U > 0$, Lieb and Wu showed that the ground state for a half-filled band is insulating for any nonzero U, and conducting for $U = 0$. Thus there is no so-called *Mott transition* between insulating and conducting state for nonzero U.

Another famous result for the Hubbard model is the absence of spontaneous magnetization in one or two dimensions at any temperature $T > 0$. This fact is known as the *Mermin-Wagner theorem of the Hubbard model*, and is named after N. D. Mermin and H. Wagner[4] who proved the theorem for the Heisenberg model. The proof for the Hubbard model uses completely analogous methods (Bogoliubov's inequality) and is due to D. K. Ghosh[5].

The attractive Hubbard model ($U(x) \leq 0$) is related to the repulsive case via a *hole-particle transformation*. It favors the building of spin pairs which indicates that the ground state should have total spin $S = 0$. This, among other things, we will prove in this chapter.

[4]N. D. Mermin, H. Wagner, Phys. Rev. Lett. **17**, 1133, (1966)
[5]D. K. Ghosh, Phys. Rev. Lett. **27**, 1584, (1971)

6.2 The Equivalent Matrix Problem

As in the previous chapter, we now reformulate the problem so that we end up dealing with matrices in a real and finite-dimensional vector space. By spin conservation, we again choose a special M subspace in which to work. All other subspaces belonging to some $\mathbf{S}^2$ value have representatives in the $M = 0$ subspace and we choose it for that reason. Let N be the (even) number of electrons in Λ. Then there are $n = N/2$ electrons of either spin up or and spin down type in this subspace.

Let $\phi_{a\uparrow}$ denote any distribution of spin up electrons on Λ. There are

$$p = \begin{pmatrix} |\Lambda| \\ n \end{pmatrix} \tag{6.2.1}$$

different such pure states and thus $\{\phi_{a\uparrow}\}_{a=1,\ldots,p}$ is a complete set for the spin up electrons. In addition, we require that each such $\phi_{a\uparrow}$ is a *real* polynomial in the $c^\dagger_{x\sigma}$'s acting on the vacuum state $|0\rangle$.

The same description holds for the spin down electrons. Therefore any configuration Φ of spin down and spin up electrons on Λ may be written as a linear combination of pure states, i.e.

$$\Phi = \sum_{a,b}^{p} a_{ab}\phi_{a\uparrow} \otimes \phi_{b\downarrow} \tag{6.2.2}$$

Here $A = [a_{ab}] \in M_{p,p}(\mathsf{C})$ and the configuration Φ is completely determined by the complex matrix A.

Let $W \in M_{p,p}$ denote some eigenstate of the Hamiltonian. By our construction, the basis vectors are real. In addition, the Hamiltonian is symmetric w.r.t. the spin index (that is, between spin up and spin down). Thus W^* is an eigenstate for the same eigenvalue. By linearity, the same holds for the sums $W + W^*$ and $i(W - W^*)$. We may therefore restrict our attention to those configurations that may be described by *Hermitian* matrices A and reduce our space to the space $\mathcal{V}$ of all Hermitian matrices, which has been introduced in definition 3.6.5. In this representation, the Hamiltonian (6.1.1) is a real $p^2 \times p^2$ matrix[6] acting on the Hermitian $p \times p$ matrices.

Let us now derive the Schrödinger equation in this representation. Suppose that $E(W)$ is the eigenvalue corresponding to the eigenstate $\Phi = \sum_{ab} W_{ab}\phi_{a\uparrow} \otimes \phi_{b\downarrow}$ which in order to streamline our notation during the following calculations will be written as $\sum_{ab} W_{ab}|a\rangle \otimes |b\rangle$. The spin directions are assumed implicitly by the

[6]Such $p^2 \times p^2$ matrices will be denoted by boldface characters such as $\mathbf{H}$.

ordering in the tensor product. Then

$$\begin{aligned} H\Phi &= \sum_{\substack{x,y\in\Lambda \\ \sigma}} t(x,y) c^\dagger_{x\sigma} c_{y\sigma} \sum_{ab} W_{ab}|a\rangle \otimes |b\rangle \\ &\quad + \sum_{x\in\Lambda} U(x) n_x \otimes n_x \sum_{ab} W_{ab}|a\rangle \otimes |b\rangle \\ &= \sum_{ab} W_{ab} \Bigg[\sum_{x,y\in\Lambda} t(x,y) \left\{ c^\dagger_x c_y |a\rangle \otimes |b\rangle + |a\rangle \otimes c^\dagger_x c_y |b\rangle \right\} \\ &\quad + \sum_{x\in\Lambda} U(x) n_x |a\rangle \otimes n_x |b\rangle \Bigg] \end{aligned}$$

We first treat the hopping part in this equation.

$$\begin{aligned} &\sum_{x,y\in\Lambda} t(x,y) \left\{ c^\dagger_x c_y |a\rangle \otimes |b\rangle + |a\rangle \otimes c^\dagger_x c_y |b\rangle \right\} \\ &= \sum_{x,y\in\Lambda} t(x,y) \sum_{c\in\Lambda} \left\{ |c\rangle\langle c| c^\dagger_x c_y |a\rangle \otimes |b\rangle + |a\rangle \otimes |c\rangle\langle c| c^\dagger_x c_y |b\rangle \right\} \\ &= \sum_c \left\{ T_{ac}|c\rangle \otimes |b\rangle + |a\rangle \otimes T_{bc}|c\rangle \right\} \end{aligned}$$

where

$$T_{ab} \equiv \langle b| \sum_{x,y\in\Lambda} t(x,y) c^\dagger_x c_y |a\rangle \tag{6.2.3}$$

Including the summation over a and b, this yields for the hopping part the following expression.

$$\sum_{abc} W_{ab} T_{ac} |c\rangle \otimes |b\rangle + \sum_{abc} W_{ca} T_{ab} |c\rangle \otimes |b\rangle$$

Since $t(x,y)$ is real and symmetric, so is T_{ab}.

The on-site interaction term can be calculated in a complete analogous fashion by inserting the identity $\sum_c |c\rangle\langle c|$ twice. Defining

$$(L_x)_{ab} \equiv \langle b| n_x |a\rangle, \tag{6.2.4}$$

this then yields

$$\sum_x U(x) \sum_{abcd} (L_x)_{ac} W_{ad} (L_x)_{db} |c\rangle \otimes |b\rangle$$

as the equation for the on-site term. In addition, $(L_x)_{ab}$ is real and symmetric, too. Thus the Schrödinger equation for the Hubbard model in matrix formulation

reads

$$\mathbf{H}W = TW + WT + \sum_x U(x) L_x W L_x \tag{6.2.5}$$

and if we suppose that $E(W)$ is the energy expectation value corresponding to the state Φ, we have

$$(\Phi, H\Phi) = \langle W, \mathbf{H}W \rangle = 2 \text{ trace}(TW^2) + \sum_x U(x) \text{ trace}(L_x W L_x W) = E(W) \tag{6.2.6}$$

where the scalar product $\langle \cdot, \cdot \rangle$ is defined via the trace as before. We have therefore succeeded in rewriting the Hamiltonian in matrix notation. For convenience, we will introduce the matrices $\mathbf{H_0}$ and $\mathbf{V}$ such that $\mathbf{H_0}W = TW + WT$ represents the hopping part of H and $\mathbf{V} = \sum_x U(x) L_x W L_x$ is the potential term. The reader should note that this definition is quite different from that in previous chapters, where $\mathbf{H_0}$ described the diagonal potential term and $\mathbf{V}$ represented the kinetic energy term (or its equivalent). We switch the meaning here not to confuse; as it will turn out, the sign of $\mathbf{V}$ (in the present notation) will again be the important ingredient of the Perron-Frobenius arguments.

We now try to find a cone in $\mathcal{V}$ such that the hopping part, as presented by the matrix $\mathbf{H_0}$, as well as the on-site energy (with a suitably chosen sign) leave it invariant. Our first choice is certainly the fat cone spanned by the p^2 state vectors $\phi_a \otimes \phi_b$ whose analogues we used so successfully in the last chapter. Unfortunately, since we did not fix the sign of the hopping matrix $t(x, y)$, we can not be sure that $TW + WT$ is again a nonnegative matrix. Even if we did fix the sign, the hopping term $c_x^\dagger c_y$ applied to some state again would invite all the dangers lurking within the violation of the ordering of the Fermi operators as discussed in section 5.2. In that section, we overcame the problem by retreating to the Jordan-Wigner transformation and restricting ourselves to nearest-neighbor hopping only. In addition, we had to reduce the dimension of our lattice to one, since generalizations of the Jordan-Wigner transformation to higher dimensions do not exist (there exists a two-dimensional version, but none for arbitrary dimensions). Thus, *the fat cone spanned by the p^2 state vectors $\phi_a \otimes \phi_b$ is not the cone of the Hubbard model.*

6.3 The No-Coupling Limit

In this section, we will restrict our attention to the hopping part $\mathbf{H_0}$.

Lemma (6.3.1) Let $\mathbf{H_0} : \mathcal{V} \to \mathcal{V}$ be given as above. Then the semigroup generated by $\mathbf{H_0}$ is

$$\exp(-t\mathbf{H_0})x = \exp(-tT)x \exp(-tT) \tag{6.3.2}$$

where $x \in \mathcal{V}$.

Proof (6.3.3) We will show by differentiation that the generator of the semigroup is indeed $\mathbf{H_0}$.

$$\begin{aligned}\frac{d}{dt}\exp(-t\mathbf{H_0})\Big|_{t=0} x &= \frac{d}{dt}\exp(-tT)\Big|_{t=0} x\exp(-tT) \\ &\quad + \exp(-tT)x\,\frac{d}{dt}\exp(-tT)\Big|_{t=0} \\ &= -(Tx + xT) \\ &= -\mathbf{H_0}x\end{aligned}$$

∎

In section 3.6.2 we studied the cone of operators $\Pi(\mathbf{PSD}) \in \mathcal{V}$ leaving invariant **PSD**. As showed in lemma 3.6.29, the semigroup generated by H_0 is an element of $\Pi(\mathbf{PSD})$. Therefore, we may apply the Krein-Rutman theorem 3.3.15 which states that there exists the Perron vector (which is thus a positive (semi)definite matrix) for the semigroup generated by $\mathbf{H_0}$.

Definition (6.3.4) The *Hilbert cone of the Hubbard model* is given as the full cone of positive semidefinite $p \times p$ matrices, i.e.

$$\mathcal{K}_{HBD} \equiv \mathbf{PSD} \tag{6.3.5}$$

This leads to the following results.

Theorem (6.3.6) Let $\mathbf{H_0}$ be the hopping part of the Hubbard model in the *no-coupling limit* acting on the space $\mathcal{V}$. Then

(a) Among the hermitian matrices in $\mathcal{V}$ that describe the ground states of $\mathbf{H_0}$, there is one that is a positive (semi)definite matrix W.

(b) There is some ground state corresponding to total spin $S = 0$.

(c) The ground state of $\mathbf{H_0}$ is not unique.

Proof (6.3.7)

to (a) We have shown that the semigroup generated by $\mathbf{H_0}$ is of the same form as the operator in lemma 3.6.29. Thus by the Krein-Rutman theorem 3.3.15, there exists a ground state matrix which is positive (semi)definite.

to (b) Let W be the positive (semi)definite ground state among the ground states of $\mathbf{H_0}$. Then we know that for some $a \in \Im[1,p]$ we have $W_{aa} > 0$ by theorem 3.6.3. Corresponding to this matrix element W_{aa} is the state $\Phi_a \equiv \phi_{a\uparrow} \otimes \phi_{a\downarrow}$ which satisfies $(\mathbf{S})^2\Phi_a = 0$. Therefore the ground state described by W has a nonzero projection into the space in which $S = 0$.

to (c) Let $\mathcal{P}_\lambda$ be the Hermitian projector into the eigenspace corresponding to the eigenvalue λ of the matrix T. Suppose that there are at least two distinct eigenvalues λ, μ– otherwise there is already degeneracy. Then the projectors P_λ and P_μ are eigenstates of $\mathbf{H_0}$ and therefore of its semigroup. Projectors are positive semidefinite such that by theorem 3.5.2, $\mathbf{H_0}$ is not $\mathcal{K}$-irreducible. In addition,

$$\begin{aligned}\langle P_\lambda, \exp(-t\mathbf{H_0})P_\mu\rangle &= \langle \exp(-tT)P_\lambda, \exp(-tT)P_\mu\rangle \\ &= \exp(-t\lambda)\exp(-t\mu)\langle P_\lambda, P_\mu\rangle \\ &= 0 \not> 0\end{aligned}$$

and thus the semigroup is not ergodic.

∎

6.4 The Attractive Case

We will now include the effects of the on-site Coulomb interaction for $U(x) \leq 0$ which, as indicated before, favors the forming of spin pairs at each lattice site.

First we show that the semigroup generated by $\mathbf{V}$ is positivity preserving w.r.t. $\mathcal{K}_{HBD}$.

Lemma (6.4.1) The on-site potential $\mathbf{V} : \mathcal{V} \to \mathcal{V}$ generates a semigroup that leaves invariant the cone $\mathcal{K}_{HBD}$.

Proof (6.4.2) We first note that L_x is a nonnegative matrix since its matrix elements $(L_x)_{ab}$ are either 0 or 1 depending whether there is an electron at site x in configuration ϕ_a and ϕ_b. Therefore by lemma 3.6.29, $L_x W L_x \in \mathcal{K}_{HBD}$ if $W \in \mathcal{K}_{HBD}$. But then, by the series expansion of the exponential, we have that

$$\begin{aligned}\langle W, \exp(-t\mathbf{V})V\rangle &= \langle W, V\rangle - t\sum_x U(x)\langle W, L_x V L_x\rangle \\ &\quad + \frac{t^2}{2}\sum_{xy} U(x)U(y)\langle W, L_y L_x V L_x L_y\rangle - \ldots \qquad (6.4.3) \\ &\geq 0 \qquad (6.4.4)\end{aligned}$$

for all $t \geq 0, W, V \in \mathcal{K}_{HBD}$, and thus the semigroup is positivity preserving, since $U(x) \leq 0$ for all x by hypothesis. ∎

So far we have shown that $\mathbf{H_0}$ and $\mathbf{V}$ do both generate positivity preserving semigroups. In order to prove uniqueness, we now need ergodicity of at least one of these operators.

Lemma (6.4.5) Assume that the hopping matrix $t(x,y)$ is connected. Then the matrix $T_{ab} = \langle b| \sum_{x,y\in\Lambda} t(x,y) c_x^\dagger c_y |a\rangle$ is connected, too.

Proof (6.4.6) We want to show that for all multi-indices $a, b \in \Im[1,p]$ there exists a sequence of multi-indices $a = c_1, c_2, \ldots, c_n = b$ such that

$$T_{ac_2} \cdot T_{c_2c_3} \cdot \ldots \cdot T_{c_{n-1}b} \neq 0 \tag{6.4.7}$$

(where the product is ordinary, not a matrix product). Thus let a be any multi-index labeling a pure configuration of electrons such that there is some $x_0 \in a$ and some $y_0 \notin a$ such that $t(x_0, y_0) \neq 0$. We note that this pair of sites x_0, y_0 exists since $t(x,y)$ is connected by hypothesis. Then $c_2 = a\backslash x_0 + y_0$ denotes a configuration identical to a except there is no electron at site x_0 but one at y_0. Thus $T_{ac_2} \neq 0$ and by the symmetry of $t(x,y)$ the same holds for T_{c_2a}. Again, we can construct a configuration c_3 such that it differs only in two sites from c_2 and $T_{c_2c_3} \neq 0$. Finally, since any lattice site x may be reached by repeated hopping $-t(x,y)$ being connected–we can reach any configuration $b \in \Im[1,p]$. Therefore T is connected. ∎

Lemma (6.4.8) The matrix $\mathbf{H_0}$ on $\mathcal{V}$, $W \mapsto TW + WT$ is connected, if T is.

Proof (6.4.9) Before we start the proof in earnest, we must dwell a bit more on the structure of $\mathcal{V}$. So far we have only said that we consider the elements of $\mathcal{V}$ as vectors and the operators on $\mathcal{V}$, such as $\mathbf{H_0}$, as matrices. But in fact these 'vectors' are $p \times p$ matrices and the 'matrices' are $p^2 \times p^2$. Therefore, in order to define how to apply such a matrix to a vector, we have to say explicitly which entry is going to be multiplied in what way. For example, we could order the $p \times p$ matrices in one column such that the first row of the original matrix constitutes the first p elements of the vector, and so on. Alternatively, we could choose the first column to do so. Any such ordering may be selected if we only keep in mind that we will have to arrange the $p^2 \times p^2$ matrices accordingly.

For convenience we will choose a less elaborate way here. We define

$$(\mathbf{H_0}W)_{ab} = \sum_{cd} \mathbf{H_0}_{abcd} W_{cd} \tag{6.4.10}$$

such that instead of one summation index, we need two. Using the explicit form of $\mathbf{H_0}$ we then have

$$\begin{aligned}(\mathbf{H_0}W)_{ab} &= \sum_{cd} \mathbf{H_0}_{abcd} W_{cd} \\ &= (TW + WT)_{ab} \\ &= \sum_{cd} (T_{ac}W_{cb} + W_{ad}T_{db}) \\ &= \sum_{cd} (T_{ac}W_{cd}\delta_{db} + \delta_{ac}W_{cd}T_{db}) \\ &= \sum_{cd} (T_{ac}\delta_{db} + \delta_{ac}T_{db})W_{cd} \end{aligned} \tag{6.4.11}$$

and therefore the matrix elements of $\mathbf{H_0}$ are given as $\mathbf{H_0}_{abcd} = (T_{ac}\delta_{db} + \delta_{ac}T_{db})$.

A configuration of electrons on Λ is thus described by the multiindex pair $\{a, b\}$, where a is for spin up and b is for spin down particles. Connectedness of $\mathbf{H_0}$ then means that starting at configuration $\{a, a'\}$, there exists a path of different configurations $\{c_i, c_i'\}$ such that we may reach the configuration $\{b, b'\}$, i.e.

$$\mathbf{H_0}_{aa'c_2c_2'} \cdot \mathbf{H_0}_{c_2c_2'c_3c_3'} \cdot \mathbf{H_0}_{c_3c_3'c_4c_4'} \cdot \ldots \cdot \mathbf{H_0}_{c_{n-1}c_{n-1}'bb'} \neq 0 \tag{6.4.12}$$

Intuitively, such a sequence of multiindex pairs is easy to construct if we keep in mind that T is connected. Thus we first connect $\mathbf{H_0}$ w.r.t. the unprimed spin down index, i.e. $\{a, a'\}$ connects to $\{b, a'\}$. Subsequently we connect the primed spin up indices, i.e. $\{b, a'\}$ connects to $\{b, b'\}$. Explicitly, we write for the unprimed sequence

$$\mathbf{H_0}_{aa'c_2a'} \cdot \mathbf{H_0}_{c_2a'c_3a'} \cdot \mathbf{H_0}_{c_3a'c_4a'} \cdot \ldots \cdot \mathbf{H_0}_{c_{l-1}a'ba'}$$

where we assume that l is the length of the path. In terms of equation (6.4.11), this gives

$$(T_{ac_2}\delta_{a'a'} + \delta_{ac_2}T_{a'a'}) \cdot (T_{c_2c_3}\delta_{a'a'} + \delta_{c_2c_3}T_{a'a'}) \cdot \ldots \cdot (T_{c_{l-1}b}\delta_{b'a'} + \delta_{c_{l-1}b}T_{a'a'})$$

$$\begin{aligned} &= T_{ac_2} \cdot T_{c_2c_3} \cdot \ldots \cdot T_{c_{l-1}b}\delta_{a'a'a'\ldots a'a'} \\ &\quad + T_{ac_2} \cdot T_{a'a'} \cdot T_{c_3c_4} \cdot \ldots \cdot T_{c_{l-1}b}\delta_{a'a'}\delta_{c_2c_3}\delta_{a'a'} \ldots \delta_{a'a'} \\ &\quad + \ldots + T_{a'a'} \cdot T_{a'a'} \cdot \ldots \cdot T_{a'a'}\delta_{ac_2c_3\ldots c_{l-1}b} \\ &= T_{ac_2} \cdot T_{c_2c_3} \cdot \ldots \cdot T_{c_{l-1}b} \\ &\neq 0 \end{aligned}$$

where $\delta_{abcd\ldots} = \delta_{ab}\delta_{bc}\delta_{cd}\ldots$ and the sequence $a, c_2, c_3, \ldots, c_{l-1}, b$ consists of distinct configurations. We now assume that $a', c_2', c_3', \ldots, c_{k-1}', b'$ connects the primed spin down configurations. Then in analogy

$$\mathbf{H_0}_{ba'bc_2'} \cdot \mathbf{H_0}_{bc_2'bc_3'} \cdot \mathbf{H_0}_{bc_3'bc_4'} \cdot \ldots \cdot \mathbf{H_0}_{bc_{k-1}'bb'}$$

$$\begin{aligned} &= T_{bc_2'} \cdot T_{c_2'c_3'} \cdot \ldots \cdot T_{c_{k-1}'b'} \\ &\neq 0 \end{aligned}$$

Thus

$$\left(\mathbf{H_0}_{aa'c_2a'} \cdot \mathbf{H_0}_{c_2a'c_3a'} \cdot \ldots \cdot \mathbf{H_0}_{c_{l-1}a'ba'}\right) \cdot \left(\mathbf{H_0}_{ba'bc_2'} \cdot \mathbf{H_0}_{bc_2'bc_3'} \cdot \ldots \cdot \mathbf{H_0}_{bc_{k-1}'bb'}\right)$$

$$\begin{aligned} &= \left(T_{ac_2} \cdot T_{c_2c_3} \cdot \ldots \cdot T_{c_{l-1}b}\right) \cdot \left(T_{bc_2'} \cdot T_{c_2'c_3'} \cdot \ldots \cdot T_{c_{k-1}'b'}\right) \\ &\neq 0 \end{aligned} \tag{6.4.13}$$

Therefore, $\mathbf{H_0}$ is connected by a path of length $n = l + k$. ∎

As for the Lieb-Mattis or the Heisenberg model, we have extracted a connected matrix that generates a positivity preserving semigroup. Using the connectedness, we now wish to show that this semigroup is also ergodic as we showed in theorem 5.3.14. Unfortunately, we are not allowed to do so. The cone $\mathcal{K}_{HBD}$ we are dealing with here is *not a fat cone* and thus the arguments presented in the theorem do not apply here.

Therefore, in order to assure positivity of the ground state– which would then again yield uniqueness–we have to adopt an alternate strategy.

Theorem (6.4.14) Let $\mathbf{H}$ be as given above and assume that $U(x) < 0$ for all $x \in \Lambda$ and $t(x, y)$ connected. Then if W is an eigenstate of $\mathbf{H}$, so is $|W|$, which is either positive or negative (semi)definite.

Proof (6.4.15) The proof uses the fact that instead of viewing W as a vector in $\mathcal{V}$, we can view it as a $p \times p$ matrix acting on $\mathbb{C}^p$. Let $Q = 2W_+ = |W| - W$, which is positive semidefinite as we showed in theorem 3.6.19. By linearity it is an eigenstate to the same eigenvalue $E(W)$, also. Denote the kernel of Q as $\ker Q = \{r \in \mathbb{C}^p | Qr = 0\}$

Then given some $r \in \ker Q$, we compute the expectation value of the Schrödinger equation (6.2.6)

$$(r, TQr) + (r, QTr) + \sum_x U(x)(r, L_xQL_xr) = E(W)(r, Qr) \tag{6.4.16}$$

The first two terms on the left hand side as well as the right hand side are zero by definition. Thus

$$\sum_x U(x)(r, L_xQL_xr) = 0 \tag{6.4.17}$$

By the hypothesis $U(x) < 0$ for all $x \in \Lambda$, and since Q is positive semidefinite and thus $(r, L_xQL_xr) \geq 0$, we conclude that $(r, L_xQL_xr) = 0$. Again by the positive

semidefiniteness of Q, this equation only holds if $QL_x r = 0$. Therefore L_x leaves $\ker Q$ invariant.

We again apply the Schrödinger equation, but this time without taking expectation values, i.e.

$$\underbrace{TQr}_{0} + QTr + \underbrace{\sum_x U(x) L_x Q L_x r - E(W) Q r}_{0} = 0 \tag{6.4.18}$$

Thus T maps $\ker Q$ into $\ker Q$, too.

Define

$$L_a \equiv \prod_{x \in a} L_x \tag{6.4.19}$$

which projects onto the basis vector e_a in C^p. Here e_a has zero entries in each of the p components except at component a, that is $(e_a)_i = \delta_{ai}$, for all $i \in \Im[1,p]$. Therefore $(L_a)_{cd} = \delta_{ac}\delta_{ad}$. By its construction, each L_a leaves $\ker Q$ invariant. In addition, $\sum_a^p L_a = \mathbb{1}$.

Assume now that $\ker Q \neq \{0\}$. Then there exists some $r \neq 0$ in $\ker Q$ and $L_a r = l_a e_a \neq 0$ for some configuration a and l_a some constant. Then we have

$$\begin{aligned}
&(L_b T L_{c_2} T \cdots T L_{c_{n-1}} T L_a r)_{ij} \\
&= \sum_{i_2,\ldots,i_{n-1}} (L_b)_{ii_2} T_{i_2 i_3} (L_{c_2})_{i_3 i_4} T_{i_4 i_5} \cdots T_{i_{n-2} i_{n-1}} (L_a)_{i_{n-1} j} r_j \\
&= T_{bc_2} \cdot T_{c_2 c_3} \cdot \ldots \cdot T_{c_{n-1} a} l_a e_a
\end{aligned}$$

Since T is a connected matrix, such a nonzero vector can be constructed for any $b \in \Im[1,p]$. Thus every basis vector of C^p is in $\ker Q$ and $\ker Q \equiv \mathsf{C}^p$ if it is not empty.

Therefore we have shown that either $\ker Q = \{0\}$ and thus $W = -|W|$, or $\ker Q = \mathsf{C}^p$ which implies that $W = |W|$. ∎

Remark (6.4.20) We emphasize that the above theorem is an adaptation of the Faris theorem 3.4.12. By proving that $|W|$ lies either in $\mathcal{K}$ or $-\mathcal{K}$, we have developed our arguments to be equivalent to the situation following equation (3.4.17).

Theorem (6.4.21) The ground state of the Hubbard model on a lattice Λ of arbitrary dimension for an even number of electrons has total spin $S = 0$. If, in addition, the hopping matrix $t(x,y)$ is connected and $U(x) < 0$ for all $x \in \Lambda$, then the ground state is unique.

Proof (6.4.22) By lemma 6.3.1 and lemma 6.4.1, we know that the semigroups generated by $\mathbf{H_0}$ and $\mathbf{V}$ are positivity preserving. Products of positivity preserving

matrices are again positivity preserving and so is $\exp(-t\mathbf{H})$ by the Trotter product formula. Thus we may again conclude that among the ground states there is one belonging to $\mathcal{K}_{HBD}$. As in theorem 6.3.6 this again implies that some ground state always has $S = 0$.

By theorem 6.4.14 we then know that if $U(x) < 0$ and $t(x, y)$ is connected, any ground state is either in $\mathcal{K}_{HBD}$ or $-\mathcal{K}_{HBD}$. Now suppose that there are two ground states W, W'. By the linearity of the Schrödinger equation, the ray $W + tW'$ for $t \in \mathsf{R}$ is a ground state, too. But by corollary 3.1.17, W and W' are multiples of each other and thus there is only a unique state. ∎

6.5 The Repulsive Case

In the attractive case, we were in the fortunate situation that the sign of the Coulomb interaction, as our essential tool, enabled a direct application of the Perron-Frobenius arguments. For the repulsive case $U(x) \geq 0$, as for the antiferromagnetic Heisenberg model, we need to employ a transformation of the Hubbard model that switches the sign of $U(x)$. A simple BCT transformation will certainly not suffice, since we already showed that this only changes the hopping term but leaves invariant the number operators. In order to see what we would want a transformation to do, let $\mathbf{U}$ be some unitary transformation matrix. Then the transformed Schrödinger equation reads

$$\begin{aligned} \mathbf{U}\mathbf{H}\mathbf{U}^{-1}\mathbf{U}W &= \mathbf{U}T\mathbf{U}^{-1}\mathbf{U}W + \mathbf{U}WT + \sum_x U(x)\mathbf{U}L_x\mathbf{U}^{-1}\mathbf{U}WL_x \\ &= T'W' + W'T + \sum_x U(x)L'_xW'L_x \end{aligned} \tag{6.5.1}$$

and therefore the conditions on the transformation $\mathbf{U}$, for leaving invariant the Hamiltonian $\mathbf{H}$, are

$$T = \mathbf{U}T\mathbf{U}^{-1} \tag{6.5.2}$$

$$L_x = -\mathbf{U}L_x\mathbf{U}^{-1} \tag{6.5.3}$$

But since $\sigma(L_x) \in \mathsf{R}^+$, the last equation can never be satisfied by a unitary matrix.

There are two ways out of this dilemma. The first one deals with a restriction of the full Hamiltonian (6.1.1) to a special, but nonetheless interesting, case, and will be covered in this section. The second possible way to circumvent the above problem is to add a one-body potential, which we will do in the next section.

The transformation we will use here is the so-called *hole-particle transformation* which we introduce in appendix A.4.3,

$$c_{x\downarrow} = d_{x\downarrow} \tag{6.5.4}$$

$$c_{x\uparrow} = \begin{cases} +d^\dagger_{x\uparrow} & \text{if } x \in \Lambda_A \\ -d^\dagger_{x\uparrow} & \text{if } x \in \Lambda_B \end{cases} \tag{6.5.5}$$

and we assume Λ to be bipartite. The net effect of this transformation–the $d_{x\sigma}$'s are again Fermi operators–is that by changing from spin up particles to spin up holes and leaving invariant spin down particles, it leaves invariant the hopping part of the Hubbard Hamiltonian. but changes the on-site potential, i.e.

$$T' = \mathbf{U}T\mathbf{U}^{-1} = T \tag{6.5.6}$$

$$L'_x = \mathbf{U}L_x\mathbf{U}^{-1} = \mathbb{1} - L_x \tag{6.5.7}$$

The Hamiltonian is then given by

$$\tilde{H} = \sum_{x,y\in\Lambda\sigma} t(x,y)d^\dagger_{x\sigma}d_{y\sigma} - \sum_{x\in\Lambda} U(x)\tilde{n}_{x\uparrow}\tilde{n}_{x\downarrow} + \sum_{x\in\Lambda} U(x)\tilde{n}_{x\downarrow} \tag{6.5.8}$$

where the number operators are now defined via the $d_{x\sigma}$'s. The corresponding matrix form of the equation should be clear by comparison with equation (6.2.5). We have therefore succeeded in flipping the sign of $U(x)$ at the cost of adding an additional term to the Hamiltonian. This term, in matrix notation $\sum_{x\in\Lambda} U(x)WL_x$, is not necessarily positivity preserving by lemma 3.6.27, if it is applied to some state W.

If we now assume that the on-site interaction parameter $U(x)$ is a constant, independent of the lattice site, then the additional term is $U\sum_{x\in\Lambda} n_{x\downarrow}$ which is a constant, too, since the number of spin down particles is conserved. This constant merely shifts the ground state energy but does not effect any of the dynamics. We may therefore drop it for the moment and consider

$$\tilde{H} = \sum_{x,y\in\Lambda\sigma} t(x,y)d^\dagger_{x\sigma}d_{y\sigma} - \sum_{x\in\Lambda} U(x)\tilde{n}_{x\uparrow}\tilde{n}_{x\downarrow} \tag{6.5.9}$$

only. This Hamiltonian is the same as the Hamiltonian in the attractive case. Thus the arguments leading to theorem 6.4.21 fully apply and we may conclude that the ground state ψ of the repulsive model is unique, too, and belongs to total spin $\tilde{S} = 0$ for connected $t(x,y)$ and $U > 0$.

Since we are not that much interested in the $\tilde{S}$ value but rather in the value of the spin operator S, we now return to the old variables. We want to show that the unique ground state $\psi(U)$ of the Hamiltonian (6.1.1) has spin value $S = (|\Lambda_B| - |\Lambda_A|)/2$, if w.l.o.g. we assume that $|\Lambda_B| \geq |\Lambda_A|$. The idea is due to E. Lieb and works for the half-filled band only. Thus let $N = |\Lambda|$ be even. Since $\psi(U)$ is unique *for all values of* U, this implies that S is independent of U — otherwise there is degeneracy for some $U > 0$. Now, as stated on page 68, the Hubbard model for

large U reduces to the antiferromagnetic Heisenberg model, for which we showed in section 5.3 that H_{HBG} has a unique ground state and $S = (|\Lambda_B| - |\Lambda_A|)/2$. For the special case of finite lattices, the uniqueness implies a gap in the energy spectrum of H_{HBG} and therefore the spin value of H_{HBD} is identical to the spin value of the Heisenberg ground state for some large enough U. We summarize

Theorem (6.5.10) The ground state of the repulsive Hubbard model on a bipartite lattice Λ of arbitrary dimension and an even number of sites is unique (up to the $2(S+1)$ degeneracy) for the half-filled band and belongs to $S = (|\Lambda_B| - |\Lambda_A|)/2$ if the hopping matrix $t(x, y)$ is connected and $U(x) = U > 0$.

Remark (6.5.11) There are three things worth considering before we leave this section. First, it may seem unfortunate that we have to restrict ourselves to half-filling. The contrary is true. For suppose we consider high-T_c superconductors (see the next chapter for more on this subject). It is now believed that superconductivity in the cooperoxides, say La_2CuO_4, takes place in the CuO planes. The Hubbard model, choosing appropriate values for t and U, has been proposed as a good starting point for modeling the behavior of these planes. The lattice is in fact bipartite w.r.t. the Cu and the O sites. Now D. Mattis [M 88] has shown that for electron occupation numbers below half-filling, the model features an antiferromagnetic state, and above half-filling, it exhibits ferromagnetism. At half-filling, the spin value equals $S = \frac{1}{2}$ in each unit cell. Thus, if the Hubbard model is a good model for these materials, theorem 6.5.10 should hold for the half-filled band only[7].

Second, we note that if $U = 0$, theorem 6.4.21 says that the total spin S equals zero and theorem 6.5.10 says that $S = (|\Lambda_B| - |\Lambda_A|)/2$. This puzzle may be solved if we observe that by the bipartiteness of the hopping matrix, rank $t(x, y) \leq 2|\Lambda_A|$ as we show in lemma A.1.20. Thus there are at most $2|\Lambda_A|$ nonzero eigenvalues and half of them are negative. Thus we put spin pairs in these negative states and fill the rest $|\Lambda| - 2|\Lambda_A| = |\Lambda_B| - |\Lambda_A|$ with, say, spin up electrons. Thus among the ground states is one with $S = |\Lambda_B| - |\Lambda_A|$ as well as one with $S = 0$.

Lastly, the one-dimensional Hubbard model is a special case of the Lieb-Mattis model of chapter 5. There we showed that the ground state belongs to $S = 0$ independent of the sign of the interaction term. This is compatible with our result for the attractive case $U \leq 0$. For the repulsive case, we note that necessarily $|\Lambda_A| = |\Lambda_B|$ in one dimension for a bipartite and connected nearest-neighbor chain. Thus $S = 0$ again and the results of this section agree with the Lieb-Mattis theorem 5.2.14.

[7]We thank E. Lieb for pointing out this fact to us.

6.6 Adding a One-Body Potential

As indicated in the last section, we may add a one-body potential to the Hamiltonian (6.1.1) such that in matrix notation, we get the Schrödinger equation

$$\mathbf{H}W = TW + WT + \sum_{x\in\Lambda} U(x)(L_x - \frac{1}{2})W(L_x - \frac{1}{2}) \tag{6.6.1}$$

Again, we perform the hole-particle transformation of above. This time the on-site term transforms as

$$\mathbf{U}(L_x - \frac{1}{2})\mathbf{U}^{-1}W'(L_x - \frac{1}{2}) = (1 - (L_x - \frac{1}{2}))W'(L_x - \frac{1}{2}) = -(L_x - \frac{1}{2})W'(L_x - \frac{1}{2}) \tag{6.6.2}$$

Thus we have flipped the sign of $U(x)$ but retained the form of the on-site term.

Before we may now apply the arguments leading to theorem 6.4.21, we have to show that the on-site potential is again positivity preserving and generates an ergodic semigroup. The first assertion is clearly true by lemma 6.4.1; the second assertion requires a careful review of the proof of theorem 6.4.14. Using the same notation as in proof 6.4.15, we see that the Schrödinger equation yields

$$(r, TQr) + (r, QTr) + \sum_x U(x)(r, (L_x - \frac{1}{2})Q(L_x - \frac{1}{2}r) = E(W)(r, Qr) \tag{6.6.3}$$

which implies that for $U(x) > 0$

$$(r, (L_x - \frac{1}{2})Q(L_x - \frac{1}{2})r) = 0 \tag{6.6.4}$$

for all x and thus $QL_xr = \frac{1}{2}Qr = 0$. Therefore we again have shown that L_x maps $\ker Q$ into $\ker Q$ and we may proceed to construct the projection operators L_a. The ground state belongs in addition to $S = 0$, by the same arguments as in the attractive case. Thus we have the following theorem:

Theorem (6.6.5) Consider the extended Hubbard model

$$H = \sum_{x,y\in\Lambda\sigma} t(x,y)c^\dagger_{x\sigma}c_{y\sigma} + \sum_{x\in\Lambda} U(x)(n_{x\uparrow} - \frac{1}{2})(n_{x\downarrow} - \frac{1}{2}) \tag{6.6.6}$$

on a lattice Λ of arbitrary dimension. Assume that the hopping matrix $t(x,y)$ is connected and bipartite. Assume further that $U(x) > 0$ but not necessarily constant. Then the ground state ψ is unique and the spin value corresponds to the $S = 0$.

Chapter 7

The tJ-Model

7.1 Introduction

The discovery of the novel high-temperature superconductivity[1] has given rise to searches for new mechanisms of superconductivity. In the standard BCS theory[2], superconductivity results from the instability of the ground state originating from effectively attractive electron-phonon interactions. Since high-T_c superconductivity seems to be created by doping an insulating state, P. W. Anderson suggested the possibility of superconductivity near a Mott transition in a model of strong repulsive on-site Coulomb interactions [A87].

Using this as a starting point, several authors have derived [S 75, ZR 88] and studied [S 87, W 88, BB 90] the *tJ-model*, which, as indicated in the preceding chapter, is a particular case of the Hubbard model. The main features of this model may be described as exhibiting

- hard-core repulsion ($\mathcal{P}$),
- nearest-neighbor hopping (t) and
- spin exchange (J)

Therefore the tJ-model may be seen as a merger of the (no-coupling) Hubbard model and the Heisenberg model.

7.2 Uniqueness of the Ground State

Consider a one-dimensional lattice Λ of N electrons, where each lattice site is capable of accommodating at most one electron; necessarily $N \leq |\Lambda|$. The dynamics

[1] J. G. Bednorz, K. A. Müller, Z. Phys. B **64**, 189, (1986)
[2] J. Bardeen, L. N. Cooper, J. R. Schrieffer, Phys. Rev. **108**, 1175, (1957)

of the tJ-model in the form of [BB 90] is given by

$$H = \mathcal{P} \sum_{\substack{(x,y) \\ \sigma=\uparrow,\downarrow}} t(x,y)\, c^{\dagger}_{x\sigma} c_{y\sigma}\, \mathcal{P} + \sum_{(x,y)} J(x,y) \left\{ \mathbf{S}(x) \cdot \mathbf{S}(y) - \frac{n_x n_y}{4} \right\}, \qquad (7.2.1)$$

where $\mathcal{P} = \prod_{x\in\Lambda}(1 - n_{x\uparrow} n_{x\downarrow})$ restricts the Hilbert space by the constraint of no double occupancy and (x,y) denotes nearest-neighbor pairs. The Hamiltonian is $SU(2)$ invariant [BB 90], i.e. the total spin S is a conserved quantity. The spin operators are given as on page 52. The hopping matrix $t(x,y)$ and the exchange matrix $J(x,y)$ are assumed to be real and symmetric.

Rewriting the Hamiltonian (7.2.1) in terms of creation, annihilation and number operators we get

$$\begin{aligned} H &= H_0 - V \\ &= -\frac{1}{2} \sum_{\substack{(x,y) \\ \sigma=\uparrow,\downarrow}} J(x,y)\, n_{x\sigma} n_{y-\sigma} \\ &\quad + \sum_{\substack{(x,y) \\ \sigma=\uparrow,\downarrow}} \left\{ t(x,y) \mathcal{P} c^{\dagger}_{x\sigma} c_{y\sigma} \mathcal{P} + J(x,y) c^{\dagger}_{x\sigma} c_{x-\sigma} c^{\dagger}_{y-\sigma} c_{y\sigma} \right\}. \qquad (7.2.2) \end{aligned}$$

In this form, it is easy to see that the spin-exchange part of the Hamiltonian describes double-hops as mentioned in chapter 5. The number of spin up and spin down particles is individually conserved[3].

By $SU(2)$ invariance, all states of allowed spin angular momentum can be rotated into the $M = 0$ subspace with no change in their energy. The ground state here is therefore the ground state of the Hamiltonian. The configurations which form a complete set in this subspace have $n = N/2$ electrons (N even) with spin up and n electrons with spin down. There are

$$p = \begin{pmatrix} |\Lambda| \\ n \end{pmatrix} \cdot \begin{pmatrix} |\Lambda| - n \\ n \end{pmatrix} \qquad (7.2.3)$$

distinct configurations, which we will denote by ϕ_a, $a \in \Im[1,p]$. We now choose the phase of the ϕ_a's such that

$$\phi_a = C \cdot c^{\dagger}_{x_1\sigma_1} c^{\dagger}_{x_2\sigma_2} \cdots c^{\dagger}_{x_N\sigma_N} |0\rangle, \qquad (7.2.4)$$

where $|0\rangle$ denotes the no-particle state, C is a positive normalization constant and the spin on each site is either up or down. Let $\mathcal{K}_{tJ}$ be the set of all linear

[3]Denoting an empty lattice site a *hole*, we may therefore view the tJ-model as a three-component quantum system [S 75]; species A are the spin up, species B the spin down particles — obeying Fermi statistics — and species C are the holes — obeying Bose statistic.

combinations of the ϕ_a with *nonnegative* superposition coefficients (the fat cone). Then the relations

$$c^\dagger_{x\sigma} c_{x+1\tau}\, \mathcal{K}_{tJ} \subset \mathcal{K}_{tJ}, \tag{7.2.5}$$
$$c^\dagger_{x+1\sigma} c_{x\tau}\, \mathcal{K}_{tJ} \subset \mathcal{K}_{tJ}, \tag{7.2.6}$$
$$c^\dagger_{x\sigma} c_{x-\sigma}\, \mathcal{K}_{tJ} \subset \mathcal{K}_{tJ}, \tag{7.2.7}$$

hold for all lattice sites x and all possible spin directions σ, τ. This can be easily seen from the anticommutation relations (here a crucial fact is the assumed nearest-neighbor coupling). We will show one of the above equations, namely equation (7.2.7). Since

$$[c^\dagger_{x\sigma} c_{y\tau}, c^\dagger_{z\rho}] = c^\dagger_{x\sigma} \delta_{yz} \delta_{\tau\rho}, \tag{7.2.8}$$

we can shift the pair $c^\dagger_{x\sigma} c_{x-\sigma}$ until we reach the xth site without changing the sign.

$$\begin{aligned}
& c^\dagger_{x\tau} c_{x-\tau}\; c^\dagger_{x_1\sigma_1} c^\dagger_{x_2\sigma_2} \cdots c^\dagger_{x_N\sigma_N} |0\rangle \\
&= c^\dagger_{x_1\sigma_1} \cdots c^\dagger_{x\tau} c_{x-\tau} c^\dagger_{x\sigma} \cdots c^\dagger_{x_N\sigma_N} |0\rangle \\
&= c^\dagger_{x_1\sigma_1} \cdots c^\dagger_{x\tau} \left[-c^\dagger_{x\sigma} c_{x-\tau} + \{c_{x-\tau}, c^\dagger_{x\sigma}\} \right] \cdots c^\dagger_{x_N\sigma_N} |0\rangle \\
&= \begin{cases} 0 & n_x = 0 \\ 0 & n_{x\sigma} = 1, \sigma = \tau \\ c^\dagger_{x_1\sigma_1} \cdots c^\dagger_{x\tau} \cdots c^\dagger_{x_N\sigma_N} |0\rangle & n_{x\sigma} = 1, \sigma = -\tau \end{cases}
\end{aligned} \tag{7.2.9}$$

The remaining relations (7.2.5) and (7.2.6) are proved in an analogous way. We have thus shown that the operator V is positivity preserving w.r.t. $\mathcal{K}_{tJ}$ if we neglect the sign of the hopping and the exchange matrix for the moment.

By the previously considered BCT transformation, the sign of the hopping matrix may again be flipped under the assumption of bipartiteness. Thus for the cases $t(x,y) \geq 0$, $J(x,y) \leq 0$ and $t(x,y) \leq 0$, $J(x,y) \leq 0$, we may again apply the Perron-Frobenius arguments as we did in the case of the Lieb-Mattis and the Heisenberg models. The connectedness insures the ergodicity of the corresponding semigroup and the H_0 operator is certainly diagonal in the chosen basis. Therefore the ground state is unique up to spin degeneracy.

Let $J(x,y) \geq 0$. Then we have to find a transformation which changes the sign of $J(x,y)$ so that we may again apply the Perron-Frobenius theorems. We first observe that

$$\begin{aligned}
c^\dagger_{x\sigma} c_{x-\sigma} c^\dagger_{y-\sigma} c_{y\sigma} &= c^\dagger_{x\sigma} c_{y\sigma} c_{x-\sigma} c^\dagger_{y-\sigma} \\
&= c^\dagger_{x\sigma} c_{y\sigma} \left[\delta_{xy} - c^\dagger_{y-\sigma} c_{x-\sigma} \right] \\
&= \begin{cases} -c^\dagger_{x\sigma} c_{y\sigma} c^\dagger_{y-\sigma} c_{x-\sigma} & \text{if } x \neq y \\ n_{x\sigma} - n_{x\sigma} n_{x-\sigma} & \text{if } x = y \end{cases}
\end{aligned} \tag{7.2.10}$$

Thus in the first case the double hops appear explicitly. In the second case ($x = y$), the double hops reduce to the number operator at each site, if we note that $n_{x\sigma}n_{x-\sigma} = 0$, since double occupancy is projected out by $\mathcal{P}$. Now we have run into a major problem, since if we denote the hopping term for spin up by u and the spin down hopping by d, we are looking for a transformation such that

$$u + d - u \cdot d \;\rightarrow\; u + d + u \cdot d \tag{7.2.11}$$

This means that an overall sign change as produced by the BCT transformation as well as a sign change in either up or down hopping operators as produced by the hole-particle transformation will not suffice.

The transformation that we will use here is the Jordan-Wigner transformation. We have already shown in chapter 5 that, restricting ourselves to one dimension and nearest-neighbor interactions only, the number as well as the hopping operators remain invariant under this transformation. The same holds now for the spin exchange part which, having been brought into the form of equation (7.2.10), is nothing but two hopping operators combined. If we now interchange the order of the operators again and bring them back into the standard form of the spin exchange, *then the sign does not change*, since the Paulions at different sites commute. Assuming $J(x,x) = 0$, we have therefore succeeded in flipping the sign of the $J(x,y)$ matrix.

$$c^\dagger_{x\sigma}c_{x-\sigma}c^\dagger_{y-\sigma}c_{y\sigma} = -c^\dagger_{x\sigma}c_{y\sigma}c_{y-\sigma}c_{x-\sigma} \;\rightarrow\; -b^\dagger_{x\sigma}b_{x-\sigma}b^\dagger_{y-\sigma}b_{y\sigma} = -b^\dagger_{x\sigma}b_{y\sigma}b_{y-\sigma}b_{x-\sigma} \tag{7.2.12}$$

We summarize these results in the next theorem.

Theorem (7.2.13) Let H_{tJ} denote the Hamiltonian of the one-dimensional nearest-neighbor tJ-model. Assume that the sign of the hopping matrix $t(x,y)$ is the same for all $x, y \in \Lambda$, and likewise for the sign of the spin exchange matrix $J(x,y)$. Then if both matrices are connected (and bipartite if necessary), the ground state of H_{tJ} is unique up to the $2S+1$ spin degeneracy.

7.3 The Spin Value of the Ground State

So far, the result is still unsatisfactory. If the tJ-model is indeed a model of high-T_c superconductors, it should exhibit an antiferromagnetic ground state for $N \approx |\Lambda|$ [ABCKS88]. One might be tempted [M 91] to argue that by the theorem 5.2.14 of Lieb-Mattis the antiferromagnetism of the ground state is implied by the one-dimensionality of the chain Λ. This is certainly not true since Lieb-Mattis [LM 62a] point out that the "theorem is not valid if there are explicitly spin-dependent forces" present. But we may again employ the mechanism which led to the spin

value for the repulsive Hubbard model. In a calculation by A. Mielke [M 91] in the same paper cited above, the author concluded that the tJ-model for $J > 0$ reduces to the antiferromagnetic $s = \frac{1}{2}$ Heisenberg chain if the spin-exchange term is treated as a perturbation. Thus, using the same arguments as for the Hubbard model, we may argue that

Conjecture (7.3.1) The unique ground state of the connected tJ-model for $J > 0$ has total spin $S = 0$ for nearest-neighbor hopping and an even number $N = |\Lambda|$ of electrons on the linear chain Λ.

A. Mielke does not restrict himself to the half-filled band in the perturbation expansion, which should therefore hold for all $N \leq |\Lambda|$. Again, this would indicate that the ground state spin value is zero for arbitrary filling. It is important to note that this result is in contrast to the behavior of real two-dimensional high-T_c superconductors [ABCKS88]. There the antiferromagnetic state changes to a spin-glass state as the number of holes is increased by doping. For an even larger number of holes such systems then exhibit a superconducting phase. The result also contrasts a recent calculation by A. Förster [F 91], who showed that on a one-dimensional lattice of $|\Lambda| = 10$ sites, the $S = 0$ state is a unique ground state for $N = 10$ only. For all other (even) values of N, the ground state is degenerate. Thus the question of the spin value of the ground state of the tJ-model apart from half-filling has to yet to be resolved.

Before we leave this chapter, we study the possible generalizations of our theorems to two or more dimensions which are decidedly much more interesting from an experimental point of view. In chapter 6 we point out that the tJ-model may be viewed as a special case of the Hubbard model; thus we may ask ourselves whether some of the arguments used there may be applied in the case of the tJ-model. Unfortunately, we cannot use the cone of the Hubbard model, $\mathcal{K}_{HBD}$, since by the hard-core repulsion of the tJ-model, the diagonal elements of the state matrices $W \in \mathcal{K}_{HBD}$ are zero and thus $\mathcal{K}_{HBD} = \emptyset$ by theorem 3.6.3.

Chapter 8

Conclusions

After having reviewed the classical Perron-Frobenius theory and introduced possible generalizations, we developed a semigroup version of the Perron-Frobenius theory following the lead of Faris [F 72]. The 'perturbation' results thus derived were used to prove the existence and, if applicable, the uniqueness of the ground state for the Lieb-Mattis model, the Heisenberg magnet and the Hubbard model on finite lattices. In addition, the tJ-model was successfully treated with the same methods.

A number of open problems remain. As the reader has probably noticed, we do not make explicit use of theorem 4.3.1 for the Hubbard model. Instead we state and prove theorem 6.4.14 without using any semigroup terminology. As indicated before, the reason for this is that the hopping term of the Hubbard model alone is not ergodic w.r.t. the cone $\mathcal{K}_{HBD}$. Thus the semigroup theorems, which use the Trotter product formula and therefore decompose the semigroup generated by H_{HBD} into products of two semigroups, are not applicable. For this reason we use the hopping *and* the on-site potential term in theorem 6.4.14 to establish ergodicity of the complete Hamiltonian H_{HBD}. A possible semigroup method for doing this might be to use the Baker-Cambell-Hausdorff-Dynkin (BCHD) formula [RS1 80]

$$\exp(A)\ \exp(B)\ = \exp(A + B + \frac{1}{2}[A, B] + \ldots)$$

If the higher commutators vanish and if in addition the semigroup generated by the commutator term is ergodic, we would have the desired form for the Perron-Frobenius arguments. Unfortunately, this is not true in the Hubbard model for the commutator of $[H_0, V]$. It is, however, certainly worthwhile to examine under what assumptions the commutator of two operators A, B is ergodic.

In applying the Perron-Frobenius arguments to the tJ-model, we had to restrict ourselves to the one-dimensional nearest neighbor version since we used the Jordan-Wigner transformation. A possible generalization to two or more dimensions may be pursued using at least two different methods. First, sticking to the

fat cone $\mathcal{K}_{tJ}$, we might apply generalizations of the Jordan-Wigner transformation to higher dimensions. Second, we might start to look for another cone. The first idea is certainly the most promising, since generalizations of the Jordan-Wigner transformation exist for at least two dimensions. With respect to the second idea, we note that, as indicated before, the cone $\mathcal{K}_{HBD}$ of the Hubbard model can not be used for the tJ-model.

The above problems are largely due to the fact that for fermionic systems, the ordering of the Fermi operators is extremely important. For the Lieb-Mattis and the tJ-model we were consequently obliged to use the Jordan-Wigner transformation, which restricted us to a linear chain. For the Hubbard model, we circumvented that problem by using a nonstandard cone, thus inviting all the problems mentioned in the above paragraph. Only for the Heisenberg model did these problems not arise, since there we were dealing with spin operators that already commute at different sites. Another most promising way out of this dilemma was presented by Gross [G 72] and Faris [F 72]. They developed a *noncommutative*, infinite-dimensional extension of the Perron-Frobenius theory by considering a suitable Clifford algebra $\mathcal{C}$ over the Hilbert space $\mathcal{W}$. Since then, not much work has been devoted to the application of C^* algebras to Perron-Frobenius type theories. It might be interesting to see whether any of the knowledge obtained from the highly successful application of C^* algebras in statistical mechanics and algebraic quantum field theory over the last 20 years might be used to further generalize and extend Perron-Frobenius theory. This in turn should yield new insights into ground state degeneracy for a larger class of Hamiltonians.

Acknowledgements A number of people have either directly or indirectly helped to bring this thesis into existence. First of all, Prof. R. Schrader suggested the original problem and had enough patience to answer my numerous questions. In addition, he and his colleagues Prof. B. Schroer and Dr. M. Karowski have provided a stimulating and lively working atmosphere. A big "Thank you!" goes to G. Haak, with whom collaboration is a pleasure. J. Mund knows about ropes *and* cones–his knowledge of both helped tremendously. I am especially indebted to Ms. A. Förster for providing me with all the background information of the tJ-model. Ms. Susanne Hauger (DUKE UNIVERSITY, Durham NC) did the linguistic checks which this thesis–written as it was by a nonnative speaker of the English tongue–desperately needed. Another "Vielen Dank!" to her.

Finally, I would like to thank all the members of the mathematical physics group at the FU Berlin for being such a healthy bunch of people.

Appendix A

Miscellaneous Results

A.1 A Short Excursion into Graph Theory

Definition (A.1.1) A matrix $A = [a_{ij}] \in M_n$ is said to have *property SC* if for every pair of distinct integers p, q with $1 \leq p, q \leq n$ there is a sequence of distinct integers $p = k_1, k_2, \ldots, k_m = q$, $1 \leq m \leq n$ such that

$$a_{k_1 k_2} \cdot a_{k_2 k_3} \cdot \ldots \cdot a_{k_{m-1} k_m} \neq 0 \tag{A.1.2}$$

(where this is an ordinary, not a matrix, product).

Definition (A.1.3) The *directed graph* of $A \in M_n$, denoted by $\Gamma(A)$, is the directed graph on $|\Gamma(A)| = n$ nodes $P_1, \ldots, P_n$ such that there is a *directed arc* in $\Gamma(A)$ from P_i to P_j iff $a_{ij} \neq 0$.

Definition (A.1.4) A *directed path* γ in a graph Γ is a sequence of arcs $P_{i_1} P_{i_2}$, $P_{i_2} P_{i_3}$, ... in Γ. The *ordered list of nodes* in the directed path γ is $P_{i_1}, P_{i_2}, \ldots$.

Definition (A.1.5) A directed graph Γ is *strongly connected* if between every pair of distinct nodes P_i, P_j in Γ there is a directed path of finite length that begins at P_i and ends at P_j.

Theorem (A.1.6) Let $A \in M_n$. Then A has property SC iff $\Gamma(A)$ is strongly connected.

Proof (A.1.7) A has property SC $\Leftrightarrow$ given any distinct pair $i, j \in \{1, \ldots, n\}$ there exists a sequence $\{k_1, \ldots, k_m\}$ such that $a_{ik_1} a_{k_1 k_2} \cdots a_{k_{m-1} j} \neq 0 \Leftrightarrow$ there exists a directed path γ between i and j of finite length $\Leftrightarrow$ $\Gamma(A)$ is strongly connected. ∎

Theorem (A.1.8) Let $A \in M_n$. Let P_i, P_j be given nodes of $\Gamma(A)$. There exists a directed path of length m in $\Gamma(A)$ from P_i to P_j iff $(|A|^m)_{ij} \neq 0$.

$$I(A) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} ; \quad \Gamma(A) =$$

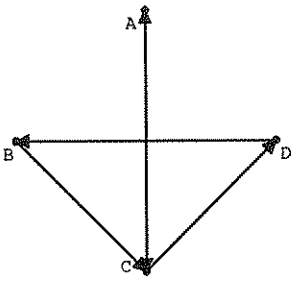

Figure A.1: The indicator matrix $I(A) \in M_4$ with property SC and its corresponding strongly connected graph $\Gamma(A)$. Note the cycle $P_A \to P_C \to P_D \to P_B \to P_C \to P_A$.

Proof by Induction (A.1.9)
$\diamond$: For $m = 1$ the theorem is clearly true. Now, let $m = 2$, then

$$\begin{aligned} (|A|^2)_{ij} &= \sum_k^n |A|_{ik}|A|_{kj} \\ &= \sum_k^n |a_{ik}||a_{kj}| > 0 \end{aligned} \tag{A.1.10}$$

iff at least one k_0 exists such that $|a_{ik_0}| \cdot |a_{k_0 j}| \neq 0$. This is equivalent to the existence of a path of length 2.
$\hookrightarrow$: In general,

$$\begin{aligned} (|A|^{m_0+1})_{ij} &= \sum_k^n (|A|^{m_0})_{ik}|A|_{kj} \\ &= \sum_k^n (|A|^{m_0})_{ik}|a_{kj}| \end{aligned} \tag{A.1.11}$$

iff at least for one k_0, $(|A|^{m_0})_{ik_0} \cdot |a_{k_0 k}| \neq 0$, and this is equivalent to a path of length $m_0 + 1$. ∎

Corollary (A.1.12) Let $A \in M_n$. Then A has property SC iff $(\mathbb{1} + |A|)^{n-1} > 0$ or, equivalently, if $(\mathbb{1} + I(A))^{n-1} > 0$.

Proof (A.1.13)

$$(\mathbb{1} + |A|)^{n-1} = \mathbb{1} + (n-1)|A| + \binom{n-1}{2}|A|^2 + \ldots + \binom{n-1}{n-2}|A|^{n-1} > 0$$

iff for each pair (i,j) of nodes with $i \neq j$ at least one of the terms $|A|, |A|^2, \ldots, |A|^{n-1}$ has a positive (i,j) entry. By the above theorem A.1.8, this is equivalent to the existence of a directed path in $\Gamma(A)$ from P_i to P_j which in turn is equivalent to $\Gamma(A)$ being strongly connected and A having property SC. ∎

Theorem (A.1.14) Let $A \in M_n(\mathsf{R}^+)$. The A has property SC iff A is irreducible.

Proof (A.1.15) We shall actually prove that A is reducible iff A^m has at least one 0 entry. Assume first that A is reducible. Let

$$P^T A P = \tilde{A} = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \tag{A.1.16}$$

where $B \in M_r$, $D \in M_{n-r}$, $C \in M_{r,n-r}$, and $0 \in M_{n-r,r}$ is a zero matrix as in definition 2.3.1. Now, $\tilde{A}^m$ has the same $0 \in M_{n-r,r}$ matrix in the lower left corner for all $m < \infty$. Therefore there is never a path between nodes P_i and P_j as long as $i \in \mathsf{N}[n-r,n]$, $j \in \mathsf{N}[1,r]$ or vice versa. Since a permutation only relabels the nodes, the argument holds for A^m, too. Thus by theorem A.1.8, A does not have property SC.

Conversely, suppose that A is not strongly connected. Then there is a pair (a,b) such that $(A^m)_{ab} \neq 0$ and no directed path from P_a to P_b exists. Let

$$N_1 \equiv \{P_i : P_i = P_b \text{ or there is a directed path from } P_i \text{ to } P_b\}$$

and $N_2 \equiv \Gamma(A) - N_1$. Then since $P_b \in N_1$, $N_2 \neq \{P_1, \ldots, P_n\}$. By construction, there cannot be a path from a node in N_2 to some node in N_1 since then this node would already belong to N_1. Relabeling the nodes such that $N_1 = \{\tilde{P}_1, \ldots, \tilde{P}_r\}$ and $N_2 = \{\tilde{P}_{r+1}, \ldots, \tilde{P}_n\}$ we notice that then we can write

$$\tilde{A} = P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \tag{A.1.17}$$

Thus A is reducible. ∎

Definition (A.1.18) A matrix $A = [a_{ij}] \in M_n$ is said to be *bipartite*, if for all $n \geq 2$ there is a permutation matrix $P \in M_n$ and some integer r, $1 \leq r \leq n-1$ such that

$$P^T A P = \begin{pmatrix} 0 & D \\ C & 0 \end{pmatrix} \tag{A.1.19}$$

where $D \in M_{r,n-r}, C \in M_{n-r,r}$ and 0 the remaining zero matrices. This is equivalent to saying that the vertices a_{ij} can be divided into two disjoint sets C, D such that $a_{ij} = 0 \; \forall \; i \in C, j \in D$ or $j \in D, i \in C$.

$$I(A) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}; \quad \Gamma(A) =$$

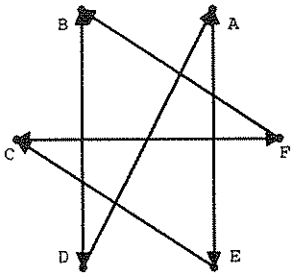

Figure A.2: The indicator matrix of a bipartite $A \in M_6$ and its graph $\Gamma(A)$.

Lemma (A.1.20) Let $A \in M_n$ be bipartite. Then, if C, D are given as in definition A.1.18, we have $\text{rank}A \leq 2\min(|\Gamma(C)|, |\Gamma(D)|)$.

Proof (A.1.21) Since A is bipartite, we write is as

$$A = \begin{pmatrix} 0 & D \\ C & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix} + \begin{pmatrix} 0 & D \\ 0 & 0 \end{pmatrix} = C' + D' \tag{A.1.22}$$

By the properties of the rank, we have that for any $A \in M_n$, $\text{rank}A \leq |\Gamma(A)|$. Now w.l.o.g. assume that $|\Gamma(C)| \leq |\Gamma(D)|$. Then

$$\begin{aligned} \text{rank}A &= \text{rank}(C' + D') \\ &\leq \text{rank}(C') + \text{rank}(D') \\ &\leq |\Gamma(C)| + |\Gamma(D)| \\ &\leq 2|\Gamma(C)| \end{aligned}$$

∎

A.2 Ergodicity and *Ergodicity*

In section 3.4 we define the concept of an ergodic operator A in a real Hilbert space. But the term "ergodic" is usually used in a different context: *Ergodic Theory* as it appears in the literature [RS1 80, CFS 82], attempts to answer the question of whether "time mean equals phase-space mean". Explicitly, let Γ be the phase space of a classical mechanical system and let a constant energy surface in Γ be denoted by Ω_E. Let $T_t : \Gamma \to \Gamma$ be a translation in time, that is $T_t x$ is the state at time t originating from the state $T_{t_0} x$ in Γ. Given an observable $f(\cdot)$ and a state $w \in \Omega_E$, we then want to insure that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(T_t w) dt \tag{A.2.1}$$

exists, is independent of w, and is equal to

$$\mu(f) = \int_{\Omega_E} f(w) d\hat{\mu}(w) \tag{A.2.2}$$

where $\hat{\mu}$ is a probability measure on Ω_E. This statement can be made in the case where T_t obeys the following definition, i.e.

Definition (A.2.3) [RS1 80] T_t is called *ergodic* if the constant functions are the only functions in $L^2(\Omega_E, d\mu_E)$ for which $f \circ T_t = f$ (as L^2 functions) for all t.

The equivalence of the two concepts is then shown by the next theorem. The cone in $L^2(M, \mu)$ is the set of a.e. nonnegative functions, as introduced on page 3.4.2.

Theorem (A.2.4) Let (M, μ) be a probability measure space, i.e. $\mu(M) = 1$. Let $T : M \to M$ be a measure preserving map, i.e. $\mu(T^{-1}F) = \mu(F)\ \forall\ F \subset M$. Let

$$(Af)(x) = f(Tx)$$

Then

(a) A is unitary,

(b) A is positivity preserving,

(c) A is never positivity improving,

(d) A ergodic in the sense of definition 3.4.8 is equivalent to T ergodic in the sense of definition A.2.3.

Proof (A.2.5)

to (a) Let $f \geq 0$. Then

$$\begin{aligned}
(Af, Ag)(x) &= \int_M f^*(Tx) g(Tx) d\mu(x) \\
&= \int_M f^*(x) g(x) d\mu(T^{-1}x) \\
&= \int_M f^*(x) g(x) d\mu(x) \\
&= (f, g)
\end{aligned}$$

since T is measure preserving. Therefore A is unitary.

to (b) $(Af)(x) = f(Tx) = f(y) \geq 0$ and thus A is positivity preserving.

to (c) $(Af)(x) = f(Tx) \geq 0$ but never > 0 for *all* $x \in M$ since $f \in L^2(M, \mu)$.

to (d) The assertion is an application of the Perron-Frobenius theorem 3.4.12. By the definition of A, we know that the constant functions are invariant, and since A is unitary, we have $(Af)(x) = f(Tx) = 1 \cdot f(x)$ for f constant. Assuming that A is ergodic as in definition 3.4.8, the Perron-Frobenius theorem states that the eigenspace is one-dimensional. The constant functions already lie in the eigenspace; they are the only functions that do so, since otherwise the eigenspace cannot be one-dimensional. The converse follows immediately by the same arguments, if we note that the Perron-Frobenius theory can be proved in the other direction; see [F 72] for details.

∎

A.3 An Operator Identity

The following lemma is needed in section 6.

Lemma (A.3.1) Given two operators A, B on a Hilbert space $\mathcal{H}$, the following formula holds for all $n \in \mathbb{N}$,

$$AB^n = \sum_{k=0}^{n-1} B^k [A, B] B^{n-1-k} + B^n A. \tag{A.3.2}$$

Proof by Induction (A.3.3)
$\diamond$: Suppose $n_0 = 1$. Then we have

$$\begin{aligned} AB^1 &= \sum_{k=0}^{0} B^k [A, B] B^{0-k} + B^1 A \\ &= [A, B] + BA \\ &= AB \end{aligned}$$

as expected.
$\hookrightarrow$: Then by induction we have for $n = n_0 + 1$

$$\begin{aligned} AB^n &= AB^{n_0} B \\ &= \left(\sum_{k=0}^{n_0-1} B^k [A, B] B^{n_0-1-k} + B^{n_0} A \right) B \\ &= \sum_{k=0}^{n_0-1} B^k [A, B] B^{n_0-k} + B^{n_0} AB \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n_0-1} B^k[A,B]B^{n_0-k} + B^{n_0}[A,B] + B^{n_0+1}A \\
&= \sum_{k=0}^{n_0} B^k[A,B]B^{n_0-k} + B^{n_0+1}A \\
&= \sum_{k=0}^{n-1} B^k[A,B]B^{n-1-k} + B^nA
\end{aligned}
$$

completing the induction. ∎

A.4 Some Transformations

A crucial problem in the application of the perturbation results presented in chapter 4 is the existence of the minus sign in the Hamiltonian (4.1.1),

$$H = H_0 - V$$

as discussed on page 43. Here we will construct some unitary operators that will enable us to perform transformations[1] on $H = H_0 + V$ such that we are left with a Hamiltonian of the form $\tilde{H} = UHU^{-1} = H_0 - V$.

As in chapter 5, let $c^\dagger_{x\sigma}, c_{x\sigma}$ denote creation and annihilation operators for fermions satisfying the anticommutation relations

$$\{c_{x\sigma}, c^\dagger_{y\tau}\} = \delta_{xy}\delta_{\sigma\tau},\ \{c_{x\sigma}, c_{y\tau}\} = \{c^\dagger_{x\sigma}, c^\dagger_{y\tau}\} = 0$$

where the subscripts x, y denote the location of the fermions in a given lattice Λ of arbitrary dimension d and σ, τ as usual denote the spin indices $\uparrow$ and $\downarrow$. Then let

$$u_{x\sigma} \equiv \exp(i\pi n_{x\sigma}),$$

where $n_{x\sigma} = c^\dagger_{x\sigma}c_{x\sigma}$ is the number operator. Consequently by Lie's series expansion,

$$
\begin{aligned}
u_{x\sigma}c_{x\sigma}u^{-1}_{x\sigma} &= \exp(i\pi n_{x\sigma})c_{x\sigma}\exp(-i\pi n_{x\sigma}) \\
&= c_{x\sigma}\exp(-i\pi n_{x\sigma}) \\
&= \begin{cases} -c_{x\sigma} & n_{x\sigma} = 1 \\ c_{x\sigma} & n_{x\sigma} = 0 \end{cases}
\end{aligned} \qquad \text{(A.4.1)}
$$

[1] A general discussion of canonical transformations on a linear chain can be found in [S 75]. The transformations (b) and (c) presented there are effectively identical to the BCT and to the Jordan-Wigner transformation used here.

where the last relation is assumed to be applied to some given state. Most of the operators we will be dealing with are either of the hopping type

$$c^{\dagger}_{x\sigma}c_{y\sigma} \tag{A.4.2}$$

or of the spin-exchange type

$$c^{\dagger}_{x\sigma}c_{x-\sigma}c^{\dagger}_{y-\sigma}c_{y\sigma} \tag{A.4.3}$$

A.4.1 BCT Transformation

We assume that the lattice Λ is bipartite, i.e. there exist disjoint sets Λ_A, Λ_B of the sites of Λ such that $\Lambda = \Lambda_A \cup \Lambda_B$. Then we define

$$U_{BCT} = \exp\left(i\pi \sum_{\substack{x\in\Lambda_A \\ \sigma=\uparrow,\downarrow}} n_{x,\sigma} \right) \tag{A.4.4}$$

where the subscript stands for *Bipartite Canonical Transformation.* The net effect of this transformation is a sign change in the creation and annihilation operators on the A sublattice, whereas the operators in the B sublattices remain invariant.

$$c_{x\sigma} = \begin{cases} -c_{x\sigma} & \text{if } x \in \Lambda_A \\ +c_{x\sigma} & \text{if } x \in \Lambda_B \end{cases} \tag{A.4.5}$$

The hopping operator $c^{\dagger}_{x\sigma}c_{y\sigma}$ changes its sign only if $x \in \Lambda_A$ and $y \in \Lambda_B$, and vice versa. The spin exchange operator $c^{\dagger}_{x\sigma}c_{x-\sigma}c^{\dagger}_{y-\sigma}c_{y\sigma}$ remains invariant for all possible x, y locations, as does the number operator $n_{x\sigma}$.

A.4.2 Jordan-Wigner Transformation

The previous transformation may be applied to more than one dimension. However, there exists an extremely useful transformation–the *Jordan-Wigner transformation* [JW 28]–applicable to one dimension only. The Jordan-Wigner transformation enables us to tackle the basic problem of fermionic systems, i.e. the ordering of the Fermi operators. Any change in this ordering may result in a change of the overall sign of the state. It should be clear that this is extremely important in the context of Perron-Frobenius theory, where we explicitly require the positivity preserving quality of operators.

The Jordan-Wigner transformation now provides us with a tool that changes the fermionic character of $c_{x\sigma}, c^{\dagger}_{x\sigma}$ to a more bosonic one. Since the new operators commute at different lattice sites and spin values, their ordering becomes less

important. Again, the transformation is built on equation A.4.1, but now a 'local' transformation matrix is used. The transformation rules are given by

$$\begin{aligned} b_{x\downarrow} &= c_{x\downarrow} \exp\left(i\pi \sum_{z=1}^{x-1} n_{z,\downarrow} \right) \\ b_{x\uparrow} &= c_{x\uparrow} \exp\left(i\pi \left[\sum_{z=1}^{|\Lambda|} n_{z,\downarrow} + \sum_{z=1}^{x-1} n_{z,\uparrow} \right] \right) \end{aligned} \tag{A.4.6}$$

The $b_{x\sigma}^{+}$ are given by the Hermitian conjugates of these defining relations. It can be verified that the so-called *paulions* $b_{i\sigma}$ obey the commutation relations [B 88]

$$[b_{x\sigma}, b_{y\tau}] = 0 \tag{A.4.7}$$

$$[b_{x\sigma}^{+}, b_{y\tau}^{+}] = 0 \tag{A.4.8}$$

In addition, paulions are hard-core particles, that is, they anticommute at equal sites such that

$$[b_{x\sigma}, b_{y\tau}^{+}] = \delta_{xy}\delta_{\sigma\tau}(\mathbb{1} - 2b_{x\sigma}^{+}b_{x\sigma}) \tag{A.4.9}$$

We note that the additional exponential factors in the transformation for $b_{x\uparrow}$ assure the compatibility of the above relations for operators of different spins.

Assuming nearest-neighbor hopping in the one-dimensional chain Λ, it is easy to see that the hopping operator remains invariant, i.e.

$$\begin{aligned} b_{x+1}^{+}b_x &= \exp(-i\pi \sum_{z=1}^{x} n_z) c_{x+1}^{\dagger} c_x \exp(i\pi \sum_{z=1}^{x-1} n_z) \\ &= \exp(-i\pi n_x) c_{x+1}^{\dagger} c_x \\ &= c_{x+1}^{\dagger} \exp(-i\pi n_x) c_x \\ &= c_{x+1}^{\dagger} \left[\mathbb{1} - i\pi n_x + (-i\pi n_x)^2 \ldots \right] c_x \\ &= c_{x+1}^{\dagger} c_x \end{aligned} \tag{A.4.10}$$

where we have dropped the spin index σ for convenience. The last equality comes from the relation $n_{x\sigma}c_{x\sigma} = 0$. For $x \neq y$ the spin exchange operator can be rewritten as a product of two hopping terms, i.e.

$$\begin{aligned} c_{x\sigma}^{\dagger} c_{x-\sigma} c_{y-\sigma}^{\dagger} c_{y\sigma} &= -c_{x\sigma}^{\dagger} c_{y\sigma} c_{y-\sigma}^{\dagger} c_{x-\sigma} \\ &= -b_{x\sigma}^{\dagger} b_{y\sigma} b_{y-\sigma}^{\dagger} b_{x-\sigma} \\ &= -b_{x\sigma}^{\dagger} b_{x-\sigma} b_{y-\sigma}^{\dagger} b_{y\sigma} \end{aligned} \tag{A.4.11}$$

We therefore may change the sign of the spin exchange operator if the corresponding transition matrix is bipartite.

A.4.3 Hole-Particle Transformation

We now consider another transformation, the *hole-particle transformation*, which is defined as

$$c_{x\downarrow} = d_{x\downarrow} \tag{A.4.12}$$

$$c_{x\uparrow} = \begin{cases} +d^{\dagger}_{x\uparrow} & \text{if } x \in \Lambda_A \\ -d^{\dagger}_{x\uparrow} & \text{if } x \in \Lambda_B \end{cases} \tag{A.4.13}$$

Thus, instead of dealing with spin up particles, we are now dealing with spin up holes; hence the name. The hopping operator structure remains invariant under the transformation, i.e. $c^{\dagger}_{x\sigma}c_{y\sigma} = d^{\dagger}_{x\sigma}d_{y\sigma}$ for a bipartite lattice. The number operators for the spin down particles remain the same, but the number operators for the spin up particles change, $n_{x\uparrow} = c^{\dagger}_{x\uparrow}c_{x\uparrow} = d_{x\uparrow}d^{\dagger}_{x\uparrow} = 1 - d^{\dagger}_{x\uparrow}d_{x\uparrow}$. This yields

$$N = n_{\uparrow} + n_{\downarrow} = \tilde{n}_{\downarrow} - \tilde{n}_{\uparrow} + 1 = \tilde{S}^3 + 1 \tag{A.4.14}$$

and

$$S^3 = n_{\uparrow} - n_{\downarrow} = 1 - \tilde{n}_{\uparrow} + \tilde{n}_{\downarrow} = 1 - \tilde{N} \tag{A.4.15}$$

Appendix B

Notation

Abbreviations

w.l.o.g.	without loss of generality
w.r.t.	with respect to

Symbols		Location
$\emptyset$	empty set	
N	set of natural numbers	
Z	set of integers	
R	real numbers	
R^+	real numbers greater than zero	
$[a,b]$	closed interval from a to b, $a \leq b$	
$]a,b[$	open interval from a to b, $a < b$	
C	complex numbers	
$\mathsf{N}[m,n]$	$\{m \in \mathsf{N}, m+1 \in \mathsf{N}, \ldots, n \in \mathsf{N}\}$	
$\Im$	(multi)index set	
$\Im[m,n]$	$\{m \in \Im, m+1 \in \Im, \ldots, n \in \Im\}$	
$M_{m,n}(\mathbf{F})$	$m \times n$ matrices over the field $\mathbf{F}$	
A^T	transpose of matrix A	
A^*	Hermitian adjoint of A	
$\lvert A \rvert$	absolute value matrix $[\lvert a_{ij} \rvert]$	
$\Gamma(A)$	directed graph of A	A.1.3
γ	directed path in Γ	A.1.4
P_{i_0}	nodes in γ	A.1.4
$I(A)$	indicator matrix of A	
$\sigma(A)$	spectrum of A	2.1.3

$\rho(A)$	spectral radius of A	2.1.3
$\tilde{A}$	reduced block form of A	2.3.1
$diag(d_1, \ldots, d_n)$	diagonal $n \times n$ matrix	
P	permutation matrix	
$\mathcal{X}$	Banach space	
$\mathcal{W}$	Hilbert space	
$(\cdot, \cdot)$	scalar product in $\mathcal{W}$	
$\mathcal{K}$	cone symbol	3.1.2
$\mathcal{K}^\circ$	interior of $\mathcal{K}$	
$\partial\mathcal{K}$	boundary of $\mathcal{K}$	
$\mathcal{K}^*$	dual cone of $\mathcal{K}$	3.3.8
$\Pi(\mathcal{K}_1, \mathcal{K}_2)$	cone of positivity preserving operators	3.1.25
$\succeq$	partial ordering	3.1.6
$\underset{\mathcal{K}}{\succeq}$	partial ordering induced by $\mathcal{K}$	3.1.6
$\sup(x, y)$	supremum induced by $\mathcal{K}$	3.3.1
$\inf(x, y)$	infimum	3.3.1
icc	"ice-cream-cone"	fig. 3.1
PSD	the cone of positive semidefinite matrices	
$\mathcal{V}$	space of $n \times n$ Hermitian matrices	3.6.5
$\langle \cdot, \cdot \rangle$	scalar product in $\mathcal{V}$	3.6.5
$A \circ B$	Hadamard product of A and B	3.6.5
$\|W\|$	modulus of W, $\|W\| = \sqrt{(W^2)}$	
Λ	lattice	5.1.1
$\|\Lambda\|$	number of sites in Λ	5.1.1
$\mathcal{H}(x)$	Hilbert space at site x	
$\mathcal{H}(\Lambda)$	Hilbert space of the whole Λ	
$\mathcal{P}$	projection operator	
$\mathbf{S}(\Lambda)^2, S_3(\Lambda)$	spin operators on Λ	p. 5.1
$S(S+1), M$	eigenvalues of spin operators	
$c^\dagger_{x\sigma}, c_{x\sigma}$	Fermi creation and annihilation operators	
$b^+_{x\sigma}, b_{x\sigma}$	Pauli operators	
$d^\dagger_{x\sigma}, d_{x\sigma}$	Fermi operators after hole-particle transformation	
$n_{x\sigma}$	number operator at site x and of spin σ	
p	number of pure states	
N	number of electrons	
n	number of spinless fermions $n = N/2$	
$a \in \Im[1, p]$	set of all pure states, a a single pure state	

$[\cdot,\cdot]$	commutator
$\{\cdot,\cdot\}$	anticommutator
$\diamond$	start-of-induction
$\hookrightarrow$	induction argument
∎	end-of-proof symbol
□	proof is found in the cited literature

Bibliography

[ABCKS88] A. A. Aharony, R. J. Birgeneau, A. Coniglio, M. A. Kastner, H. E. Stanley, *Magnetic Phase Diagram and Magnetic Pairing in Doped La_2CuO_4*, Phys. Rev. Lett. **60**, 1330, (1988)

[A87] P. W. Anderson, *The Resonating Valence Bond State in La_2CuO_4 and Superconductivity*, Science, **235**, 1196, (1987)

[B 67] G. Birkhoff, *Linear Transformations with Invariant Cones*, Amer, Math. Monthly **74**, 274, (1967)

[B 73] A. Berman, *Cones, Matrices and Mathematical Programming*, Springer Verlag, (1973)

[B 88] U. Busch, *Exakte Berechnungen zum Wolff-Modell: Endliche Systeme*, Diploma Thesis, FU Berlin, Department of Physics, (1988)

[BB 73] A. Berman, A. Ben-Israel, *Linear Equations over Cones with Interior: A Solvability Theorem with Applications to Matrix Theory*, Lin. Algebra and its Appl. **7**, 139, (1973)

[BB 90] P. A. Bares, G. Blatter, *Supersymmetric tJ-Model in One Dimensions: Seperation of Spin and Charge*, Phys. Rev. Lett. **64**, 2567, (1990)

[CFS 82] I. P. Cornfeld, S. V. Fomin, Ya. G. Sinai, *Ergodic Theory*, Springer Verlag, (1982)

[DLS 78] F. J. Dyson, E. Lieb, B. Simon, *Phase Transitions in Quantum Spin Systems with Isotropic and Nonisotropic Interactions*, J. Stat. Phys. **18**, 336, (1978)

[E 79] V. J. Emery, *Theory of the One-Dimensional Electron Gas* in *Highly Conducting One-Dimensional Solids*, Ed. J. T. Devreese, R. P. Evrard, V. E. van Doren, Plenum University Press, (1979)

[F 08] G. Froebenius, *Über Matrizen aus positiven Elementen*, S.-B. Preuss. Akad. Wiss. 471, (1908) and 514, (1909)

[F 12] G. Froebenius, *Über Matrizen mit nichtnegativen Elementen*, S.-B. Preuss. Akad. Wiss. 456, (1912)

[F 69] Ky Fan, *Asymptotic Cones and Duality of Linear Relations*, J. of Approx. Theory **2**, 152, (1969)

[F 72] W. Faris, *Invariant Cones and Uniqueness of the Ground State for Fermion Systems*, J. Math. Phys. **13**, 1285, (1972)

[F 91] A. Förster, *The Quantum Case of the tJ-model*, Ph. D. Thesis, FU Berlin, Department of Physics, (1991), in preparation

[FS 75] W. Faris, B. Simon, *Degenerate and Non-degenerate Ground States for Schrödinger Operators*, Duke Math. J. **42**, 559, (1975)

[FSS 76] J. Fröhlich, B. Simon, T. Spencer, *Infrared Bounds, Phase Transitions and Continuous Symmetry Breaking*, Comm. Math. Phys. **50**, 79, (1976)

[G 59] F. R. Gantmacher, *Applications of the Theory of Matrices* (translated by J. L. Brenner), Interscience New York, (1959)

[G 72] L. Gross, *Existence and Uniqueness of Physical Ground States*, J. Func. Ana. **10**, 52, (1972)

[GJ 70] J. Glimm, A. Jaffe, *The $\lambda(\phi^4)_2$ quantum field theory without cutoffs: II. The field operators and the approximate vacuum*, Ann. Math. **91**, 362, (1970)

[GJ 87] J. Glimm, A. Jaffe, *Quantum Physics*, 2nd Edition, Springer Verlag, (1987)

[HJ 85] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, (1985)

[JW 28] P. Jordan, E. Wigner, *Über das Paulische Äquivalenzverbot*, Z. f. Phys. **47**, 631, (1928)

[K 64] M. A. Krasnosel'skiĭ, *Positive Solutions of Operator Equations*, Noordhoff Ltd., (1964)

[K 66] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag, (1966)

[KR 48] M. G. Krein, M. A. Rutman, *Linear operators leaving invariant a cone in a Banach space*, Uspechi Mathematičeskich Nauk (N.S.) **3**, no. 1(23), 3–95, (1948)
translated and published by AMS in 1950 as Translation Number 26

[L 89] E. Lieb, *Two Theorems on the Hubbard Model*, Phys. Rev. Lett. **64**, 1201, (1989)

[LM 62a] E. Lieb, D. Mattis, *Theory of Electromagnetism and the Ordering of Electronic Energy Levels*, Phys. Rev. **125**, 164, (1962)

[LM 62b] E. Lieb, D. Mattis, *Ordering Energy Levels of Interacting Spin Systems*, J. Math. Phys. **3**, 749, (1962)

[LW 68] E. Lieb, F. Y. Wu, *Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension*, Phys. Rev. Lett. **20**, 1445, (1968)

[M 88] D. Mattis, *Statistical Mechanics of CuO_2 plane in the presence of localized holes*, Phys. Rev. B **38**, 7061, (1988)

[M 91] A. Mielke, *The One-Dimensional Hubbard Model for Large or Infinite U*, J. Stat. Phys. **62**, 509, (1991)

[P 07] O. Perron, *Grundlagen für eine Theorie des Jakobischen Kettenbruchalgorithmus*, Math. Ann. **64**, 248, (1907)

[P 53] R. S. Phillips, *Perturbation theory for semi-groups of linear operators*, Trans. Amer. Soc. **74**, 199, (1953)

[PT91] PHYSICS TODAY, *High-Temperature Superconductivity*, June (1991)

[R 40] F. Riesz, *Sur quelques notions fondamentales dans le théorie générale des opérations linéaires* Ann. of Math. **41**, 174, (1940)

[R 88] H. L. Royden, *Real Analysis*, Maxmillan Publishing Company, (1988)

[RS1 80] M. Reed, B. Simon, *Functional Analysis*, 2nd Edition, Academic Press, (1980)

[RS2 78] M. Reed, B. Simon, *Fourier Analysis, Self-Adjointness*, Academic Press, (1978)

[RS4 78] M. Reed, B. Simon, *Analysis of Operators*, Academic Press, (1978)

[S 73] E. Seneta, *Non-negative matrices and Markov chains*, Springer Verlag, (1973)

[S 74] H. H. Schäfer, *Banach Lattices and Positive Operators*, Springer Verlag, (1974)

[S 75] B. Sutherland, *Model for a multicomponent quantum system*, Phys. Rev. B **12**, 3795, (1975)

[S 85] E. R. Spear, *Failure of Reflection Positivity in the Quantum Heisenberg Ferromagnet*, Lett. Math. Phys. **10**, 41, (1985)

[S 87] J. Schröder, *Operator Inequalities and Applications* in *Inequalities: Fifty Years on from Hardy, Littlewood and Pólya*, Ed. W. N. Everitt, M. Dekker Inc., (1987)

[S 91] J. Schröder, private communication, (1991)

[S 87] P. Schlottmann, *Integrable narrow-band model with possible relevance to heavy-fermion systems*, Phys. Rev. B **36**, 5177, (1987)

[SV 70] H. Schneider, M. Vidyasagar, *Cross-positive Matrices*, SIAM J. Numer. Anal. bf 7, 508, (1970)

[T 66] O. Taussky, *Positive-Definite Matrices* in *Inequalities I*, Ed. O. Shisha, Academic Press, 309, (1966)

[T 67] O. Taussky, *Positive-Definite Matrices and Their Role in the Study of the Characteristic Roots of General Matrices* in *Inequalities II*, Ed. O. Shisha, Academic Press, 175, (1967) or O. Taussky, Adv. in Math. **2**, 175, (1968)

[T 72] O. Taussky, *Automorphs of Quadratic Forms as Positive Operators* in *Inequalities III*, Ed. O. Shisha, Academic Press, 341, (1972)

[V 68] J. S. Vandergraft, *Spectral Properties of Matrices which have Invariant Cones*, SIAM J. Appl. Math. **16**, 1208, (1968)

[W 72] B. Weiss, *Positive Cones in Hilbert Space and a Maximal Inequality* in *Inequalities III*, Ed. O. Shisha, Academic Press, 353, (1972)

[W 88] P. B. Wiegmann, *Superconductivity in Strongly Correlated Electronic Systems and Confinement versus Deconfinement Phenomenon*, Phys. Rev. Lett. **60**, 821, (1988)

[ZR 88] F. C. Zhang, T. M. Rice, *Effective Hamiltonian for superconducting Cu oxides*, Phys. Rev. B **37**, 3759, (1987)