Promotor:
Prof. Dr. Ir. A. van der Avoird
INTERMOLECULAR FORCES
AND THE
GROUP THEORY OF MANY-BODY SYSTEMS
INTERMOLECULAR FORCES
AND THE
GROUP THEORY OF MANY-BODY SYSTEMS

PROEFSCHRIFT
TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE
WISKUNDE EN NATUURWETENSCHAPPEN AAN DE KA-
THOLIEKE UNIVERSITEIT TE NIJMEGEN, OP GEZAG VAN
DE RECTOR MAGNIFICUS PROF. MR. F.J.F.M. DUYNSTEE
VOLGENS BESLUIT VAN HET COLLEGE VAN DECANEN
IN HET OPENBAAR TE VERDEDIGEN OP DONDERDAG
11 DECEMBER 1975 DES NAMIDDAGS TE 2.00 UUR
PRECIES.

DOOR

PAUL EDUARD STEPHAN WORMER

GEBOREN TE AMSTERDAM

1975

DRUK: STICHTING STUDENTENPERS NIJMEGEN
Aan mijn moeder
en Jenny
CONTENTS

PREFACE

PART I: GROUP THEORY OF MANY-BODY SYSTEMS.

CHAPTER I.1. INTRODUCTION

CHAPTER I.2. A REVIEW OF SOME ALGEBRA

2.1. Definition of a few different algebras
2.2. Semi-simple associative algebras
2.3. Representations of an algebra and its commutator algebra

CHAPTER I.3. THREE CLASSICAL (LIE) GROUPS

3.1. A review of some Lie group theory
3.2. The general linear group GL(n) and the unitary group U(n)
3.3. Connection between the irreps of GL(n) and U(n)
3.4. The special unitary group SU(n)

CHAPTER I.4. TENSOR REPRESENTATIONS OF GL(n) AND S_N

4.1. Tensor products of vector spaces
4.2. The space $V_n^\otimes N$ as a carrier space for representations of GL(n)
4.3. The space $V_n^\otimes N$ as a carrier space for representations of the Lie algebras of the classical groups
4.4. The space $V_n^\otimes N$ as a carrier space for representations of CS_N
4.5. Bisymmetric operators on $V_n^\otimes N$
4.6. The interconnection between the tensor irreps of GL(n) and S_N
4.7. A note on the Casimir invariants of GL(n)

CHAPTER I.5. THE DECOMPOSITION OF TENSOR SPACE

5.1. Young units
5.2. Young orthogonal units
5.3. Yamanouchi units

1
4
9
9
13
17
22
22
27
31
33
37
37
40
42
44
46
52
53
58
59
66
68
5.4. Graphical representation of projected tensors
5.5. Weyl-Rumer basis
5.6. Gelfand basis
5.7. Explicit construction of matrix representations of GL(n)
5.8. A note on the tensor irreps of the Lie algebra of GL(n)

CHAPTER I.6. PHYSICAL APPLICATIONS
6.1. N-particle state vectors
6.2. Symmetry of N-particle systems
6.3. A model for N-electron systems
6.4. Symmetry restrictions imposed by spin
6.5. A note on the calculation of matrix elements

REFERENCES PART I

REPRINTED PAPERS:
- Russell-Saunders states arising from configurations of equivalent electrons,
- Classification of interaction operators with respect to many-particle permutation symmetry,
- Transformation properties of antisymmetric spin eigenfunctions under linear mixing of the orbitals,
  J. Chem. Phys. 57, 2498 (1972)
- Transformation properties of many-electron wave functions with special attention to the relation between pair-correlated DODS and configuration interaction,
  Int. J. Quant. Chem. 8, 715 (1974)

PART II: INTERMOLECULAR FORCES.

CHAPTER II.1. INTRODUCTION
CHAPTER II.2. THE MULTIPOLAR EXPANSION OF \(1/r_{12}\)
CHAPTER II.3. LONG RANGE FORCES
REFERENCES PART II

REPRINTED PAPERS:
- Ab-initio valence bond calculations on the He-He potential curve using small bases, Mol. Phys. 29, 1181 (1975)

SUMMARY
SAMENVATTING
ACKNOWLEDGEMENTS
CURRICULUM
This thesis consists of two parts, which at first sight do not show much connection. However, the parts are related in that many of the group theoretical techniques, expounded in part I, have been applied in the calculations of part II. To account for the somewhat uncommon division of this dissertation into two parts, it is necessary to relate some of the history of the work.

My interest in the group theory of many-body systems was first aroused when I was working in Delft, under the supervision of Prof. L.L. van Reijen and in cooperation with Ir. J.J.M. Potters, on a spin-projected DODS (different orbitals for different spin) approach to the bonding in LiH. We worked within the AMO (alternant molecular orbital) scheme, which is a simplified version of the DODS method. (The results of this work are presented in ref. 15 of part I).

When later the work in Nijmegen on the calculation of intermolecular forces was started, we decided that it could be worthwhile to see what the AMO method would give for the London dispersion interaction between two He-atoms, because it is known that such an interaction is due to interatomic correlation and DODS methods are meant to give part of the correlation energy. So, the AMO programs were rewritten in FORTRAN (the original programs were in ALGOL), using the formula manipulation language FORMAC to derive and punch the really terrible formulas arising in a straightforward spin-projection (a Gelfand basis, discussed in sec. I.5.6, was used).

The AMO calculations gave remarkable results: using an atomic orbital basis consisting of s-orbitals only, a potential curve was obtained which closely resembled the experimental curves. This was most astonishing as it is known from London's theory (ch. II.3) that polarization functions (p-, d-, etc. orbitals) must be present on both the He-atoms to obtain a Van der Waals minimum. So we immediately knew the AMO results to be spurious. Furthermore, the AMO energies showed some unexpected regularities in their behaviour as a function of the variation parameters.

In order to explain these observations we had to look deeper into the transformation properties of spin-projected functions and to consider especially the behaviour of such functions under blocked orbital transformations, since the AMO method makes use of such a transformation.
(The results of this theoretical work are presented in two of the papers reprinted in this dissertation). It could now easily be proved that the He-He potential curve was due to a distance dependent intra-atomic correlation contribution to the AMO energy, thus showing that the AMO method was completely unsuited for this kind of work. The regularities in the energy as a function of the variation parameters could also be accounted for by the theory. In fact the AMO energy was proved to be a surprisingly simple function of these parameters, which also made clear that the formulas derived by FORMAC in a straightforward manner were unnecessarily complicated.

While doing this group theoretical work, I had to read parts of Weyl's and Boerner's books which I found an arduous job. But after having made this effort, it seemed that my experiences could be made useful by writing a survey that was self-contained and less condensed than the treatments in the books just mentioned. Since this review contained some material that could not be found explicitly in other places, we thought it worthwhile to include it in the first six chapters of this thesis.

Part II of this dissertation contains the results of the work on intermolecular forces which actually formed my main research interest during the past few years. For some introductory remarks on this part of the work the reader is referred to chapter II.1.
PART I:

GROUP THEORY OF MANY-BODY SYSTEMS
CHAPTER I.1. INTRODUCTION

Because of the great complexity of the Schrödinger equation anyone performing molecular calculations is forced to introduce from the outset a number of simplifying assumptions, some of them rather severe, others not too strong. No quantum chemist will doubt the usefulness of the Born-Oppenheimer approximation or question the truncation of one-particle Hilbert space (orbital space). The Russell-Saunders coupling scheme, too, is generally accepted (except, of course, in cases where strong spin-orbit coupling occurs). Stated in mathematical terms these three widely applied approximations constitute the replacement of an exact Hamilton operator by a zeroth order N-electron model Hamiltonian, from which all energy contributions arising from spin are removed, in which all the nuclei are clamped, and which operates on a finite-dimensional Hilbert space that is a tensor product of N orbital spaces. This model is called the multi-configuration method, and includes most of the current computational models in quantum chemistry, such as valence bond, configuration interaction, and other methods.

It is after this point that consensus ceases to exist, a multitude of different approximations simplifying the model has been proposed, some of them so drastic that they virtually allow the Schrödinger equation to be solved by hand. However, after the advent of fast and large computers none of these further simplifications is in principle indispensable, at least not for simple molecules. Moreover, if computers and computing methods keep on developing at their present rate, the time is not far off when chemically more interesting systems will become tractable within the multi-configuration model. Hence it becomes a point of more than just academic interest to study the details of the model and to assess all its properties.

Fortunately we do not have to start from scratch; as so often before in physics, mathematicians have already paved the way. To clarify this we must note that, apart from possible spatial symmetry groups, two symmetry groups are always associated with the multi-configuration Hamiltonian, viz. the permutation group $S_N$ and the special unitary group $SU(2)$. Furthermore, the general linear group $GL(n)$ enters the model in its rôle of dynamical group (a concept which is defined in chapter I.6). Of course, the representations of these three groups are not at all unrelated. As a matter of fact, one of the main purposes of the following exposition is the presentation of a coherent and unifying approach to the representation theory of
$S_n$ and the Lie groups $GL(n)$, $U(n)$ and $SU(n)$ together with their Lie algebras. These groups have indeed received much attention by mathematicians, and, after their bearing on quantum mechanics had been established, also by theoretical physicists.

The representation theory of $S_n$ is linked with the names of Frobenius, Schur, Young, Littlewood and others. Weyl discussed the relation between $S_n$ and the global representation theory of $GL(n)$, whereas Wigner, Racah and other physicists concentrated mainly on the subgroup $SU(2)$ of $GL(2)$. Casimir, for instance, showed that the total spin angular momentum operator $S^2$ commutes with the elements of the Lie algebra of $SU(2)$, thus tying the physical concept of spin with the mathematical concept of a Lie algebra, developed by Lie, Killing, Cartan, Weyl and others. During the last decade Biedenharn, Moshinksky, Louck and other workers have been generalizing the infinitesimal approach of $SU(2)$ to $SU(n)$ and $GL(n)$, for arbitrary $n$. Clearly an enormous group theoretical machinery is in existence, ready for application in quantum chemistry.

However, the question may come to mind: do we really need all this group theory for handling the multi-configuration model? The traditional approach, which employs (spin-projected) Slater determinants is known to go a long way and has the advantage of not requiring any difficult mathematics. But if this traditional approach is pushed a little bit further, especially into the direction of atomic theory, one cannot avoid using shift operators, vector coupling coefficients etc. and so willy-nilly to employ Lie algebra theory, in particular the part that is related to $SU(2)$. One may object that even if it is true that there is no way around $SU(2)$, it does not necessarily mean that we must go further and introduce $S_n$ and $GL(n)$ also. There are some answers to that. In the first place it seems a waste not to explore an existing mathematical theory so ideally applying to our model. More important, however, is that we deepen our insight considerably by studying this theory; that is to say, the researcher who does the studying deepens his own personal understanding. For this to be useful for the whole of theoretical chemistry something more has to come out of it, though. What we would really like is a set of tools that cannot be furnished by the Slater determinant-plus-$SU(2)$ approach alone. In atomic theory many instances of such results may be found, one only has to look into the books of Judd [1] or Wybourne [2] to find numerous examples. Especially the search, initiated by Racah, for complete sets of quantum
numbers has proved to be fruitful.

In the field of molecular quantum mechanics it is only very recently that people are beginning to realize that GL(n) and its unitary subgroup U(n) form a unifying framework from which arise such applications as for example:
- An algorithm for writing down the possible linearly independent anti-symmetric eigenstates of $\mathcal{S}^2$.
- A closed formula for the total number of such states.
- Explicit relations for the transformation properties of such states under orbital mixings.
- Different sets of rules to evaluate matrix elements of the multi-configuration Hamiltonian.

Still, the full power of the Lie algebraic approach to GL(n) and U(n) remains to be explored with molecular applications in mind.

The permutation group has received a good deal more thought by quantum chemists over the years, with the attention focusing on the systematic derivation of the Hamilton matrix elements over states derived from non-orthogonal orbitals.

It is sometimes argued that group theory going beyond simple point groups is too abstract and complicated for use in chemistry. However, one must realize that every mathematical theory looks abstruse and complex before it has found general acceptance. To illustrate this point, I will relate some of the conceptual problems mathematicians experienced when letters carrying signs came into use for denoting unknowns in simple algebraic equations. For example, Hieronymus Cardanus (1501-1576), one of the leading mathematicians of his time, published in 1570 a "proof" of the fact that the rule \((-a) \cdot (-b) = +a \cdot b\) does not hold for letters. Mind you, this rule had been known for numbers since Diophantus' time (third century). Thirty eight years later another mathematician, Christophorus Clavius (1537-1612), still could not quite grasp the use of letters, as is witnessed by the following citation from his famous textbook (translated from the Latin text quoted by Tropfke [3]):
"It seems that one must forgo the reason why the unknowns and their signs multiply as they do: this must be attributed to the shortcomings of the human mind that cannot comprehend in which manner it can be possibly true."

One of the main lessons that can probably be drawn from the history of science, is not to be deterred by an apparently esoteric theory.
The first four chapters of the subsequent review contain a rather thorough discussion concerning the tensor representations of GL(n), U(n) and SU(n) and the relation with the corresponding theory for the respective Lie algebras. Also the connection with the representation theory of $S_N$ is discussed in detail. Much of this material can be found in the books of Weyl [3, 4] and Boerner [5]. These books being written in a highly condensed style, the main objective of the present work is to give a self-contained account that is readable by the average quantum chemist, without demanding too much effort. To that end two introductory chapters, one on associative algebras and one on linear Lie groups, precede the chapters I.4 and I.5, which together form the backbone of the present review.

The relation between electron spin, SU(2), rotational and permutational symmetry is gradually worked out in a number of examples. In chapter I.6 the bearing of the foregoing chapters on the quantum mechanics of $N$-particle systems is discussed.

For reasons of space many of the proofs are omitted, but always with a reference (including page or theorem number) to a source where the proof can be found. The proofs that are given explicitly satisfy one or both of the two criteria:
- The theorem is of such importance that an explicit proof, clarifying the content of the theorem, is justified.
- No proof could be found in the literature.

When writing a review such as this it is unavoidable that one comes across numerous blank spots that must be filled out in the course of writing. For instance the following points do not come directly from other sources:
- A proof that every linear operator on tensor space, commuting with all permutations, can be written as a polynomial in the generators of GL(n) (sec. I.4.5).
- A proof that the (tensor) Casimir operators of GL(n) can be expressed in terms of the class sum operators of $S_N$ (sec. I.4.7).
- An explicit and simple formula for the matrix representations of $S_N$ carried by Young units (sec. I.5.1).
- An algebraic proof of the fact that Young orthogonal units and Yamanouchi units are identical (sec. I.5.3).
- A proof of the completeness and linear independence of the set of standard Weyl-Rumer tensors (sec. I.5.5).
- A proof of the completeness and linear independence of the set of standard Gelfand tensors (sec. I.5.6).
- An explicit formula for the matrix representations of GL(n) carried by Gelfand or Weyl-Rumer tensors (sec. I.5.7).
- The discussion of N-particle symmetry (sec. I.6.2).
- The realization that GL(n) is a dynamical group of the Hamiltonian (sec. I.6.3).
- The antisymmetrizer expressed in terms of Young units (sec. I.6.4).
- The explicit relation between the conventional approach via spin-bonded functions and the spin-free approach using Weyl-Rumer tensors (sec. I.6.5).

Of course many of these points are already implicitly contained in the work of others, and not having made a complete literature survey I may have easily missed earlier explicit proofs of the same facts. In this connection a series of stimulating lectures by professor Matsen, given at this university in October 1974, must especially be mentioned. That was when my attention was first drawn to several of the points listed above.

The review is finally followed by a few papers published earlier. Each of these carries its own introduction; the reader is referred to those for the purpose and scope of these articles.
CHAPTER I.2. A REVIEW OF SOME ALGEBRA

This chapter summarizes some abstract algebra which is required in the later sections, and which is probably not general knowledge among chemists. Not reviewed are the following better known subjects, also having much bearing on the first half of this thesis:
- finite groups and their irreducible representations; for instance ex-pounded in ref. 4, chapters I.1 - I.5 and II.1 - II.5;
- linear operators (and their matrices) on finite dimensional vector spaces (see e.g. ref. 5, chapters I, II, V and VI).

Most of the theory of this chapter can be found scattered and interwoven with other material in the books of Weyl [6, 7] and Boerner [8]. References 9, 10 and 11 should be consulted especially for the properties of semi-simple associative algebras, reviewed briefly in sec. I.2.2.

2.1. Definition of a few different algebras

**Definition.**
A set $A$ is an **associative algebra** over the field of complex numbers $C$ if it is a vector space over $C$ and has an internal binary operation (multipli-
cation) satisfying the properties:

A1. $a(b + c) = ab + ac$; $(b + c)a = ba + ca$;
A2. $a(bc) = (ab)c$ (the associativity property);
A3. $(\lambda a)b = \lambda(ab) = a(\lambda b)$;
for all $a, b, c \in A$ and $\lambda \in C$.

**Definition.**
A set $\Lambda$ is a **Lie algebra** over the field of real numbers $R$ if it is a vector space over $R$ and has an internal binary operation (commutator bracketing) satisfying the properties:

L1. $[a, b + c] = [a, b] + [a, c]$; $[b + c, a] = [b, a] + [c, a]$;
L2. $[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$;
L3. $[a, a] = 0$;
L4. $[\lambda a, b] = \lambda[a, b] = [a, \lambda b]$;
for all $0, a, b, c \in \Lambda$ and $\lambda \in R$.

When the field $R$ is extended to $C$, $\Lambda$ becomes a complex Lie algebra: the **complexification** of $\Lambda$. 

9
Theorem. 
\[ [a, b] = -[b, a]. \]

Proof. Substitute \( a = a + b \) in L3 and use L1 and L3.

Notes.
1. Although it is not strictly necessary, we will assume an algebra (Lie or associative) to be finite-dimensional in the vector space sense. Further we assume associative algebras to have a unit element.
2. In the sequel we will almost exclusively be dealing with algebras of linear operators. Linear operators on a vector space can be multiplied by a scalar, added and multiplied, in all cases yielding again a linear operator. (The multiplication of linear operators is associative). Thus, if a set of linear operators is closed under these three algebraic operations the set is an associative algebra.
3. The commutator bracket:
\[ [A, B] = AB - BA \]
of two linear operators \( A \) and \( B \) is easily shown to satisfy the requirements L1 - L4. Thus, if a set of operators is closed under multiplication by a scalar, addition and commutator bracketing, the set is a Lie algebra.

Definition.
Let \( \{a_i\} \) be a basis of the associative algebra \( A \). Since \( a_i a_j \in A \) we can write:
\[ a_i a_j = \sum_k \gamma_{ijk} a_k, \quad \gamma_{ijk} \in C. \]
The constants \( \gamma_{ijk} \) are the structure constants of \( A \). Analogously one defines the structure constants of Lie algebras.

Example.
Define the following three matrices, representing the three components of the angular momentum operator \( \mathbf{\hat{L}} \):
\[ L_x = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L_z = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \]
The well-known commutation relations of angular momentum operators yield:
\[ [L_i, L_j] = i \sum_{k=1}^{3} \epsilon_{ijk} L_k, \quad i, j, k = x, y, z = 1, 2, 3. \]
The constants $\varepsilon_{ijk}$ are the antisymmetric Levi-Civita symbols defined by:

\[
\varepsilon_{ijk} = +1 \quad \text{if } i < j < k \text{ and any even permutation of } i, j \text{ and } k.
\]

\[
= -1 \quad \text{for an odd permutation of } i, j \text{ and } k.
\]

\[
= 0 \quad \text{if two or more indices are equal.}
\]

The 3-dimensional linear space generated by $L_x$, $L_y$, $L_z$ is a (matrix) Lie algebra with structure constants $i\varepsilon_{ijk}$. Later we will show this algebra to be the Lie algebra of $SO(3)$ (sec. I.3.1).

Theorem.
The set of all linear operators on an n-dimensional vector space $V_n$ is an $n^2$-dimensional associative algebra, denoted by $AL(n)$.

Proof. Ref. 5, theorems 2.6 and 2.8.

Linear operators on an n-dimensional space are faithfully represented by $n \times n$ matrices. The set of all $n \times n$ matrices $AL(n,C)$ is therefore an $n^2$-dimensional associative algebra too.

Definition.
Let $S$ be a subset of an associative algebra $A$. Close $S$ multiplicatively. That is, augment $S$ with all possible different products of elements in $S$, where the product is the associative product of $A$. This yields a new subset $S'$ of $A$: the multiplicative closure of $S$ (a semigroup). Addition and multiplication by scalars are defined on $S'$, since it is a subset of $A$. Close $S'$ linearly with respect to these vector space operations. That is, form the vector space $\mathcal{E}(S)$ generated by the maximal number of linearly independent elements in $S'$. The algebra $\mathcal{E}(S)$, the smallest subalgebra of $A$ containing $S$, is the enveloping algebra of $S$ (in $A$).

Example.
Let $\Lambda \subset AL(n)$ be the set of all skew-Hermitean operators. We will show in sec. I.3.2 that $\Lambda$ is a Lie algebra over $R$, and hence is by definition linearly closed. Extending the field $R$ to $C$, $\Lambda$ ceases to be closed, for $i\Lambda$, $A \in \Lambda$, is Hermitean. It is easily shown that $i\Lambda$ is the set of all Hermitean operators. Since it is quite easy to prove that every linear operator can be written uniquely as the sum of a skew-Hermitean and a Hermitean operator, we find:

\[AL(n) = \Lambda \oplus i\Lambda\]
and so $\text{AL}(n)$ is the enveloping algebra (over $\mathbb{C}$) of $\Lambda$. Because, as we will see later (sec. I.3.2), $\Lambda$ is the Lie algebra of $U(n)$, and because $\text{AL}(n)$ is the Lie algebra of $\text{GL}(n)$ (sec. I.3.2) we have found the interesting result that the Lie algebra of $\text{GL}(n)$ is the complexification of the Lie algebra of $U(n)$. We will use this fact later, when discussing Weyl's "Unitarian Trick" (sec. I.3.2).

**Definition.**

Let $G$ be a finite abstract group of order $\,^o G$. Since by definition $G$ is already closed under multiplication, its linear closure, consisting of the formal sums:

$$
\sum_{g \in G} c(g)g, \quad g \in G, \ c(g) \in \mathbb{C},
$$

is an algebra of dimension $\,^o G$. This algebra is the **group algebra** of $G$, denoted by $\mathbb{C}G$.

---

**Note.**

In this definition the adjective 'formal' is essential (see e.g. ref. 12, p. 4 for a rigorous definition of the formal sum). If it happens that addition and multiplication by scalars have already been defined on $G$, then the formal sum must be clearly distinguished from these operations. For example, on the group $O(3)$, consisting of all $3 \times 3$ real orthogonal matrices, addition and multiplication by a scalar are defined as is usual for matrices. The enveloping algebra of $O(3)$ in $\text{AL}(3)$ is finite-dimensional. However, the group algebra of $O(3)$ is of infinite dimension, since $O(3)$ is of infinite order. (In fact this group algebra is a pre-Hilbert space; see ref. 13, p. 408).

**Definitions.**

1. Let $\phi$ be a mapping of a group $G$ into a group $G'$, then $\phi$ is a **group homomorphism** if it satisfies:

$$
\phi(gg') = \phi(g)\phi(g'), \quad g, g' \in G.
$$

2. Let $\phi$ be a mapping of an algebra $A$ into an algebra $A'$, then $\phi$ is an **algebra homomorphism** if it satisfies:

$$
\phi(aa') = \phi(a)\phi(a')
$$

$$
\phi(\lambda a + \mu a') = \lambda\phi(a) + \mu\phi(a'), \quad a, a' \in A; \lambda, \mu \in \mathbb{C}.
$$

If in these two cases $\phi$ is a 1-1 mapping and onto, we speak of a group or an algebra **isomorphism** respectively.
Notes.
1. From here on only enveloping algebras of sets of linear operators on n-dimensional spaces will be considered. It is understood that these algebras are always imbedded in $\mathfrak{A}(n)$, without mentioning this explicitly.

2. The enveloping algebra of an $n$-dimensional irrep $D$ of a finite group $G$ is always equal to the full matrix algebra $\mathfrak{A}(n)$. This follows from Burnside's theorem [8, p. 65], which states that $D(G)$ contains exactly $n^2$ linearly independent operators.

We will find several examples in the sequel where the irreducibility of a group representation is proved by first ascertaining the same fact for its enveloping algebra. The validity of this procedure derives from the following theorem, holding for general groups.

Theorem.
The enveloping algebra $\mathcal{E}(G)$ of an operator group $G$ is irreducible if and only if $G$ is irreducible.

Proof. Apply the argument of ref. 9, p. 45.

Commutator algebra.
If $S$ is an arbitrary set of linear operators on a vector space $V$, it is easy to show that the set:

$$S^c = \{s^c \mid s^c s = ss^c, \forall s \in S\}$$

consisting of all operators on $V$ that commute with the elements of $S$ is an algebra. The algebra $S^c$ is called the intertwining, commutator or commuting algebra of $S$.

2.2. Semi-simple associative algebras

Almost all associative algebras encountered in quantum mechanics have completely reducible representations. Just as is the case with groups, this property of the representation is solely determined by the structure of the algebra itself. Unfortunately it requires too many pages to present something remotely resembling a self-contained account of the structure of algebras, and so we have to make do with only a sketch of the theory.
Definitions.
1. A linear subspace $L$ of an associative algebra $A$ is a left ideal if it is stable under left multiplication with the elements of $A$. So, $L$ has the property:
   \[ aL \subseteq L, \quad \forall a \in A. \]
2. If $L$ does not contain any left ideals except itself and $(0)$, it is a minimal left ideal.
3. In an analogous way one defines a (minimal) right ideal.
4. An element $e \in A$ is an idempotent if $e$ has the property: $e^2 = e$.
5. Two elements $a', a'' \in A$ are orthogonal if $a'a'' = 0$.
6. Two subsets $S'$ and $S''$ of $A$ are orthogonal if $S'S'' = (0)$; that is: $s's'' = 0$, $\forall s' \in S'$ and $\forall s'' \in S''$.
7. An idempotent $e$ is primitive if no two non-zero and orthogonal idempotents $e'$ and $e''$ exist, such that $e = e' + e''$.

The set $Aa$, for some arbitrary fixed $a \in A$, is a left ideal generated by $a$.

Theorem.
Let $L$ be the left ideal generated by the idempotent $e$. $L$ is minimal if, and only if, $e$ is primitive.

Definitions.
1. A set $A' \subseteq A$ which is simultaneously a left and a right ideal in $A$ is a two sided ideal (other names: ideal, invariant subalgebra).
2. A finite dimensional algebra that contains no ideals except itself and $(0)$ is a simple algebra.

Theorem.
(i) The unit $e$ of a simple algebra $A$ can be decomposed thus:
   \[ e = e_1 + e_2 + \ldots + e_f, \]
   where the $e_i$ are primitive orthogonal idempotents.
(ii) A simple algebra $A$ is a direct sum of orthogonal minimal left ideals generated by the $e_i$:
   \[ A = L_1 \oplus L_2 \oplus \ldots \oplus L_f, \]
   where the left ideals are all of the dimension $f$; so, $A$ is of dimension $f^2$. 
Theorem: (Wedderburn).

(i) The minimal left ideal $L_i$ occurring in the decomposition of a simple algebra $A$ has a basis: $\{e_{ji} | j = 1, \ldots, f \}$, $i = 1, \ldots, f$. This basis has the property:

$$e_{ij} e_{kl} = \delta_{jk} e_{il}, \quad i, j, k, l = 1, \ldots, f.$$ 

(ii) If $a \in A$ is expressed with respect to this basis as follows:

$$a = \sum_k \sum_{\ell} D(a)_{k\ell} e_{k\ell}$$

then:

$$ae_{ij} = \sum_k \sum_{\ell} D(a)_{ki} e_{\ell j}, \quad a \in A, \; D(a)_{ki} \in \mathbb{C}.$$ 

(iii) The set $\{D(a) | a \in A\}$ comprises all $f \times f$ matrices and is therefore the full matrix algebra $\text{AL}(f, \mathbb{C})$, which hence is a faithful representation of $A$.

Note.

The matrix representation $D(a)$ of $a \in A$ does not depend on the choice of the minimal left ideal that carries the representation.

Example.

The full matrix algebra $\text{AL}(n, \mathbb{C})$ is the best known example of a simple associative algebra; it is of the dimension $n^2$. The subset $L_i$, consisting of matrices with zeros everywhere except in column $i$, is an $n$-dimensional minimal left ideal with basis $\{E^{ji} | j = 1, \ldots, n\}$, where $E^{ji}$ is a matrix with all elements equal to zero, except the $(j,i)$-element, which is equal to one. The decomposition of $\text{AL}(n, \mathbb{C})$ into minimal left ideals is obvious, the corresponding decomposition of the unit is:

$$E = \sum_{i=1}^{n} E^{ii}.$$ 

One can proceed analogously with the minimal right ideals, which are sets of matrices with one row filled and zeros elsewhere.

Definitions.

1. An algebra $A$ is **semi-simple** if it is a direct sum of simple algebras:

$$A = \sum_{\lambda} \Phi A^{(\lambda)},$$

with:

$$A^{(\lambda)} A^{(\mu)} = (0).$$
The second condition implies that the representations of A afforded by the different simple algebras in this decomposition are non-equivalent.

2. The simple algebras $A^{(\lambda)}$ are called the simple components of A.

**Theorem.**
The unit $e$ of a semi-simple algebra A can be written as:

$$e = \sum_{\lambda} e^{(\lambda)}$$

where $e^{(\lambda)}$ is the unit of the simple component $A^{(\lambda)}$ of A.

The following theorem provides a criterion enabling us to decide whether an algebra is semi-simple.

**Theorem.**
An algebra A is semi-simple if and only if each of its ideals is direct summand. That is, if $A'$ is an ideal in A, there exists always an ideal $A''$ in A, such that $A = A' \oplus A''$.

**Examples.**
1. Let A be a decomposable operator algebra on a vector space $V$. This means that each A-invariant subspace of V has an A-invariant complement in V.

So, if $U \subseteq V$ is A-invariant, there exists an A-invariant space $W \subseteq V$ such that $V = U \oplus W$. An element $\alpha \in A$ decomposes accordingly as:

$$\alpha = \alpha_U \oplus \alpha_W,$$

where $\alpha_U$ and $\alpha_W$ are the restriction of $\alpha$ to U and W, respectively. The sets $A_U = \{\alpha_U \oplus 0\}$ and $A_W = \{0 \oplus \alpha_W\}$ are ideals in A with the property:

$$A = A_U \oplus A_W.$$

So, a decomposable algebra is semi-simple.

2. The most important example of a semi-simple algebra is the group algebra CG of a finite group G. Recall [14] that CG has a basis of elements ("Wigner operators") of the form:

$$e^{(\lambda)}_{ij} = \sum_{g \in G} D^{(\lambda)}(g^{-1})_{ji} g,$$

where $D$ is an irreducible matrix representation of G.

The set $\{e^{(\lambda)}_{ij} | i, j = 1, \ldots, f(\lambda); \lambda \text{ fixed}\}$ spans the simple component $A^{(\lambda)}$ of CG. The element
\[ e^{(\lambda)} = \sum_{i=1}^{f(\lambda)} e^{(\lambda)}_{ii} = \sum_{g \in G} \chi^{(\lambda)}(g^{-1}) \ g, \]

where \( \chi^{(\lambda)}(g^{-1}) \) is the trace of \( D^{(\lambda)}(g^{-1}) \), is the unit of \( A^{(\lambda)} \). The unit \( e^{(\lambda)} \) is often called a "character projector".

3. Later in this chapter it will be shown that the commutator algebra of a semi-simple operator algebra is semi-simple too.

The following two theorems concern the representations of semi-simple algebras. The first is the analogue of Maschke's theorem [4, p. 95], and the second is a generalization of a well-known theorem holding for the regular representation of finite groups [4, p. 108].

**Theorem.**
Every linear representation of a semi-simple algebra is decomposable.

**Theorem.**
Every linear irreducible representation (irrep) of a semi-simple algebra is equivalent to one of the irreps carried by the algebra itself.

### 2.3. Representations of an algebra and its commutator algebra

In this section we will consider how \( V \) decomposes under two sets of commuting operators, and see under which conditions the decompositions involved are complete.

**Definition.**
The Kronecker product \( R \otimes S \) of two arbitrary matrices \( R \) and \( S \) is defined as the matrix:

\[
R \otimes S = \begin{pmatrix}
R_{11}S & R_{12}S & \cdots & R_{1m}S \\
R_{21}S & \cdots & \cdots & \cdots \\
\vdots & \cdots & \cdots & \cdots \\
R_{n1}S & \cdots & \cdots & R_{nm}S
\end{pmatrix}
\]

So \( R \otimes S \) is a matrix with the matrix \( R_{ij}S \) at position \((i,j)\). Another way of looking at \( R \otimes S \) is to regard it as a matrix with its elements (scalars) labelled by four indices:
\[(R \otimes S)_{ij;kl} = R_{ik} S_{jl}.\]

The following multiplication rule for Kronecker products can be easily proved:

\[(R \otimes S)(R' \otimes S') = (RR') \otimes (SS').\]

**Note.**
Kronecker products can also be defined for more than two factors, postulating associativity.

Now we state an important theorem, which will be one of the cornerstones of the theory expounded in I.4.

**Theorem.**
Let \(A_n\) be a semi-simple algebra of operators on the vector space \(V_n\). Let \(V_n\) be decomposed in \(A_n\)-irreducible subspaces:

\[V_n = \sum_{\lambda=1}^{p} \sum_{j=1}^{f(\lambda)} \oplus U_{j}^{(\lambda)}\]

(1)

Let the spaces \(U_{j}^{(\lambda)}\), \(j = 1, \ldots, f(\lambda)\) carry identical irreducible matrix representations of dimension \(n^{(\lambda)}\).

If \(A^n_c\) is the commutator algebra of \(A_n\), then we will prove that a basis of \(V_n\), adapted to the decomposition (1), is simultaneously adapted to the decomposition:

\[V_n = \sum_{\lambda=1}^{p} \sum_{i=1}^{n^{(\lambda)}} \oplus W_{i}^{(\lambda)}\]

(2)

The spaces \(W_{i}^{(\lambda)}\), \(i = 1, \ldots, n^{(\lambda)}\) carry identical irreducible \(f(\lambda)\)-dimensional matrix representations of \(A^n_c\).

Note that the frequency factor \(f(\lambda)\) and the dimension \(n^{(\lambda)}\) have changed roles on going from decomposition (1) to (2).

**Proof.** Let

\[\{u_{k\ell}^{(\lambda)} | \lambda = 1, \ldots, p; k = 1, \ldots, f(\lambda); \ell = 1, \ldots, n^{(\lambda)}\}\]

be a basis of \(V_n\) corresponding to decomposition (1). The matrix \(A\) of \(A \in A_n\) is defined by:

\[A_{i\mu;kl}^{(\lambda)} = \sum_{\mu=1}^{p} \sum_{j=1}^{f(\lambda)} \sum_{\ell=1}^{n^{(\lambda)}} A_{ij;kl}^{(\mu;\lambda)} u_{i\mu}^{(\lambda)} u_{j\ell}.\]
Since $U_k^{(\lambda)}$ is stable under $A \in A_n$, the matrix of $A$ takes the form:

$$A_{ij;k}^{(\mu;\lambda)} = \delta_{\mu,\lambda} \delta_{ik} D^{(\lambda)}(A)_{j,k}.$$ 

The matrix $D^{(\lambda)}(A)$ represents $A$ on $U_k^{(\lambda)}$. Clearly

$$\delta_{ik} D^{(\lambda)}(A)$$

is the $(i,k)$-element of

$$E_f^{(\lambda)} \otimes D^{(\lambda)}(A),$$

where $E_f^{(\lambda)}$ is the $f^{(\lambda)} \times f^{(\lambda)}$ unit matrix. So, the matrix $A$ of $A \in A_n$ can be written as:

$$A = \sum_{\lambda=1}^{P} E_f^{(\lambda)} \otimes D^{(\lambda)}(A),$$

or alternatively:

$$A = \sum_{\lambda=1}^{P} E_f^{(\lambda)} \otimes D^{(\lambda)}(A),$$

where $E_p$ is the $p \times p$ unit matrix.

We will now show that the matrix $B$ of $B \in A_n^C$ has a very similar form. From $AB - BA = 0$ follows $AB - BA = 0$, or:

$$\sum_{\lambda=1}^{P} \sum_{q=1}^{n^{(\lambda)}} \sum_{r=1}^{n^{(\lambda)}} \left[A_{ij;r}(\mu;\nu) B_{rq;k}(\nu;\lambda) - B_{ij;r}(\mu;\nu) A_{rq;k}(\nu;\lambda)\right] = 0,$$

and substituting the special form of $A$ it follows that:

$$\sum_{\lambda=1}^{n^{(\lambda)}} \left[D_{ij;k}(\mu;\lambda) B_{iq;k}(\lambda;\mu) - B_{ij;k}(\mu;\lambda) D^{(\lambda)}(A)_{q,k}\right] = 0,$$

which reads in matrix notation:

$$D^{(\mu)}(A) B^{(\mu;\lambda)}_{i;k} - B^{(\mu;\lambda)}_{i;k} D^{(\lambda)}(A) = 0,$$

where $B^{(\mu;\lambda)}_{i;k}$ is an $n^{(\lambda)} \times n^{(\mu)}$ matrix with the element $B_{iq;k}(\lambda;\mu)$ at the $(q,k)$ position.

Now by Schur's lemma [4, p. 98] we get:

$$B^{(\mu;\lambda)}_{i;k} = c E_n^{(\lambda)} \quad \text{if } \lambda = \mu$$

$$= 0 \quad \text{if } \lambda \neq \mu.$$ 

The Schur constant $c$ depends on $B$, $\lambda$, $i$ and $k$, and so we write $D^{(\lambda)}(B)_{ik}$ rather than $c$, thus making clear that these constants form an $f^{(\lambda)} \times f^{(\mu)}$ matrix, when keeping $B$ and $\lambda$ fixed.
The matrix $B$ of $B$ can finally be written as:

$$B = \sum_{\lambda=1}^{p} \otimes D^{(\lambda)}(B) \otimes E^{(\lambda)}_n$$

or alternatively:

$$B = \sum_{\lambda=1}^{p} \otimes D^{(\lambda)}(B) \otimes E^{(\lambda)}_n$$

Up to this point we never used the fact that $A_n^C$ is an algebra, and indeed the result just found holds for an arbitrary set of operators commuting with $A_n$. To prove, as the next step, the irreducibility of $\{D^{(\lambda)}(B)\}$ we do need the fact that $A_n^C$ contains all operators commuting with $A_n$, which makes $A_n^C$ automatically an algebra.

Every operator on $V_n$ with a matrix of the form:

$$E^{(\lambda)}_p \otimes C \otimes E^{(\lambda)}_n$$

arbitrary $C \in \text{AL}(f^{(\lambda)}, C)$,

which is defined with respect to the basis $\{u_{k\ell}^{(\lambda)}\}$, belongs to $A_n^C$. This follows from the commutation of such a matrix with the matrices of $A_n$:

$$\left(E^{(\lambda)}_p \otimes C \otimes E^{(\lambda)}_n\right) \left(E^{(\lambda)}_p \otimes E^{(\lambda)}_{f^{(\lambda)}} \otimes D^{(\lambda)}(A)\right) = \left(E^{(\lambda)}_p \otimes E^{(\lambda)}_{f^{(\lambda)}} \otimes D^{(\lambda)}(A)\right) \left(E^{(\lambda)}_p \otimes C \otimes E^{(\lambda)}_n\right).$$

Hence the matrix $E^{(\lambda)}_p \otimes C \otimes E^{(\lambda)}_n$ is the matrix of an element in $A_n^C$.

Since this holds for all possible $f^{(\lambda)} \times f^{(\lambda)}$ matrices $C$, we find that the set $\{D^{(\lambda)}(B)\}$ is the full matrix algebra. The full matrix algebra being by definition irreducible, it follows that the set $\{D^{(\lambda)}(B)\}$ constitutes an irreducible matrix representation of $A_n^C$.

As a last step we show which spaces carry these irreps of $A_n^C$. The matrix $B$ of $B \in A_n^C$ is defined by:

$$B_{k\ell}^{(\lambda)} = \sum_{\lambda=1}^{p} \sum_{i=1}^{f^{(\lambda)}} B^{(\mu;\lambda)}_{ik;\ell} u_{i\ell}^{(\mu)}.$$ 

Inserting the expression just found for $B$:

$$B_{k\ell}^{(\lambda)} = \sum_{i=1}^{f^{(\lambda)}} D^{(\lambda)}(B)_{ik} u_{i\ell}^{(\lambda)}$$

we see that the space $W_{k\ell}^{(\lambda)}$ spanned by $\{u_{i\ell}^{(\lambda)} | i = 1, \ldots, f^{(\lambda)}\}$ carries the $f^{(\lambda)}$-dimensional irreducible matrix representation $D^{(\lambda)}(B)$ of $A_n^C$.

To clarify the meaning of this theorem somewhat further we write the basis elements of $V_n$ belonging to one fixed $\lambda$ in a rectangular scheme,
following ref. 15.

\[
\begin{array}{cccccc}
  u^{(\lambda)}_{11} & u^{(\lambda)}_{12} & \cdots & \cdots & u^{(\lambda)}_{1n} \\
  u^{(\lambda)}_{21} & \cdots & \cdots & \cdots & \cdots \\
  \vdots & \cdots & \cdots & \cdots & \cdots \\
  \vdots & \cdots & \cdots & \cdots & \cdots \\
  u^{(\lambda)}_{n1} & \cdots & \cdots & \cdots & u^{(\lambda)}_{n1} \\
  f^{(\lambda)}_1 & \cdots & \cdots & \cdots & f^{(\lambda)}_n \\
\end{array}
\]

The rows in this scheme span the \( A_n \)-irreducible spaces \( U^{(\lambda)}_j \) occurring in the decomposition (1) of \( V_n \), and the columns span the \( A_n^c \)-irreducible spaces \( W^{(\lambda)}_i \) occurring in the decomposition (2) of \( V_n \).

Corollary.

The commutator algebra \( A_n^c \) of a semi-simple operator algebra \( A_n \) is semi-simple:

Proof. The theorem shows that \( A_n^c \) is decomposable; we have seen earlier that a decomposable operator algebra is semi-simple.

Corollary.

For a semi-simple operator algebra \( A_n \), the property of being a commutator algebra is symmetric, or:

\[ (A_n^c)^c \equiv A_n. \]

Proof. If \( V_n \) is decomposed under \( A_n \), it is also decomposed under \( A_n^c \) (by the foregoing theorem). Copying the proof of the foregoing theorem with \( A_n^c \) taking the place of \( A_n \), we find that the algebra \((A_n^c)^c\) is represented by all matrices of the form: \( E_p \otimes E^{(\lambda)}_f \otimes D^{(\lambda)}(A), A \in (A_n^c)^c \). However, \( A_n \) is semi-simple, and so it contains all the matrices of this form (by Wedderburn's theorem). In other words \( A_n \) and \((A_n^c)^c\) coincide.

As we will see in sec. I.4.7, this corollary has the interesting consequence that the Casimir operators of \( GL(n) \) and \( U(n) \) are linear combinations of permutation operators.
CHAPTER I.3. THREE CLASSICAL (LIE) GROUPS

The group theoretical approach to many body theory depends extensively on the representation theory of the Lie group GL(n) and its Lie subgroups. This family of linear Lie groups has been christened Classical Groups by Weyl [7]. Besides GL(n), the classical groups U(n) and SU(n) always play a rôle in systems of indistinguishable particles, irrespective of whether the system has geometrical symmetry or not.

In this chapter these three groups and their Lie algebras are presented, but first some of the necessary Lie group theory will be briefly outlined. For more details on this the books of Miller [13] and of Hausner and Schwartz [16] are especially recommendable.

3.1. A review of some Lie group theory

**Definition.**
Consider an open subset V of the r-dimensional Euclidean space \( \mathbb{R}^r \); the field \( \mathbb{F} \) is either the complex field \( \mathbb{C} \), or the real field \( \mathbb{R} \). Let the zero vector \( \vec{0} = (0, 0, \ldots, 0) \) belong to V. Let \( \vec{A} : \hat{g} \to \hat{A}(\hat{g}) \) be a mapping from V into the full matrix algebra \( \mathbb{A}(n, \mathbb{C}) \), where \( \hat{A}(\hat{g}) \) is a matrix depending on the r parameters \( \hat{g} \equiv (g_1, \ldots, g_r) \). Let V be such that the \( \hat{A}(\hat{g}) \) form a group; more precisely the set \( G = \{ \hat{A}(\hat{g}) | \hat{g} \in V \} \) is an r parameter linear Lie group, if it satisfies the following postulates:

G1. \( \hat{A}(\hat{0}) = \mathbb{1} \) (the n x n unit matrix).

G2. G is a group with matrix multiplication as the group multiplication. (So, the matrices in G must be non-singular).

G3. The matrix elements \( \hat{A}(\hat{g})_{ij}, i, j = 1, \ldots, n \), are analytic functions of the r parameters \( g_1, g_2, \ldots, g_r \).

G4. The r parameters are essential; this is expressed by requiring the \( n^2 \times r \) Jacobi matrix \( \hat{J} \) with elements:

\[
\hat{J}_{kl} \equiv \frac{\partial \hat{A}_k}{\partial g_l}, \quad k \equiv (i,j) = 1, \ldots, n^2; \; l = 1, \ldots, r
\]

to have rank r. (Hence \( r \leq n^2 \)).

**Notes.**
1. Postulate G4 implies immediately that the r matrices (the columns of \( \hat{J} \)):
   \( \partial \hat{A}/\partial g_1, \partial \hat{A}/\partial g_2, \ldots, \partial \hat{A}/\partial g_r \) are linearly independent. This statement
can be turned around: if these $r$ matrices are linearly independent, $J$
has rank $r$ and the $r$ parameters are essential.

2. If the field $F$ is the complex field, the Lie group is a complex Lie
group. If $F$ is the field $R$, $G$ is a real Lie group. Notice, however,
that in both cases the matrices themselves, which constitute the group,
can be complex.

3. Often different mappings

$$\mathbb{A} : V \rightarrow G \subseteq \text{AL}(n,C)$$

of different parameter spaces $V$ onto the linear Lie group $G$ can be con-
structed. These are different parametrizations of $G$.

4. The definition of a linear Lie group given above is usually for so-
called local linear Lie groups. For global Lie groups it is not always
possible to parametrize the whole group with one open set $V \subseteq F^r$, often
more (overlapping) subsets of $F^r$ are required.

**Examples.**

1. The full rotation group $SO(3)$, consisting of all orthogonal $3 \times 3$ ma-
trices with determinant $+1$, is a Lie group. The elements of $SO(3)$ rep-
resent rotations of $R^3$. Using the parametrization of Euler [17] to
describe these rotations, one finds that the matrix elements contain
only sines and cosines of the angles $\alpha$, $\beta$ and $\gamma$, and hence are analytic
functions of the parameters. By differentiating an arbitrary rotation
matrix $R$ with respect to $\alpha$, $\beta$ and $\gamma$ it is easily proved that the Euler
angles are essential, for the three matrices: $\partial R/\partial \alpha$, $\partial R/\partial \beta$ and $\partial R/\partial \gamma$
are linearly independent. The parameter space $V$ is given by $0 \leq \alpha < 2\pi$,
$0 \leq \beta < \pi$, and $0 \leq \gamma < 2\pi$.

2. Other examples of Lie groups are $GL(n)$, $U(n)$ and $SU(n)$. These groups
will be considered in more detail below.

Let us next define the Lie algebra of a Lie group. Abstract Lie
algebras have already been introduced in sec. I.2.1. One may wonder why
Lie algebras are considered in this work, because the irreducible repres-
tentations of the classical groups just mentioned will be computed by
purely "global" means; that is, we will be dealing with the groups them-
selves rather than with their Lie algebras. But Lie algebras, besides being
useful tools for obtaining representations, usually also possess physical
meaning. In particular this is true for the elements of the algebras of
Lie symmetry groups: these are constants of motion. For instance, as we will see later, the Pauli spin matrices are elements of the Lie algebra of SU(2); it is just this fact that ties electron spin to Lie group theory. So we must discuss the Lie algebra of SU(2) anyway, and it does not require much extra effort to keep the discussion fairly general. One is rewarded, not only by a much broader view of the subject, but also by a much shorter proof of the fact that GL(n) and U(n) share many of their irreps (Weyl's Unitarian Trick).

Definitions.
1. Let \( N(0) \equiv (-t_0, t_0) \) be an interval on the real line, and let \( G = \{ A'(\mathbf{g}) \} \) be a Lie group of \( n \times n \) matrices. The mapping \( \mathbf{A} \) from \( N(0) \) into \( G \):

\[
\mathbf{A} : N(0) \rightarrow V \rightarrow G
\]

is an analytic curve in \( G \), provided the components of \( \mathbf{g}(t) \in V, t \in N(0) \), are analytic functions of \( t \). Clearly the matrix elements of \( \mathbf{A}(t) = A'(\mathbf{g}(t)) \) are then also analytic in \( t \).

In the sequel only curves with \( \mathbf{A}(0) = \mathbf{E} \), i.e. curves passing through \( \mathbf{E} \), will be considered.

2. An analytic curve in \( G \) through \( \mathbf{E} \) has a tangent vector \( \mathbf{a} \) at \( \mathbf{E} \) defined by:

\[
\mathbf{a} = \left. \frac{d\mathbf{A}(t)}{dt} \right|_{t=0}
\]

Notes.
1. The name tangent vector may be somewhat confusing; a tangent vector is of course still an \( n \times n \) matrix, not a vector in the ordinary sense.
2. A tangent vector represents a linear operator on an \( n \)-dimensional space; this operator is called an infinitesimal operator of the Lie group \( G \).

Theorem.
The set \( \Lambda(G) \) of all tangent vectors at \( \mathbf{E} \) of all analytic curves through \( \mathbf{E} \) in an \( r \) parameter Lie group \( G \) is a Lie algebra of dimension \( r \).

Proof. Let \( \mathbf{a} \) be a tangent vector at \( \mathbf{E} \) of the analytic curve \( \mathbf{A}(t) \). Let \( \mathbf{b} \) be the same for \( \mathbf{B}(t) \). Choose two arbitrary scalars \( \lambda \) and \( \mu \), and take \( t \) so small that \( \mathbf{A}(\lambda t) \) and \( \mathbf{B}(\mu t) \) are defined. Because \( G \) is a Lie group, it follows that:

\[
\mathbf{C}(t) \equiv \mathbf{A}(\lambda t) \mathbf{B}(\mu t)
\]
is an analytic curve in \( G \). The tangent vector \( \mathbf{c} \) of \( \mathbf{C}(t) \) belongs by definition to \( \Lambda(G) \). Differentiation of the definition equation of \( \mathbf{C}(t) \) gives:

\[
\mathbf{c} = \lambda \mathbf{a} + \mu \mathbf{b}.
\]

Hence \( \Lambda(G) \) is a vector space.

Let \( \mathbf{A}(t) \) and \( \mathbf{B}(s) \) be analytic curves in \( G \), then:

\[
\mathbf{C}(s,t) \equiv \mathbf{B}(s) \mathbf{A}(t) \mathbf{B}(s)^{-1}
\]

is an analytic curve in \( G \), if we keep \( s \) fixed. Therefore its tangent vector:

\[
\mathbf{c}(s) \equiv \frac{d\mathbf{C}(s,t)}{dt}
\]

\[
_{t=0} = \mathbf{B}(s) \mathbf{a} \mathbf{B}(s)^{-1}
\]

belongs by definition to \( \Lambda(G) \). Also

\[
\frac{1}{s}[\mathbf{c}(s) - \mathbf{c}(0)] = \frac{1}{s} [\mathbf{B}(s) \mathbf{a} \mathbf{B}(s)^{-1} - \mathbf{B}(0) \mathbf{a} \mathbf{B}(0)^{-1}]
\]

belongs to \( \Lambda(G) \), because \( \mathbf{B}(0) = \mathbf{E} \) and \( \Lambda(G) \) is a vector space. Therefore, using \( \frac{d\mathbf{B}(s)}{ds} = -d\mathbf{B}(s)^{-1}/ds \), we find that:

\[
\lim_{s \to 0} \frac{1}{s}[\mathbf{c}(s) - \mathbf{c}(0)] = \mathbf{b} \mathbf{a} - \mathbf{a} \mathbf{b} \equiv [\mathbf{b}, \mathbf{a}]
\]

belongs also to \( \Lambda(G) \). So, \( \Lambda(G) \) is closed under Lie multiplication.

So, this theorem states that a unique Lie algebra is associated with every Lie group. Conversely, one can ask if every Lie algebra corresponds to a Lie group. Indeed this is true: if \( \Lambda \) is a matrix Lie algebra, then the multiplicative closure of the set:

\[
\{ e^{\mathbf{a}} | \mathbf{a} \in \Lambda \}, \quad e^{\mathbf{a}} \equiv \sum_{n=0}^{\infty} \frac{1}{n!} (\mathbf{a})^n
\]

is a Lie group with \( \Lambda \) as its algebra. But this correspondence is not unique, there can be other groups that also have \( \Lambda \) as their Lie algebra. All the Lie groups with the same Lie algebra are homomorphic images of another group: the universal covering group \( \tilde{G} \), which is the "largest" group with \( \Lambda \) as its Lie algebra. It goes beyond the scope of this survey to say more about this, but we will give later in this chapter an example of a Lie group and its covering group.

**Example.**

Define the following analytic curves through \( \mathbf{E} \) in \( \text{SO}(3) \):

\[
\begin{align*}
\mathbf{R}_x(t) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}, & \mathbf{R}_y(t) &= \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}, & \mathbf{R}_z(t) &= \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{align*}
\]
with corresponding tangent vectors at \( E \):

\[
\begin{align*}
I_x &= \left. \frac{dR_x}{dt} \right|_{t=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \\
I_y &= \left. \frac{dR_y}{dt} \right|_{t=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \\
I_z &= \left. \frac{dR_z}{dt} \right|_{t=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{align*}
\]

In sec. I.2.1 we have shown that \( \{ I_x \equiv iI_x, I_y \equiv iI_y, I_z \equiv iI_z \} \) spans a 3-dimensional Lie algebra. We now see that that algebra is (i times) the Lie algebra of \( \text{SO}(3) \).

It can easily be proved by expanding the following expressions in Taylor series:

\[
\begin{align*}
tI_x, & \quad tI_y, & \quad tI_z \\
e^{tI_x}, & \quad e^{tI_y}, & \quad e^{tI_z}
\end{align*}
\]

and by comparing with the Taylor expansions of \( \sin t \) and \( \cos t \), that these exponentials are another way of writing the matrices \( R_x(t), R_y(t) \) and \( R_z(t) \) respectively. So this example illustrates that by differentiation of the analytic curves at \( E \) we go from the Lie group to its algebra, and by exponentiation we go back from the algebra to the group.

We end this synopsis of Lie group theory with a few remarks on the representations of Lie groups and their algebras. Let the Lie group \( G \) have the representation \( D \) (see ref. 13, p. 186 for the definition). The set \( D(G) \) is also a Lie group, and has an algebra consisting of the tangent vectors:

\[
\bar{a} = \left. \frac{dD(g(t))}{dt} \right|_{t=0}, \quad g(t) \in G.
\]

It can be shown that this algebra is always a Lie algebra representation of \( \Lambda(G) \), the Lie algebra of \( G \), and hence every representation of the group yields a representation of the algebra. If the representation of the group is decomposable (irreducible), then the representation of the algebra is also decomposable (irreducible).

Conversely, let \( D \) be a representation of \( \Lambda(G) \). Now, the multiplicative closure of the set

\[
\{ e^{\bar{a}} | \bar{a} \in D(\Lambda(G)) \}
\]

may or may not be a single-valued representation of \( G \). The representation thus obtained can be multiple-valued and is then, strictly speaking, not a representation for \( G \). However, it will always constitute a single-valued representation for the universal covering group \( \hat{G} \) of \( G \), and therefore one often considers first \( \hat{G} \) rather than \( G \). If the representation \( D \) of \( \Lambda(G) \) is
irreducible, then so is the corresponding representation of \( \tilde{G} \). Thus, the problem of computing all irreps of \( G \) can be solved by computing all the irreps of \( \Lambda(G) \), and herewith of \( \tilde{G} \), provided we check every representation of \( \tilde{G} \) to see that it actually determines a single-valued representation of \( G \).

Note.

At the end of this chapter we will meet \( SU(2) \) as an example of a universal covering group (of \( SO(3) \)). There is a 2-1 homomorphism from \( SU(2) \) onto \( SO(3) \). This explains the existence of double-valued representations ("double groups") of \( SO(3) \) and its subgroups.

3.2. The general linear group \( GL(n) \) and the unitary group \( U(n) \)

In this section the general linear group \( GL(n) \) and its unitary subgroup \( U(n) \) are defined together with their Lie algebras.

Definition.

The general linear group \( GL(n) \) is the set of all non-singular linear operators on an \( n \)-dimensional vector space \( V_n \) defined over \( \mathbb{C} \). In the sequel this group will not be distinguished from the group consisting of all complex non-singular \( n \times n \) matrices.

Note.

\( GL(n) \) is a subset of the associative algebra \( AL(n) \), defined in sec. I.2.1, which has a basis \( \{ E^{i\ell}_{\jmath} | i, j = 1, \ldots, n \} \) introduced in sec. I.2.2. Endowing \( AL(n) \) with the Lie product:

\[
[a, b] = ab - ba, \quad a, b \in AL(n)
\]

this algebra becomes also a Lie algebra. Its Lie algebra structure constants are given by:

\[
[E^{i\ell}_{\jmath}, E^{k\ell}_{\jmath}] = \delta^{jk} E^{i\ell}_{\jmath} - \delta^{li} E^{k\ell}_{\jmath}.
\]

Theorem.

(i) \( GL(n) \) is an \( n^2 \) parameter complex Lie group.
(ii) \( AL(n) \) is the Lie algebra of \( GL(n) \).
Proof.
(i) Parametrize $GL(n)$ in the following way:

$$A = \mathbf{E} + A' \quad \text{with} \quad A' = A - \mathbf{E}.$$  

The set of $n^2$ matrix elements $\{A'_{ij}\}$ of $A'$ is a vector in $C^{n^2}$, the parameter space of $GL(n)$. Now $GL(n)$ is a Lie group, for:
G1. Set $A'_{ij} = 0$, $i$, $j = 1$, $\ldots$, $n \Rightarrow A = \mathbf{E}$.
G2. $GL(n)$ is a group in the algebraic sense.
G3. The matrix elements are linear, hence analytic, functions of the parameters.
G4. The $n^2$ parameters are essential. (If they were not, it would imply that a relation between the matrix elements of $A$ would exist, in contradiction with the definition of $GL(n)$).

(ii) Every element $a \in AL(n)$ is a tangent vector at $\overline{\mathbf{E}}$ of the analytic curve $e^{ta}$, which, of course, is also an $n \times n$ matrix. Because $e^{ta}$ has the inverse $e^{-ta}$, it belongs to $GL(n)$ (see ref. 13, p. 159 for a proof of $(e^{ta})(e^{-ta}) = \mathbf{E}$).

To introduce $U(n)$ and its Lie algebra, the following two lemmas are needed.

Lemma 1.
The set $\Lambda$ of all skew-Hermitean $n \times n$ matrices is a real Lie algebra of dimension $n^2$.

Proof. Write the complex matrix element $a_{ij}$ of $a \in \Lambda$ as:

$$a_{ij} = \text{Re}(a_{ij}) + i\text{Im}(a_{ij}).$$

It is then easy to show that every element $a \in \Lambda$ can be written as follows:

$$a = \sum_{i<j} \text{Re}(a_{ij})(E_{ij} - E_{ji}) + i \sum_{i>j} \text{Im}(a_{ij})(E_{ij} + E_{ji}) + i \sum_i \text{Im}(a_{ii})E_{ii}$$

and so the $n^2$ matrices:

$$a_{ij} \equiv \frac{1}{2}(E_{ij} - E_{ji}), \quad \text{for} \quad i < j$$

$$a_{ij} \equiv \frac{1}{2i}(E_{ij} + E_{ji}), \quad \text{for} \quad i \geq j$$

constitute an $n^2$-dimensional basis of $\Lambda$. The expansion coefficients of $a$ with respect to this basis are real; $\Lambda$ is a real vector space.

From $[a, a']^\dagger = -[a, a']$, $a, a' \in \Lambda$, follows then that the set $\Lambda$ is closed under Lie multiplication.
Note.
One can express a skew-Hermitean matrix $\mathbf{a}$ also directly in terms of $\{g^{ij}\}$. Then its components with respect to this basis obey $a^{*,\ i}_j = -a^{\ i}_j$.

Lemma 2.
Any unitary $n \times n$ matrix $\mathbf{U}$ can be written as:

$$\mathbf{U} = e^{\mathbf{a}}$$

with $\mathbf{a} \in \Lambda$.

Proof.
Ref. 18, p. 55.

Note.
The matrix $\mathbf{a}$ is not uniquely determined by $\mathbf{U}$.

In the following theorem no distinction will be made between linear operators and matrices.

Theorem.
(i) The set $\mathbf{U}(n)$, which consists of all unitary operators on $V_n$, is an $n^2$ parameter real Lie group.

(ii) The Lie algebra $\Lambda(\mathbf{U}(n))$ of $\mathbf{U}(n)$ is the algebra $\Lambda$ of all skew-Hermitean $n \times n$ matrices.

Proof.
(i) According to the lemmas 1 and 2, every element $\mathbf{U} \in \mathbf{U}(n)$ can be expressed in $n^2$ real parameters $x_{ij}$:

$$\mathbf{U} = \exp\left( \sum_{i,j} x_{ij} a^{ij} \right), \quad x_{ij} \in \mathbb{R}, \quad a^{ij} \in \Lambda.$$

Now:

G1. Set $x_{ij} = 0$, $i, j = 1, \ldots, n \Rightarrow \mathbf{U} = \mathbf{E}$.

G2. $\mathbf{U}(n)$ is a group in the algebraic sense.

G3. Expansion of the exponential function shows that the elements of $\mathbf{U}$ are analytic functions of the $x_{ij}$.

G4. The $n^2$ matrices:

$$\frac{\partial \mathbf{U}}{\partial x_{ij}} = a^{ij} \mathbf{U}, \quad i, j = 1, \ldots, n$$

are linearly independent, for suppose that:

$$\sum_{i,j} C_{ij} \frac{\partial \mathbf{U}}{\partial x_{ij}} = \left( \sum_{i,j} C_{ij} a^{ij} \right) \mathbf{U} = 0.$$
Multiply with $U^{-1}$ and it follows that $C_{i,j} = 0$, $i, j = 1, \ldots, n$, because the matrices $a_{i,j}^i$ are linearly independent. (See e.g. ref. 13, p. 158 for the differentiation of exponential matrix functions).

(ii) a. An analytic curve $U(t)$ through $E$ obeys:

$$U(t)^+ U(t) = E$$

and so, using $U(0) = E$:

$$\left[ \frac{dU(t)}{dt} \right]_{t=0}^+ + \left[ \frac{dU(t)}{dt} \right]_{t=0}^- = 0.$$

Hence the tangent vector of $U(t)$ at $E$ is skew-Hermitean.

So: $\Lambda(U(n)) \subseteq \Lambda$.

b. The arbitrary matrix $a \in \Lambda$ is a tangent vector of the curve $e^{ta}$ in $U(n)$ and so $\Lambda \subseteq \Lambda(U(n))$.

Conclusion: $\Lambda \equiv \Lambda(U(n))$.

Notes.

1. In ref. 19, chapter 2 it is shown that any $n \times n$ unitary matrix can be factorized into a product of $\frac{1}{2} n(n-1)$ complex $2 \times 2$ rotations, each containing two angles, and a diagonal matrix containing $n$ phase factors. Of the $n^2$ angles thus arising $n$ are in the closed and bounded interval $[-\pi, +\pi]$, and $n(n-1)$ are in the closed and bounded interval $[-\frac{3\pi}{2}, +\frac{3\pi}{2}]$. Hence, by the Heine-Borel theorem [20, p. 35], $U(n)$ is compact. Because integrals of analytic functions over a compact set are well-defined, many of the results known from the representation theory of finite groups also hold for compact groups. One has only to replace the summations over group elements by integrations over the parameter space of the compact group. In particular it must be noted that Maschke's theorem holds for compact groups, and so we are assured that all analytic representations of $U(n)$ are decomposable.

2. The basis $\{a_{i,j}^i\}$ of the real Lie algebra $\Lambda(U(n))$ is obtained by a non-singular transformation of the basis $\{E_{i,j}^i\}$ of the complex Lie algebra $\text{AL}(n)$, hence $\{a_{i,j}^i\}$ can also serve as a basis for $\text{AL}(n)$. If one extends the field $\mathbb{R}$, over which $\Lambda(U(n))$ has been defined, to the complex field $\mathbb{C}$, then $\Lambda(U(n))$ becomes $\text{AL}(n)$. So, $\text{AL}(n)$ is the complexification of $\Lambda(U(n))$. (We have met this case already in an example in sec. I.2.1).
3.3. Connection between the irreps of GL(n) and U(n)

The group GL(n) with its unbounded parameter space \( \mathbb{C}^{n^2} \) is obviously noncompact, and yet many of its representations are decomposable. This is no coincidence, but follows from the fact that the complex Lie group GL(n) shares its irreps with U(n); that is: the irreps of GL(n) do not decompose upon subduction to U(n), and conversely any irrep of U(n) extends to an irrep of GL(n).

Before discussing this further we prove a theorem establishing the same fact for the corresponding Lie algebras. We need the following lemma in the proof.

**Lemma.**

Let \( A \) be a skew-Hermitean \( n \times n \) matrix and \( B \) an arbitrary \( n \times n \) matrix. If \( \text{Tr}(A B) = 0 \) it follows that \( B = 0 \).

**Proof.** Ref. 8, p. 165.

**Theorem.**

(i) An irrep of the Lie algebra \( \mathfrak{a}(U(n)) \) stays irreducible under restriction to \( \mathfrak{a}(U(n)) \).

(ii) An irreducible representation of \( \Lambda(U(n)) \) stays irreducible under extension to \( \mathfrak{a}(U(n)) \), the complexification of \( \Lambda(U(n)) \).

**Proof.**

(i) Let \( D \) be an irreducible linear matrix representation of \( \mathfrak{a}(U(n)) \), and assume that \( D \), possibly after a similarity transformation, reduces under the restriction to \( \Lambda(U(n)) \). This implies that \( D(a) \) has one or more matrix elements \( D(a)_{pq} \) that are equal to zero for all \( a \in \Lambda(U(n)) \). Expand \( a \) in the basis \( \{ E_{ij} \} \):

\[
a = \sum_{i,j=1}^{n} a_{ij} E_{ij}
\]

and so:

\[
D(a)_{pq} = \sum_{i,j=1}^{n} a_{ij} D(E_{ij})_{pq} = 0
\]

with the condition \( a_{ij}^* = -a_{ji} \). Writing \( D(E_{ij})_{pq} = b_{ji} \), we get:

\[
\sum_{i,j=1}^{n} a_{ij} b_{ji} = \text{Tr}(a b) = 0
\]

and from the lemma \( b = 0 \). This means that the matrix elements \( D(E_{ij})_{pq} \)
are zero for the whole basis of $\Lambda(n)$, which in turn implies that $D$ has the same blocked form for $\Lambda(n)$ as for $\Lambda(U(n))$. Contradiction!

(ii) Let $D$ be an irreducible linear matrix representation of $\Lambda(U(n))$. We have shown earlier that $a \in \Lambda(U(n))$ can be written thus:

$$a = \sum_{i,j} a_{ij} a_{ij}^*, \quad a_{ij} \in \mathbb{R}$$

and so:

$$D(a) = \sum_{i,j} a_{ij} D(a_{ij}^*), \quad a_{ij} \in \mathbb{R}.$$

Allowing the coefficients $\{a_{ij}\}$ to become complex, the representation $D$ extends to a representation of $\Lambda(n)$. If $D$ became reducible upon this extension, it would be so for all elements of $\Lambda(n)$, in particular for the elements of $\Lambda(U(n))$. Contradiction!

Returning to the Lie groups, we note that not all representations of $\Lambda(n)$ and $\Lambda(U(n))$ yield after exponentiation single-valued representations of their respective Lie groups. It can, however, be proved [13, p. 330] that if a representation $\rho$ of $\Lambda(n)$ does yield a single-valued representation $D$ of $GL(n)$, then the restriction $\rho'$ of $\rho$ to $\Lambda(U(n))$ will give upon exponentiation a single-valued representation $D'$ of $U(n)$. It is further almost trivial to demonstrate that $D'$ is then the restriction of $D$ to $U(n)$. Now the converse holds also: if $\rho'$ of $\Lambda(U(n))$ yields the single-valued representation $D'$ of $U(n)$, then $D'$ extends to $D$ of $GL(n)$; $D$ corresponds to $\rho$, which in turn is the complex extension of $\rho'$. Since we have seen in sec. I.3.1 that all irreducible Lie group representations arise from irreducible Lie algebra representations, it follows from these considerations and the theorem above that any irrep $D$ of $GL(n)$ stays irreducible under subduction to $U(n)$ and conversely that any irrep $D'$ of $U(n)$ extends to an irrep of $GL(n)$. So, the irreps of the non-compact complex Lie group $GL(n)$ have all the usual properties of compact Lie groups.

Notes.

1. This procedure of looking at a unitary and hence compact subgroup of a non-compact Lie group rather than at the Lie group itself, is called by Weyl [7] the Unitarian Trick.

2. One can consider $GL(n)$ also as a $2n^2$ parameter real Lie group, with a $2n^2$-dimensional real Lie algebra. Proceeding this way one can obtain non-decomposable irreps of $GL(n)$. These do not have any bearing on this work, because we will always consider $GL(n)$ as a complex Lie group.
3.4. The special unitary group SU(n)

Lemma 1.
The set $\Lambda$ of traceless skew-Hermitean matrices is an $(n^2-1)$-dimensional real Lie algebra.

Proof. By definition $\Lambda$ is a subset of $\Lambda(U(n))$, which has the basis $\{a^{ij}\}$ introduced in sec. I.3.2. So any $\underline{a} \in \Lambda$ can be written thus:

$$\underline{a} = \sum_{i,j} a^{ij} \underline{a}^{ij}, \quad a^{ij} \in R.$$  

The condition $\text{Tr}(\underline{a}) = 0$ then reads $\sum_i a_{ii} = 0$. This linear relation on the components of the elements in $\Lambda$ defines an $(n^2-1)$-dimensional linear subspace in $\Lambda$, which can easily be shown to be closed under Lie multiplication.

Lemma 2.
Any unitary matrix $\underline{U}$, with $\det(\underline{U}) = +1$, can be written as:

$$\underline{U} = e^\underline{a}$$

with $\underline{a} \in \Lambda$.

Proof. Write $\underline{U} = e^\underline{a}$ (lemma 2, sec. I.3.2) and from $\det(\underline{U}) = e^{\text{Tr}(\underline{a})}$ (e.g. ref. 13, p. 156) follows $\text{Tr}(\underline{a}) = 0$.

Theorem:
(i) The set: $SU(n) = \{\underline{U} \mid \underline{U} \in U(n), \det(\underline{U}) = +1\}$ is an $(n^2-1)$ parameter real Lie group.

(ii) The Lie algebra $\Lambda(SU(n))$ of $SU(n)$ is the algebra $\Lambda$ of all traceless skew-Hermitean $n \times n$ matrices.

Proof.
(i) Define the following basis for $\Lambda$ in terms of the basis $\{a^{ij}\}$ of $\Lambda(U(n))$:

$$s^{ij} = a^{ij} - \delta^{ij}/n \sum_{i=1}^{n} a_{ii}, \quad i, j = 1, \ldots, n.$$  

Clearly:

$$\sum_{i=1}^{n} s_{ii} = 0$$

and so the basis is $(n^2-1)$-dimensional.

Write: $\underline{U} = \exp[\sum_{i,j} x_{ij}s^{ij}], \quad x_{ij} \in R$.

Check off the Lie group axioms G1 through G4:
G1, G2, G3 as for \( U(n) \).

G4. The \( n^2 \) matrices:

\[
\frac{\partial U}{\partial x_{ij}} = s_{ij} U, \quad i, j = 1, \ldots, n
\]

obey one linear relation: \( \sum_{i=1}^{n} s_{ii} U = 0 \), and so there are \((n^2 - 1)\) linearly independent ones among them: \( SU(n) \) has \( n^2 - 1 \) essential parameters.

(ii) a. From lemma 2 it follows immediately that any curve \( U(t) \) through \( U \) in \( SU(n) \) can be written as:

\[
U(t) = e^{a(t)} , \quad a(t) \in \Lambda \text{ and we require } a(0) = 0.
\]

From this:

\[
\left[ \frac{dU(t)}{dt} \right]_{t=0} = \left[ U(t) \frac{da(t)}{dt} \right]_{t=0} = \left[ \frac{da(t)}{dt} \right]_{t=0}
\]

and since:

\[
\text{Tr} \left[ \frac{da(t)}{dt} \right]_{t=0} = \frac{d}{dt} \text{Tr}(a(t)) \bigg|_{t=0} = \frac{d}{dt} (0) = 0
\]

it follows that:

\[
\left[ \frac{dU(t)}{dt} \right]_{t=0} \in \Lambda \Rightarrow \Lambda(SU(n)) \subset \Lambda.
\]

b. One proves in the same way as for \( U(n) \): \( \Lambda \subset \Lambda(SU(n)) \).

Example.

We will outline some properties of \( SU(2) \), a group which is of great importance for electron systems. (This will be discussed in ch. I.6). We have just met a basis for \( SU(n) \); for the case \( n = 2 \) this basis consists of the following three matrices:

\[
\begin{align*}
\mathbf{s}^{11} &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \equiv h_{1} \sigma_{z}, \\
\mathbf{s}^{21} &= \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix} \equiv h_{1} \sigma_{x}, \\
\mathbf{s}^{12} &= \begin{pmatrix} 0 & 1 \\ -i & 0 \end{pmatrix} \equiv h_{1} \sigma_{y}.
\end{align*}
\]

(The fourth matrix \( \mathbf{s}^{22} \) is equal to \( -\mathbf{s}^{11} \)). The matrices \( h_{i} \), \( i = x, y, z \), are the well-known Pauli spin matrices [21, p. 545]. They obey the commutation relations:

\[
[h_{i}, h_{j}] = i \sum_{k=1}^{3} \epsilon_{ijk} h_{k}, \quad i, j, k = x, y, z = 1, 2, 3.
\]

The constants \( \epsilon_{ijk} \) are the Levi-Civita symbols, defined in the discussion of the Lie algebra \( \Lambda(\text{SO}(3)) \) of \( \text{SO}(3) \) (sec. I.2.1 and sec. I.3.1). Turning
back to these discussions, one sees that $\Lambda(\text{SU}(2))$ and $\Lambda(\text{SO}(3))$ are Lie algebra isomorphic. (Note the factors $\frac{1}{2}$ in front of the Pauli matrices. As we will see shortly, they make SU(2) a "double" rotation group).

It has been pointed out in sec. I.3.1 that a Lie algebra isomorphism does not necessarily carry through for the associated groups, that is, SU(2) is not necessarily isomorphic with SO(3). To explore this point further we define a 1-1 mapping between $\Lambda(\text{SU}(2))$ and $\Lambda(\text{SO}(3))$ by writing an element of $\Lambda(\text{SO}(3))$ as $\mathbf{n} \cdot \mathbf{L}$, and the corresponding element of $\Lambda(\text{SU}(2))$ as $\mathbf{n} \cdot \hat{\mathbf{a}}/2$, with in both cases the same vector $\mathbf{n} \in \mathbb{R}^3$. Almost exactly as for SU(2), it can be shown that every element $\mathbf{R}(\mathbf{n}) \in \text{SO}(3)$ can be written thus:

$$\mathbf{R}(\mathbf{n}) = e^{\mathbf{n} \cdot \mathbf{L}}, \quad \mathbf{n} \cdot \mathbf{L} \in \Lambda(\text{SO}(3)).$$

Every element $\mathbf{U}(\mathbf{n}) \in \text{SU}(2)$ can be written in the form

$$\mathbf{U}(\mathbf{n}) = e^{\mathbf{n} \cdot \hat{\mathbf{a}}/2}, \quad \mathbf{n} \cdot \hat{\mathbf{a}}/2 \in \Lambda(\text{SU}(2)).$$

Let us now see if the map: $\mathbf{U}(\mathbf{n}) \rightarrow \mathbf{R}(\mathbf{n})$ of SU(2) onto SO(3) is possibly an isomorphism. To that end we expand, and after some manipulation we get:

$$\mathbf{R}(\mathbf{n}) = \mathbf{E} + (\mathbf{i}\mathbf{n} \cdot \mathbf{L}) (\sin n) + (\mathbf{i}\mathbf{n} \cdot \mathbf{L})^2 (1 - \cos n)$$

$$\mathbf{U}(\mathbf{n}) = \mathbf{E}(\cos n/2) + (\mathbf{i}\mathbf{n} \cdot \hat{\mathbf{a}}/2)(\sin n/2),$$

where $\mathbf{E} = \mathbf{n} \mathbf{n}$ and $|\mathbf{n}| = 1$ (see for a derivation of these relations ref. 22, p. 12 and p. 25 resp.). It is easy to show [23, p. 8] that the matrix $\mathbf{R}(\mathbf{n})$ represents a rotation around the axis $\mathbf{n}$ (= unit vector along $\mathbf{n}$) over an angle $n = |\mathbf{n}|$.

The first relation shows that SO(3) is covered completely when the range of $\mathbf{n}$ is restricted to a sphere of radius $\pi$ around the origin in $\mathbb{R}^3$; each point $\mathbf{n}$ of the sphere represents a rotation $\mathbf{R}(\mathbf{n})$ uniquely, except when $n = \pi$, for the two points $\pi \mathbf{n}$ and $-\pi \mathbf{n}$ represent the same rotation.

The second relation shows that SU(2) is covered completely when the range of $\mathbf{n}$ is restricted to a sphere of radius $2\pi$ around the origin in $\mathbb{R}^3$; each point $\mathbf{n}$ of the sphere represents a 'rotation' $\mathbf{U}(\mathbf{n})$ uniquely, except that the whole surface of the sphere ($n = 2\pi$) represents $\mathbf{U} = -\mathbf{E}$.

We note in particular that as we go over the large sphere with radius $2\pi$, we cover SU(2) once and SO(3) twice in such a way that the two points $\pi \mathbf{n}$ and $-(2\pi - n) \mathbf{n}$ correspond to the same rotation of SO(3) and to elements of SU(2) of opposite sign.
Summarizing, we have found that two different elements of SU(2) map onto the same element of SO(3):

\[ \frac{U(n\hbar)}{U(-(2\pi - n)\hbar)} \rightarrow R(n\hbar), \quad 0 \leq n < 2\pi. \]

From the Campbell-Baker-Hausdorff formula [13, p. 161] and the isomorphism of \( \Lambda(SO(3)) \) and \( \Lambda(SU(2)) \) follows that this map is a group homomorphism; SU(2) is a double-valued representation of SO(3).

Now, it can be proved by topological arguments that SU(2) is the "largest" group with a Lie algebra which is isomorphic to \( \Lambda(SO(3)) \), and so SU(2) is the universal covering group of SO(3).
CHAPTER I.4. TENSOR REPRESENTATIONS OF GL(n) AND S_N

The space $V_n \otimes^N$, which is the N-fold tensorial product of $V_n$ with itself, carries representations of the permutation group $S_N$ on the one hand and of the general linear group GL(n) on the other. In this chapter it will be shown that the representations of these two groups are closely related: decomposition of $V_n \otimes^N$ under the one group induces simultaneous decomposition under the other.

This remarkable fact was discovered by Schur in 1901 and later explored in depth by Weyl [6, 7], who also recognized its importance for quantum mechanics. Because Weyl's expositions make extremely difficult reading, and because Boerner's exposition of the same subject [8] is not much more transparent, some of the more important proofs will be presented in the hope that I have succeeded in clarifying the derivations for non-mathematical readers. At the same time the theory is extended to include also the Lie algebra of GL(n), thus leading up to a rigorous proof of the fact that the Casimir operators of GL(n) are represented by linear combinations of permutation operators. This relation between Casimir operators and permutation operators has been noted before [24, 25], but it seems that no general proof has been given earlier.

4.1. Tensor products of vector spaces

Definition.

Let $U$ be an n-dimensional vector space with basis $\{u_i\}$ and $V$ an m-dimensional vector space with basis $\{v_j\}$. Associate with each element $(u_i, v_j)$ in the Cartesian product $U \times V$ an element $u_i \otimes v_j$, and let $W$ be the nm-dimensional linear closure consisting of all formal sums:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} u_i \otimes v_j, \quad c_{ij} \in \mathbb{C}.$$ 

If we further postulate:

T1. $(cu_i) \otimes v_j = u_i \otimes (cv_j) = c(u_i \otimes v_j), \quad c \in \mathbb{C}$

T2 a. $u_k \otimes (v_i + v_j) = u_k \otimes v_i + u_k \otimes v_j$

T2 b. $(u_i + u_j) \otimes v_k = u_i \otimes v_k + u_j \otimes v_k$

for all basis vectors $u_i$ of $U$ and $v_j$ of $V$, then $W$ is a tensor product space, commonly denoted by $U \otimes V$. 

37
Notes.
1. From the postulates follows immediately:
   \[ 0 \otimes v = u \otimes 0 = 0 \otimes 0 \in U \otimes V, \quad \forall u \in U, \forall v \in V \]
and \( u \otimes v = 0 \) implies either: \( u = 0 \) or \( v = 0 \) or \( u = v = 0 \). It then follows that \( U \otimes V \) is nm-dimensional.
2. A tensor product \( u \otimes v \in U \otimes V \) is not commutative.
3. Nothing has been required of the vector spaces \( U \) and \( V \) except finite dimensionality, and so these spaces can be tensor product spaces themselves. Postulating:
   T3. \( (U \otimes V) \otimes W = U \otimes (V \otimes W) \)
   the tensor product of vector spaces can be extended to an arbitrary number of factors.
4. The \( N \)-fold tensor product \( V \otimes V \otimes \ldots \otimes V \) of the vector space \( V \) with itself will in the sequel be denoted by \( V \otimes^N \).
5. Elements of \( V \otimes^N \) are called tensors.
6. The definition of a tensor product space is formulated in terms of bases of the vector spaces constituting the product. (See e.g. ref. 12, p. 85 for a basis free definition, which has the advantage of holding for infinite dimensional spaces as well. This latter definition will tacitly be assumed in ch. I.6, where tensor products of infinite dimensional spaces are considered).

Strictly speaking, it must now be proved that a tensor product space is independent of the choice of bases in the different vector spaces. Since this proof is a trivial generalization of the following theorem it will be forgone. Summation convention will be used in the rest of this section.

Theorem.
Let \( \{v_i\} \) and \( \{v'_i\} \) be bases of \( V_n \) connected via:
\[ v'_j = a^i_j v_i. \]
Then the corresponding bases of \( V_n \otimes^N \) satisfy the following transformation equation:
\[ v'_1 \otimes v'_2 \otimes v'_3 \otimes \ldots \otimes v'_N = a^i_{j_1} a^i_{j_2} a^i_{j_3} \ldots a^i_{j_N} v_{i_1} \otimes v_{i_2} \otimes \ldots \otimes v_{i_N}. \]
Proof. Direct consequence of the postulates.
Note.

If \( a_{ij} \) is the \((i,j)\)-element of \( A \), then \( a_{i_1 j_1} a_{i_2 j_2} \ldots a_{i_N j_N} \) is the 
\((i_1, i_2, \ldots, i_N ; j_1, j_2, \ldots, j_N)\)-element of the Kronecker product matrix 
\( [A]^N \equiv A \otimes A \otimes A \otimes \ldots \otimes A \), (see sec. I.2.3 for the definition).

Notation.

1. In the sequel we will often write \( E_I \) instead of \( v_{i_1} \otimes v_{i_2} \otimes \ldots \otimes v_{i_N} \) for a basis vector of \( V_n \otimes^N \); \( I \) then stands for the index set 
\( \{i_1, i_2, \ldots, i_N\} \). The theorem just stated reads in this notation:
\[
 v_j = a_{ij} v_i \Rightarrow E_J = A_I^J E_I,
\]
where \( A_I^J \) is the \((I;J)\)-element of \([A]^N\).

2. A tensor \( T \in V_n \otimes^N \) can be expanded with respect to the basis \( \{E_I\} \):
\[
 T = E_I t^I.
\]
Here \( t^I \equiv t_1^i t_2^{i_2} \ldots t_N^{i_N} \) is the component of \( T \) along \( E_I \).

Theorem.

Let \( \{E'_J\} \) and \( \{E_I\} \) be two different bases of \( V_n \otimes^N \) connected via:
\[
 E'_J = A_J^I E_I
\]
and let \( T \) have the components \( t_{IJ} \) and \( t^I \) with respect to the basis \( \{E'_J\} \) and \( \{E_I\} \) respectively. Then
\[
 t^I = A_I^J t_{IJ}.
\]

Proof. \( T = E_I t^I = E'_J t_{IJ} = E_I A_I^J t_{IJ} \Rightarrow t^I = A_I^J t_{IJ} \).

Notes.

1. An array consisting of \( n^N \) scalars \( t^I \) can be taken to represent a tensor 
in \( V_n \otimes^N \) if it behaves under a change of basis in \( V_n \) as in the theorem 
above. This leads to the definition of a tensor most commonly found in 
books on vector and tensor analysis.

2. Tensors are important for this work, because orbital products are but 
an example of tensors. It is common in quantum chemistry to denote 
tensors of this kind by particle labels rather than by tensor product 
symbols. Keeping track of particle labels one can commute factors in a
tensor product. For instance \( \phi(1) \psi(2) \) and \( \psi(2) \phi(1) \) stand both for the same tensor \( \phi \otimes \psi \).

4.2. The space \( V_n \otimes^N \) as a carrier space for representations of GL(n)

Definitions.
1. Let \( \alpha \) and \( \beta \) be linear operators on \( V_n \). The tensor product \( \alpha \otimes \beta \) is the linear operator on \( V_n \otimes^2 \) defined by:
   \[
   \alpha \otimes \beta (u \otimes v) = \alpha(u) \otimes \beta(v), \quad u, v \in V_n.
   \]
   If \( \alpha \) has the matrix \( A \) and \( \beta \) has the matrix \( B \), then \( \alpha \otimes \beta \) has the matrix \( A \otimes B \).
2. Tensor products of operators multiply thus:
   \[
   (\alpha \otimes \beta)(\alpha' \otimes \beta') = (\alpha \alpha')(\otimes) (\beta \beta')
   \]
   (c.f. the multiplication of Kronecker product matrices).
3. Tensor products of operators can easily be defined for more than two factors, postulating associativity.

Theorem.
The set of linear operators on \( V_n \otimes^N \):

\[
GL(n) \otimes^N = \{ \alpha' \otimes \alpha'' \otimes \ldots \otimes \alpha^{(N)} \mid \alpha', \alpha'', \ldots, \alpha^{(N)} \in GL(n) \}
\]
is a group: the outer direct product group of GL(n).

Proof. Multiplication of linear operators is associative. The identity is \( e \otimes e \otimes \ldots \otimes e \), \( e \in GL(n) \). The inverse of \( \alpha' \otimes \alpha'' \otimes \ldots \otimes \alpha^{(N)} \) is

\[
(a')^{-1} \otimes (a'')^{-1} \otimes \ldots \otimes (a^{(N)})^{-1}.
\]

Theorem.
(i) The set of linear operators on \( V_n \otimes^N \):

\[
[GL(n)]^N = \{ \alpha \otimes \alpha \otimes \ldots \otimes \alpha \mid \alpha \in GL(n) \}
\]
is a group: the inner direct product group of GL(n).
(ii) Let \( C_N \) be the cyclic group of order \( N \), then:

\[
GL(n)/C_N \cong [GL(n)]^N.
\]

Proof.
(i) In the same way as for the outer direct product of GL(n).
(ii) Define the mapping \( \tau : GL(n) \to [GL(n)]^N \) by
\[ \tau(\alpha) = \alpha \otimes \alpha \otimes \ldots \otimes \alpha, \quad \alpha \in \text{GL}(n), \]

which is readily seen to be a group homomorphism and an onto mapping.

The kernel of \( \tau \) consists of the elements \( \alpha \) with

\[ \tau(\alpha) \equiv \alpha \otimes \alpha \otimes \ldots \otimes \alpha = e \otimes e \otimes \ldots \otimes e. \]

From this follows: \( \alpha = x e \) with \( x^N = 1 \). So the kernel of \( \tau \) is the set:

\[ C^N_N = \{ x_k e | x_k = e^{(2\pi i/N)k}, \ e \in \text{GL}(n) \}, \]

which is a cyclic subgroup of \( \text{GL}(n) \). By theorem 15 of ref. 4, it follows that \( \text{GL}(n)/C^N_N \cong [\text{GL}(n)]^N \).

We have now found that the inner direct product group \([\text{GL}(n)]^N\) represents \( \text{GL}(n) \) on \( V^N_n \otimes \ldots \otimes V^N_n \). This representation is non-faithful. In the following example we will furthermore see that it is a reducible one.

Example.

Two well-known processes often applied to second rank tensors are symmetrization and contraction. Both procedures have a profound group theoretical meaning: the decomposition of a second rank tensor into a symmetric and an antisymmetric component adapts the tensor to \( \text{GL}(n) \), i.e. decomposes it into components belonging to irreducible carrier spaces of \( \text{GL}(n) \). Contraction of indices gives a further adaptation of the tensor to the orthogonal group \( \text{O}(n) \subseteq \text{GL}(n) \). For more about the latter point the reader is referred to ref. 26, sec. 10-5 to 10-7.

We will presently show that the representation of \( \text{GL}(n) \) on \( V^N_n \otimes \ldots \otimes V^N_n \) can be reduced by decomposing this tensor space into a direct sum of a symmetric and an antisymmetric subspace. Define a new basis of \( V^N_n \otimes \ldots \otimes V^N_n \) by:

\[ w^+_{ij} = v_i \otimes v_j + v_j \otimes v_i, \quad i \leq j = 1, \ldots, n \]

\[ w^-_{ij} = v_i \otimes v_j - v_j \otimes v_i, \quad i < j = 1, \ldots, n. \]

Now:

\[ \alpha \otimes \omega(w^+_{ij}) = a^k_{i j} v_k \otimes v_\ell + a^\ell_{j i} v_\ell \otimes v_k \]

\[ = \{ a^k_{i j} v_k \otimes v_\ell + a^\ell_{j i} v_\ell \otimes v_k \}
\]

\[ = \{ a^k_{i j} v_k \otimes v_\ell + a^\ell_{j i} v_\ell \otimes v_k \}
\]

\[ = \{ a^k_{i j} + a^\ell_{i j} \} v_k \otimes v_\ell + \{ a^k_{j i} + a^\ell_{i j} \} v_\ell \otimes v_k
\]

\[ = \{ a^k_{i j} + a^\ell_{i j} \} (v_k \otimes v_\ell + v_\ell \otimes v_k)
\]

\[ = T^+_{ij} w^+_{kl}. \]
So, the \( \binom{1}{n}(n+1) \)-dimensional subspace of \( \mathbb{V}_n \otimes^2 \) spanned by the symmetric tensors \( w_{ij}^+ \) is invariant under \( \alpha \otimes \alpha \). The matrix \( T^+(\alpha)_{ij}^{kl} = \tfrac{1}{2} \left[ a_{i}^{k} a_{j}^{l} + a_{j}^{k} a_{i}^{l} \right] \) is an example of an irreducible tensor representation of \( \text{GL}(n) \). In chapter 1.5 general expressions for tensor representations will be derived.

For the antisymmetric tensors one gets similarly:

\[
\alpha \otimes \alpha (w_{ij}^-) = \tfrac{1}{2} \left[ a_{i}^{k} a_{j}^{l} - a_{j}^{k} a_{i}^{l} \right] (v_{k} \otimes v_{l} - v_{l} \otimes v_{k}) = T^- (\alpha)_{ij}^{kl} w_{kl}^-
\]

and so the \( \frac{1}{2}n(n-1) \)-dimensional subspace of \( \mathbb{V}_n \otimes^2 \) spanned by the antisymmetric tensors \( w_{ij}^- \) is also invariant under \( \alpha \otimes \alpha \).

The main purpose of this chapter is twofold: the processes of symmetrization and antisymmetrization will be generalized to tensors of arbitrary rank \( N \), and the representations of \( \text{GL}(n) \) herewith obtained will be shown to be irreducible.

### 4.3. The space \( \mathbb{V}_n \otimes^N \) as a carrier space for representations of the Lie algebras of the classical groups

Let \( G \) be a classical group. By definition \( G \) consists of linear operators on \( \mathbb{V}_n \). Applying the argument of the foregoing section, one sees that the inner direct product group \( [G]^N \) forms a (possibly non-faithful) representation of \( G \), carried by \( \mathbb{V}_n \otimes^N \). The Lie algebra \( \Lambda([G]^N) \) of \( [G]^N \) consists also of linear operators on \( \mathbb{V}_n \otimes^N \). We will show that this Lie algebra is a faithful representation of \( \Lambda(G) \), the Lie algebra of \( G \).

**Definition.**

Let \( \mathbf{a} \) be an arbitrary \( n \times n \) matrix and \( \mathbf{E} \) the \( n \times n \) unit matrix, then we define:

\[
\mathbf{a}(k) \equiv \mathbf{E} \otimes \mathbf{E} \otimes \ldots \otimes \mathbf{E} \otimes \mathbf{a} \otimes \mathbf{E} \otimes \ldots \otimes \mathbf{E}
\]

where the matrix on the right hand side is an \( N \)-fold Kronecker product matrix having the matrix \( \mathbf{a} \) as the \( k \)-th factor and unit matrices as the other factors.

**Theorem.**

(i) The Lie algebra \( \Lambda([G]^N) \) consists of all possible elements of the form:

\[
\sum_{k=1}^{N} \mathbf{a}(k), \quad \mathbf{a} \in \Lambda(G).
\]
(ii) The mapping \( \sum_{k=1}^{N} a(k) \) of \( \Lambda(G) \) onto \( \Lambda([G]^N) \) is a Lie algebra isomorphism.

**Proof.**

(i) Let \( \underline{a} \) be the tangent vector at \( \underline{E} \) of the analytic curve \( \underline{A}(t) \) in \( G \), so \( \underline{a} \) belongs to \( \Lambda(G) \). Now \( \underline{A}(t) \otimes \underline{A}(t) \otimes \ldots \otimes \underline{A}(t) \) (\( N \) factors) is obviously an analytic curve in \([G]^N\) with tangent vector:

\[
\left( \frac{d}{dt}[\underline{A}(t) \otimes \underline{A}(t) \otimes \ldots \otimes \underline{A}(t)] \right)_{t=0} = \sum_{k=1}^{N} \underline{a} \otimes \underline{E} \otimes \underline{a} \otimes \ldots \otimes \underline{a} = \sum_{k=1}^{n} \underline{a}(k).
\]

This follows immediately by application of the chain rule, which can easily be asserted for Kronecker product matrices.

(ii) The mapping of \( \Lambda(G) \) onto \( \Lambda([G]^N) \) is clearly one-to-one; it is also easy to show that it is linear and preserves commutator brackets (Lie products).

Subsequently, an example of the Lie algebra of an inner direct product group will be presented. In this example we will need the following lemma.

**Lemma.**

Let \( a, b, \ldots, z \) be \( m \) arbitrary \( n \times n \) matrices. Then:

\[
e^{-a} \otimes e^{-b} \otimes \ldots \otimes e^{-z} = \exp[a(1) + b(2) + \ldots + z(m)].
\]

**Proof.**

\[
e^{-a} \otimes e^{-b} = \sum_{i,j} \frac{1}{i!j!} (a)_i \otimes (b)_j = \sum_{i,j} \frac{1}{i!j!} (a \otimes \underline{E})_i (\underline{E} \otimes b)_j = e^{-a} \otimes e^{-b} = \exp[a \otimes \underline{E} + \underline{E} \otimes b] = \exp[a(1) + b(2)] .
\]

By induction one extends this result to \( m \) factors.

**Example.**

An element of the classical group \( SU(2) \) can be written as:

\[
\underline{U}(n) = e^{in \cdot \underline{S}}
\]

(sec. I.3.4; here we employ the one-electron spin operator \( \underline{S} \) rather than its matrix \( \underline{\sigma}/2 \)).

An element of \([SU(2)]^N\) has the form:

\[
[U(n)]^N = U(n) \otimes U(n) \otimes \ldots \otimes U(n)
\]

and from the lemma follows:

\[
[U(n)]^N = \exp[in \cdot \underline{S}]
\]
with:

\[ \vec{S} = \sum_{k=1}^{N} \vec{s}(k). \]

The components \( iS_x, iS_y, iS_z \) of \( \vec{s} \) span clearly the Lie algebra \( \mathfrak{sl}(2,\mathbb{C})^N \), which is isomorphic to \( \Lambda(SU(2))^N \). Quantum mechanically \( \vec{S} \cdot \vec{S} \) is the observable representing the spin of an N-electron system.

### 4.4. The space \( V_{\mathbb{C}} \otimes^N \) as a carrier space for representations of \( CS_N \)

The tensor space \( V_{\mathbb{C}} \otimes^N \) carries in a very natural way representations of the permutation group \( S_N \) and its group algebra \( CS_N \). However, different conventions may be chosen to define the action of permutation operators on tensors and since erroneous results are obtained when not strictly adhering to a convention once chosen, we must define unambiguously how a permutation acts on a tensor.

**Definition.**

A permutation \( \pi \in S_N \) acts on a set \( A \) of \( N \) objects, which are not necessarily different. The positions (not the objects!) in \( A \) are numbered from 1 to \( N \). The action of

\[
\pi = \begin{pmatrix}
1 & 2 & \ldots & k & \ldots & N \\
j_1 & j_2 & \ldots & j_k & \ldots & j_N
\end{pmatrix}
\]

on \( A \) is described by the rule: Put object standing in position \( k \) into position \( j_k \), for \( k = 1, 2, \ldots, N \) successively.

**Example.**

From here on the well-known cycle notation will be used to denote permutations (see ref. 26, p. 13 for a definition of this notation).

Let \( A \) be the set \{a b c d b\} and number its positions, counting from 1 to 5 going from left to right. Now:

\[
\begin{pmatrix}
1 & 2 & 3 & 4 & 5 \\
3 & 2 & 4 & 5 & 1
\end{pmatrix}A \equiv (1345)A = \{b b a c d\} = A'
\]

\[(123)(1345)A = (123)A' = \{a b b c d\} = (2345)A.
\]

So we have the multiplication rule: \((123)(1345) = (2345)\).
Definition.
The action of an arbitrary permutation $\pi \in S_N$ on an index set
$I = \{i_1, i_2, \ldots, i_N\}$ is defined as above, after numbering the positions
in $I$ from left to right. Writing the basis of $V_n \otimes^N$ as
\[ \{E_I | \{I = i_1, \ldots, i_N\} \} \]
and an arbitrary tensor $T$ with respect to this
basis as $t^I E_I$, we can give two alternative definitions of a linear oper-
ator $\hat{\pi}: V_n \otimes^N \to V_n \otimes^N$ associated with $\pi \in S_N$:
\[
\begin{align*}
&\hat{\pi} E_I = E_{\pi(I)}, \\
&\hat{\pi} t = t^{\pi^{-1}(I)}.
\end{align*}
\]

Notes.

1. It is easy to show that the two definitions of $\hat{\pi}$ give the same linear
operator on $V_n \otimes^N$. For map an arbitrary tensor $T \in V_n \otimes^N$:
\[
\hat{\pi} T = \hat{\pi} (t^I E_I) = t^I E_{\pi(I)} = t^{\pi^{-1}(I)} E_I = (\hat{\pi} t) E_I,
\]
where we used the fact that $I$ runs over all possible index sets. So we
can define $\hat{\pi}$ by its action on the covariant (lower) indices or by its
action on the contravariant (upper) indices of a tensor.

2. The mapping $\pi \to \hat{\pi}$ is a group homomorphism, that is to say:
\[
\begin{align*}
&(i) \hat{\pi} \hat{\pi} E_I = \hat{\pi} E_{\pi(I)} = E_{\pi \pi(I)} \\
&(ii) \hat{\pi} \hat{\pi} t = \hat{\pi} t^{\pi^{-1}(I)} = t^{\pi^{-1} \pi^{-1}(I)} = t^{(\pi \pi)^{-1} I}.
\end{align*}
\]

Example.
In the following chapter Young operators of the NP-type will often be used.
An example of such an operator is:
\[
\hat{Y} = [(1) - (13)](1) + (12)] = (1) + (12) - (13) - (123).
\]

Now:
\[
\begin{align*}
\hat{Y} E_{121} &= E_{121} + E_{112} - E_{121} - E_{112} = E_{211} - E_{112} \\
\hat{Y} t^{121} &= t^{121} + t^{211} - t^{121} - t^{211} = 0.
\end{align*}
\]

Note that $\hat{Y}$ acting on contravariant indices gives zero, whereas acting on
covariant indices it gives a result not equal to zero. Had we used a Young
operator of the PN-type, we would have found the converse of this result.

Having defined representations for $S_N$, we can now turn to $CS_N$. 

45
Definition.
The enveloping algebra of the set of permutation operators on $V_n \otimes^N$ will be denoted by $A_{\text{S}_N}$. This algebra represents the group algebra $CS_N$ on $V_n \otimes^N$.

In the sequel we drop the caret over the permutation operators, not distinguishing in the notation any further between permutations (acting on arbitrary sets) and permutation operators (acting on $V_n \otimes^N$).

4.5. Bisymmetric operators on $V_n \otimes^N$

The linear operators on $V_n \otimes^N$ that belong to the commutator algebra $A_{\text{S}_N}^C$ of $A_{\text{S}_N}$ are often called bisymmetric operators. In this section it will be demonstrated that the algebra of bisymmetric operators is generated by $[N(E)]^N$, the inner direct product of the local Lie subgroup $N(E)$ of $\text{GL}(n)$.

This has two consequences, viz.:
- $A_{\text{S}_N}^C$ is the enveloping algebra of $[\text{GL}(n)]^N$.
- $A_{\text{S}_N}^C$ is the enveloping algebra of $\Lambda([\text{GL}(n)]^N)$; in other words: every bisymmetric operator can be written as a polynomial in the infinitesimal generators of $[\text{GL}(n)]^N$.

Definition.
$N(E)$ is the set of all complex $n \times n$ matrices with matrix elements satisfying:

$$|a_{ij} - \delta_{ij}| < 1/n, \quad (a_{ij}) \in N(E).$$

Lemma 1.
The matrices in $N(E)$ are non-singular, so: $N(E) \subseteq \text{GL}(n)$.

Proof. Define a matrix norm via:

$$\|A\| = \left[\text{Tr}(A^\dagger A)\right]^{1/2}$$

(ref. 18, p. 53). Here $A^\dagger$ is the Hermitean adjoint of $A$ and $\text{Tr}A$ stands for the trace of $A$.

We show that $\|A - E\|^2 < 1$ for all $A \in N(E)$:

$$\|A - E\|^2 = \sum_{i,j} (a_{ij} - \delta_{ij})^\dagger (a_{ij} - \delta_{ij}) = \sum_{i,j} |a_{ij} - \delta_{ij}|^2 < \sum_{i,j} 1/n^2 = 1.$$  

A matrix $A$ with $\|A - E\|^2 < 1$ is non-singular, for:
A = E + (A - E) \equiv E + A' = e^{\ln(E + A')} and A^{-1} = e^{-\ln(E + A')}.

This way of writing A and A^{-1} is permitted since \ln(E + A') is well-defined if \|A'\| < 1 (ref. 18, p. 55).

Notes.
1. N(E), being a subset of GL(n) and a neighbourhood of E, is a local Lie group (ref. 13, p. 162).
2. In the same way as for GL(n), one defines the inner direct product [N(E)]^N of N(E).

We now turn to bisymmetric operators.

Definition.
The elements of AS_C^N, the commutator algebra (sec. I.2.1) of AS_N', are called bisymmetric operators.

The name bisymmetric is due to Weyl [7] and was inspired by the following lemma.

Lemma 2.
Let the linear operator \beta on V_n^N have the matrix (B^\pi_j^I), then:

B^\pi_j^I = B^\pi_j^I, \quad \forall \pi \in S_N

if and only if \beta \in AS_C^N.

Proof.
(i) Assume \beta \in AS_C^N:

B^\pi_j^I E^\pi_j^I = \beta E^\pi_j^I = \pi^{-1} \beta E^\pi_j^I = \pi^{-1} B^\pi_j^I E^\pi_j^I = B^\pi_j^I E^\pi_j^I \Rightarrow

\Rightarrow B^\pi_j^I = B^\pi_j^I.

(ii) Assume B^\pi_j^I = B^\pi_j^I:

\beta E^\pi_j^I = \beta E^\pi_j^I = B^\pi_j^I E^\pi_j^I = B^\pi_j^I E^\pi_j^I \Rightarrow \pi B^\pi_j^I E^\pi_j^I = \pi \beta E^\pi_j^I.

This lemma enables us to recognize bisymmetric operators, as in the following examples.
Examples.
1. The elements of $[\text{GL}(n)]^N$ are bisymmetric, for the $(I,J)$-element of $[A]^N \in [\text{GL}(n)]^N$ is:

$$A^I_J = a^I_{j_1} a^I_{j_2} \ldots a^I_{j_N}.$$ 

The multiplication of the factors on the right hand side is commutative, and so $A^\pi(I) = A^I_J$ for all $\pi \in S_N$.

Note that the elements of the enveloping algebra (sec. I.2.1) of $[\text{GL}(n)]^N$ are also bisymmetric; in fact we will show later in this section that this enveloping algebra is an improper subalgebra of $\mathcal{A}_S^C_N$, i.e. $\mathcal{E}( [\text{GL}(n)]^N ) = \mathcal{A}_S^C_N$.

2. The elements of the Lie algebra $\mathfrak{a}([\text{GL}(n)]^N)$ are bisymmetric. The enveloping algebra of this Lie algebra is also contained in $\mathcal{A}_S^C_N$.

3. A Hamiltonian describing a system of $N$ identical particles is bisymmetric. The physical consequences of this observation will be worked out in ch. I.6.

By definition the matrix of a linear operator on $V_n \otimes^N$ belongs to $\text{AL}(n^N,C)$, which is an algebra of dimension $(n^N)$. From the equality of dimensions follows that this full matrix algebra is equal to $\text{AL}(n,C) \otimes^N$, for this tensor product has dimension $(n^N)$.

The algebra of bisymmetric mappings is a proper subspace of $\text{AL}(n,C) \otimes^N$. This is the content of the following lemma.

---

**Lemma 3.**

$\mathcal{A}_S^C_N$ is a proper subspace of $\text{AL}(n,C) \otimes^N$ of dimension $(n^2 + N - 1)$.

**Proof.** If one considers $\beta \in \mathcal{A}_S^C_N$ as a vector, then its matrix elements $B^I_J$ are its components with respect to the basis of $\text{AL}(n^N,C)$.

Note that $(B^I_J)$ is not necessarily a Kronecker product matrix, or in other words $B^I_J \equiv b^I_{j_1} \ldots b^I_{j_N}$ is not necessarily a product of the form $b^I_{j_1} b^I_{j_2} \ldots b^I_{j_N}$.

Recalling from linear algebra that the number of linearly independent components of an arbitrary vector is equal to the dimension of the subspace to which the vector belongs, we see that the bisymmetric operators certainly belong to a proper subspace of $\text{AL}(n^N,C)$; for the relation
$B^\pi(I) = B^I_J$ shows that many of the components of $\beta$ are equal and hence linearly dependent.

Restrict now the index sets labelling the components $B^i_{j_1 \ldots j_N}$ of $\beta$ in the following manner:

(i) Replace the double index $(i_p, j_p)$ by a single index $k_p = (n - 1)i_p + j_p$ for all $p = 1, 2, \ldots, N$.

(ii) Consider only ordered index sets $K$, i.e. impose on the elements of $K$ the condition:

$$k_1 \leq k_2 \leq \ldots \leq k_N.$$ 

The bisymmetry of $\beta$ yields in this notation:

$$B^\pi(K) = B^K, \quad \forall \pi \in S_N,$$

and so the set $\{B^K\}$ contains only the components of $\beta$ which do not follow from each other by permutation. Any component of $\beta$ not contained in the set $\{B^K\}$ is equal to one of the $B^K$ in the set.

Since no condition other than commutation with all permutation operators has been imposed on bisymmetric operators, we have exhausted all the relations between the components of $\beta$: the elements of $\{B^K\}$ are linearly independent and so the dimension of $A_S^C$ is equal to the order of $\{B^K\}$.

Finally, it is known from combinatorics (see e.g. ref. 27, p. 488) that the number of ordered index sets $K = \{k_1 \leq k_2 \leq \ldots \leq k_n^2\}$ with $k_p = 1, 2, \ldots, n^2$ is equal to $\binom{n^2 + N - 1}{N}$.

We now come to the crucial theorem of this section. The following lemma is needed in the proof.

\textbf{Lemma 4.}

Let $T_i$, $i = 1, \ldots, n^2$, be $n^2$ infinite subsets of $C$. Let $P(x^1, x^2, \ldots, x^{n^2})$ be a polynomial in $n^2$ variables in $C$. If for each $i$ ($i = 1, \ldots, n^2$) $P(x^1, x^2, \ldots, x^{n^2}) = 0$ for all points $x^i \in T_i$ independently of the remaining variables, then $P \equiv 0$. That is, the coefficients of $P$ are all zero.

\textbf{Proof.} Ref. 28, p. 121, corollary 2.

\textbf{Theorem.}

The algebra $A_S^C$ is the enveloping algebra of $[N(E)]^N$. 

49
Proof. Note first that the enveloping algebra $\mathfrak{A}(\{N(\mathbb{E})\}^N)$ of $\{N(\mathbb{E})\}^N$ is contained in $\mathbb{A}^C_N$, for we have: $\{N(\mathbb{E})\}^N \subseteq \{\text{GL}(n)\}^N \subseteq \mathbb{A}^C_N$, and products and linear combinations of bisymmetric operators are bisymmetric, so closing $\{N(\mathbb{E})\}^N$ multiplicatively and linearly does not bring us outside $\mathbb{A}^C_N$.

Note secondly that both $\mathbb{A}^C_N$ and $\mathfrak{A}(\{N(\mathbb{E})\}^N)$ are subspaces of $\text{AL}(n,\mathbb{C}) \otimes \mathbb{N}^N$. As we will show that these spaces have the same dimension, they must coincide.

To compute the dimension of $\mathfrak{A}(\{N(\mathbb{E})\}^N)$ we consider how many of the matrix elements of an arbitrary matrix $[X]^N \in \{N(\mathbb{E})\}^N$ are linearly independent. Write to that end first the elements of $[X]^N$ thus:

$$
i_1, i_2, \ldots, i_N \quad j_1, j_2, \ldots, j_N = x_{i_1}^{j_1} \cdot x_{i_2}^{j_2} \cdot \ldots \cdot x_{i_N}^{j_N} \equiv K$$

with $k_p = (n - 1)i_p + j_p$, $p = 1, \ldots, N$, $k_p = 1, \ldots, n^2$. In the manner of the proof of lemma 3 it now follows immediately that only the $\binom{n^2 + N - 1}{N}$ elements $x^K$ with ordered index set $K$ are essentially different.

The matrix elements $x_{i_p}^{j_p}$ satisfy the condition:

$$\frac{|x_{i_p}^{j_p} - \delta_{i_p}^{j_p}|}{1/n, \quad p = 1, \ldots, N},$$

demonstrating that the diagonal elements belong to a neighbourhood $T_1 \subseteq \mathbb{C}$ of 1 and that the non-diagonal elements belong to a neighbourhood $T_0 \subseteq \mathbb{C}$ of 0. Each element $x_{i_p}^{j_p}$ may be chosen independently of the remaining elements from its neighbourhood; any choice of $n^2$ elements yields a non-singular (by lemma 1) matrix belonging to $N(\mathbb{E})$.

The monomials $x^K$ are linearly independent. To prove this we note that a linear relation among them is necessarily trivial (i.e. has all coefficients equal to zero), because the following linear relation between the essentially different matrix elements $x^K$ of $[X]^N$:

$$\sum_{k_1 < k_2 \ldots < k_N} c_{k_1, k_2 \ldots, k_N} x_{k_1}^{k_2} \ldots \cdot x_{k_N} = 0$$

implies by lemma 4 that all $c_{k_1, \ldots, k_N}$ are zero.

Summarizing: an arbitrary matrix $[X]^N \in \{N(\mathbb{E})\}^N$ has $\binom{n^2 + N - 1}{N}$ linearly independent matrix elements (the components of $[X]^N$ with respect to the basis of $\text{AL}(n,\mathbb{C}) \otimes \mathbb{N}^N$), hence $\{N(\mathbb{E})\}^N$ generates an $\binom{n^2 + N - 1}{N}$-dimensional subspace (the algebra of bisymmetric operators) of $\text{AL}(n,\mathbb{C}) \otimes \mathbb{N}^N$. 

50
Note.
In the proof just given it is essential that each of the elements \( x^i \) (i = 1, ..., n^2) can be chosen independently of the remaining elements. If this is not the case, lemma 4 is not applicable. Therefore, the same proof does not hold for Lie groups, consisting of n x n matrices, that have less than n^2 parameters.

Corollary.
The algebra \( AS^C_N \) is the enveloping algebra of \( [GL(n)]^N \).

Proof. The enveloping algebra of \( [GL(n)]^N \) is obviously contained in \( AS^C_N \). The subset \( [N(E)]^N \) of \( [GL(n)]^N \) already generates all of \( AS^C_N \) and hence \( [GL(n)]^N \) does so a fortiori.

Theorem.
The algebra \( AS^C_N \) is the enveloping algebra of \( \Lambda([GL(n)]^N) \).

Proof. The enveloping algebra of \( \Lambda([GL(n)]^N) \) is obviously contained in \( AS^C_N \). We will show that \( [N(E)]^N \) is a subset of this enveloping algebra, from which the theorem follows immediately.
The matrix \( \ln([X]^N), [X]^N \in [N(E)]^N \), is well-defined and belongs to \( \Lambda([GL(n)]^N) \):

\[
\ln([X]^N) = \ln(X \otimes X \otimes \ldots \otimes X)
= \ln(X(1)X(2) \ldots X(N))
= \ln_{X(1)} + \ln_{X(2)} + \ldots + \ln_{X(N)}
\equiv a(1) + a(2) + \ldots + a(N)
\equiv A \in \Lambda([GL(n)]^N).
\]

The second line follows from the definition of \( X(k) \) (sec. I.4.3). The fourth line follows from \( \ln_{X(k)} \equiv a(k) \) being well-defined; for write \( X(k) = E(k) + X'(k) \), then:

\[
\ln_{X}(k) = \ln(E(k) + X'(k)) = \sum (-1)^{q-1} \frac{q}{q} (E \otimes \ldots \otimes X' \otimes \ldots \otimes E)^q
= E \otimes E \otimes \ldots \otimes \ln_{X'} \otimes \ldots \otimes E
\]

and, as is shown in the proof of lemma 1, \( \ln X' \) exists; so it belongs to \( \Lambda(GL(n)) = AL(n,C) \). The last line follows from the theorem of sec. I.4.3.

Writing finally:
\[
[x]^N = e^{\frac{x}{n}[x]}^N = e^{\sum_{q=1}^{\infty} \frac{1}{q!} \frac{x^q}{n^q}} , \quad n \in N([GL(n)]^N)
\]

we see that \([x]^N \in [N(\mathbb{Z})]^N\) belongs to the enveloping algebra of \(N([GL(n)]^N)\).

4.6. The interconnection between the tensor irreps of GL(n) and \(S_N\)

We have now laid the ground for the main theorem of this chapter. In fact everything regarding this theorem has already been proved and so the following result is not new, but a summary of the foregoing theory; it is for that purpose that it will be stated and proved explicitly.

**Theorem.**
Let \(V_n \otimes^N\) be decomposed into \(S_N\)-irreducible subspaces \(U_j^{[\lambda]}\) thus:

\[
V_n \otimes^N = \sum \Theta R(\lambda) , \quad R(\lambda) = \sum \Theta U_j^{[\lambda]} \quad (1)
\]

Let the spaces \(U_j^{[\lambda]}\), \(j = 1, \ldots, n_{<\lambda>}\), span identical \(f_{<\lambda>\cdot}\)-dimensional irreps of \(S_N\). Then the basis of \(V_n \otimes^N\) adapted to the decomposition (1) is simultaneously adapted to the decomposition:

\[
V_n \otimes^N = \sum \Theta R(\lambda) , \quad R(\lambda) = \sum_{i=1}^{f_{<\lambda>\cdot}} \Theta W_i^{<\lambda>} \quad (2)
\]

where the spaces \(W_i^{<\lambda>}\), \(i = 1, \ldots, f_{<\lambda>\cdot}\), span identical \(n_{<\lambda>\cdot}\)-dimensional irreps of \(GL(n)\), \(U(n)\) and \(SU(n)\).

Conversely, a decomposition of \(V_n \otimes^N\) according to (2) induces a simultaneous decomposition according to (1).

**Proof.** The interdependence of the different representations is given by the following scheme:

Inside the box the permutation group \(S_N\) (and its group algebra \(\mathbb{C}[S_N]\)) standing on the left, and the Lie groups \(GL(n)\), \(U(n)\) and \(SU(n)\) standing on the right, are represented by linear operators on \(V_n \otimes^N\).

Explanation of the different steps:
1. \(S_N\) and its group algebra \(\mathbb{C}[S_N]\) share their irreps (see e.g. ref. 9, 10 or 14).
2. \( \text{CS}_N \) is represented on \( \mathbb{V}_n \otimes^N \) by \( \text{AS}_N \) (see sec. I.4.4).

3. Decomposition of \( \mathbb{V}_n \otimes^N \) under \( \text{AS}_N \) induces a simultaneous decomposition of \( \mathbb{V}_n \otimes^N \) under \( \text{AS}_N^C \), the commutator algebra of \( \text{AS}_N \). Also the converse of this statement holds (see sec. I.2.3).

4. \( \text{AS}_N^C \) is the enveloping algebra of \( [\text{GL}(n)]^N \) (see sec. I.4.5).

5. The enveloping algebra of an operator group shares its irreps with the group (see sec. I.2.1).

6. \( \text{GL}(n) \) is represented on \( \mathbb{V}_n \otimes^N \) by \( [\text{GL}(n)]^N \) (see sec. I.4.2).

7. \( \text{GL}(n) \) and \( \text{U}(n) \) share their irreps (see sec. I.3.3).

8. An irreducible tensor representation of \( \text{U}(n) \) stays irreducible under restriction to \( \text{SU}(n) \).

We show step 8:

The irreducible tensor representation \( T^{<\lambda>}(\mathbb{U}), \mathbb{U} \in \text{U}(n) \), is obtained from a similarity transformation on the matrix \( [\mathbb{U}]^N \in [\text{U}(n)]^N \). Hence, the matrix elements of \( T^{<\lambda>}(\mathbb{U}) \) are homogeneous polynomials of the order \( N \) in the matrix elements \( u_{ij} \) of \( \mathbb{U} \). Now, any matrix \( \mathbb{U} \in \text{U}(n) \) can be written as \( \mathbb{U} = \alpha \mathbb{U}' \) with \( \mathbb{U}' \in \text{SU}(n) \) (take \( \alpha = (\det(\mathbb{U}))^{1/n} \)), and so:

\[
T^{<\lambda>}(\mathbb{U}) = T^{<\lambda>}(\alpha \mathbb{U}') = \alpha^N T^{<\lambda>}(\mathbb{U}').
\]

Applying the argument of ref. 9, p. 45 it follows that \( T^{<\lambda>} \) is an irreducible representation of \( \text{SU}(n) \) if and only if \( T^{<\lambda>} \) is an irreducible representation of \( \text{U}(n) \).

Note.

In this section only the irreps of the global groups have been discussed. Knowing that these irreps also constitute irreps of the corresponding Lie algebras (see sec. I.3.1), it follows that everything that has been stated in this section about the groups \( \text{GL}(n), \text{U}(n) \) and \( \text{SU}(n) \) also holds for the respective Lie algebras: \( \Lambda(\text{GL}(n)), \Lambda(\text{U}(n)) \) and \( \Lambda(\text{SU}(n)) \).

4.7. A note on the Casimir invariants of \( \text{GL}(n) \)

A well-known quantum mechanical problem is the determination of a complete set of commuting operators. The elements of such a set possess a set of common eigenvectors, which are uniquely determined (up to phase). Once a complete set of operators has been determined, the state labelling problem has been solved; i.e. every state is unambiguously characterized by a complete set
of eigenvalues. For instance the problem of labelling the atomic states arising from configurations of equivalent electrons led Racah in 1943 to the introduction of the seniority operator $Q$, which together with the angular momentum operators $L^2$ and $L_z$ and spin operators $S^2$ and $S_z$ yields an unambiguous labelling of states arising from $d^n$-configurations. In 1949 Racah was able to prove that $Q$, just like $L^2$, $L_z$, $S^2$ and $S_z$, is a Casimir invariant of a certain Lie group. This discovery initiated the search for, what is now called, canonical chains of groups, for the Casimir invariants, associated with such chains, form complete sets of commuting operators.

In this section we will show that the operators, representing the Casimir invariants of $\text{GL}(n)$ on tensor space, belong to the center of $A_N^\infty$. But before that, Casimir invariants will briefly be discussed in a more general context.

**Definitions.**

1. Any (abstract) Lie algebra $\Lambda$ can be imbedded into an (abstract) associative algebra $\mathfrak{g}_U(\Lambda)$ in the following manner. Define a formal associative product on $\Lambda$ and close $\Lambda$ multiplicatively with respect to this product, thus constructing a semigroup of infinite order out of $\Lambda$. Close this semigroup linearly while defining $[a, b] = ab - ba$, $a, b \in \Lambda$, to relate the Lie product to the new associative product. The resulting infinite dimensional associative algebra $\mathfrak{g}_U(\Lambda)$ is the universal enveloping algebra of $\Lambda$.

2. The elements in the center of $\mathfrak{g}_U(\Lambda)$ are the Casimir invariants of $\Lambda$, i.e. they are elements $I \in \mathfrak{g}_U(\Lambda)$ such that $aI = Ia$, $\forall a \in \Lambda$; and hence $I$ commutes with all elements of $\mathfrak{g}_U(\Lambda)$.

**Notes.**

1. See for a mathematically more satisfying definition of $\mathfrak{g}_U(\Lambda)$ ref. 29, p. 155.

2. A linear representation $D$ of $\Lambda$ extends to a representation of $\mathfrak{g}_U(\Lambda)$. $\mathfrak{g}_U(\Lambda)$ is represented by the enveloping algebra of $D(\Lambda)$.

3. If $D$ is irreducible then, by Schur's lemma, $D(I) = \lambda I E$, for all Casimir invariants $I$ of $\Lambda$.

Returning to the classical groups, we first note that $\text{GL}(n)$ shares its Casimir invariants with $\text{U}(n)$. This is obvious since the Lie algebra of
GL(n) is the complexification of the Lie algebra of U(n) (sec. I.3.2). So anything stated in the sequel about U(n) holds also for GL(n) (when considered as an n²-dimensional complex Lie group).

Gelfand and later Biedenharn [30] have derived explicit expressions for a set of basic invariants of U(n). ('Basic' means here that any Casimir invariant of U(n) can be expressed as an analytic function of these invariants, which themselves are functionally independent). It can be proved [30] that U(n) has n and that SU(n) has n-1 basic invariants. Denoting the operators representing the invariants \( \{ \hat{I}^{(n)}_i \} \) of U(n) on \( V_n \otimes \) by \( \{ \hat{I}^{(n)}_i \} \) it follows from Schur's lemma that the \( (f_{(\lambda)} \times n_{(\lambda)}) \)-dimensional spaces \( R^{(\lambda)} \), defined in sec. I.4.6 as the direct sum of all the spaces carrying the irrep \( \lambda \) of \( S_N \), are eigenspaces of the n operators \( \{ \hat{I}^{(n)}_i \} \). Biedenharn proved that the n corresponding eigenvalues give a unique designation of the irreps of U(n). These eigenvalues have been computed explicitly by Louck [31]; they are given as rather complicated expressions in terms of partial hooks [31]. (Recently Hudson [32] derived another system of basic invariants for U(n), of which the eigenvalues are simple polynomials in the components of the highest weight of the irrep).

Of course the n operators \( \{ \hat{I}^{(n)}_i \} \) representing the Casimir invariants of U(n) do not constitute a complete set of operators. However, the \( h n(n+1) \) invariants belonging to the canonical chain:

\[
U(n) \supset U(n-1) \supset \ldots \supset U(1)
\]

form a complete set of operators (we come back to this point in sec. I.5.6).

The operator \( S^2 \), representing the (only) Casimir invariant of SU(2) on N-electron spin space, can be expressed as a linear combination of permutation operators. This can be derived from the Dirac identity [33, p. 222]. The following theorem can be considered as an alternative derivation and a generalization of this result.

Theorem.
The operators on \( V_n \otimes \) that represent the Casimir invariants of GL(n) and U(n) belong to the center \( \mathfrak{c} S_N \) of \( \mathfrak{a} S_N \).

Proof. An invariant I belongs by definition to the center of the universal enveloping algebra of the Lie algebra of GL(n). Hence \( \hat{I} \), representing I on \( V_n \otimes \), belongs to the center of the enveloping algebra of \( M[GL(n)]^N \).

In sec. I.4.5 it has been proved that this enveloping algebra coincides with \( \mathfrak{a} S_N \), the commutator algebra of \( \mathfrak{a} S_N \). The center of an algebra being
the intersection of the algebra and its commutator algebra, we get:

\[ \hat{\mathbf{i}} \in A S_N^C \cap (A S_N^C)^C. \]

From sec. I.2.3 we know that \((A S_N^C)^C = A S_N^C\) (\(A S_N\) is semi-simple, as it represents the finite group algebra \(CS_N\)), and therefore:

\[ \hat{\mathbf{i}} \in (A S_N^C \cap A S_N^C) \equiv A S_N. \]

**Notes.**

1. The map of the Casimir invariants of \(U(n)\) is usually into \(A S_N^C\). So, in general, the set operators \(\{\hat{\mathbf{i}}^{(n)}_i\}\) spans a proper subspace of \(A S_N^C\).

2. Two possible bases of \(A S_N^C\) are the set of class sum operators and the set of character projectors of \(S_N\) (see e.g. ref. 10, theorems 5.3, 5.4). The operators \(\{\hat{\mathbf{i}}^{(n)}_i\}\) can be expressed in terms of either of these two sets.

**Example.**

The two spin orbitals \(\alpha\) and \(\beta\) span the 2-dimensional spin space \(V_2\), which carries an irrep of \(\Lambda(SU(2))\) and its complexification \(\Lambda^C(SU(2))\). To clarify the theory of this chapter we discuss some points regarding the decomposition of \(V_2 \otimes^4\), the 4-electron spin space.

From standard angular momentum theory [17, sec. 3.1] it is known how to decompose \(V_2 \otimes^4\) under \(\Lambda^C(SU(2))\). Briefly, the procedure is the following:

(i) Choose the vector with highest eigenvalue of \(S^z\) (there is only one such vector in spin space).

(ii) Apply the step-down operator:

\[ S^- = \sum_k s_-(k), \quad s_-(k) = s_x(k) - i s_y(k) \]

repeatedly onto this vector until the zero vector is generated.

Because \(S^- \in \Lambda^C([SU(2)]^N)\) (see sec. I.4.3), the space \(V^S\) so generated carries an irrep of \(\Lambda^C(SU(2))\) and of \(\Lambda(SU(2))\). \(SU(2)\) being the universal covering group (sec. I.3.1 and I.3.4) of all the groups with this same Lie algebra, \(V^S\) carries also an irrep of \(SU(2)\).

(iii) Choose a vector with the highest but one eigenvalue of \(S^z\), orthogonalize to \(V^S\), apply the step-down operator, etc.

Proceeding in this manner, one finds that \(V_2 \otimes^4\) decomposes into \(SU(2)\)-irreducible spaces thus:

\[ V_2 \otimes^4 = V^S_{S=2} \oplus V^S_{S=1} \oplus V^S_{S=1} \oplus V^S_{S=0} \oplus V^S_{S=0} \]
with corresponding basis:

\[
V'_{S=2} : \; \begin{array} {c|c|c|c|c}
2,2,1 & 2,1,0 & 2,0,1 & 2,-1,0 & 2,-2,1 \\
\end{array}
\]

\[
V'_{S=1} : \; \begin{array} {c|c|c|c}
1,1,1 & 1,0,1 & 1,-1,1 \\
\end{array}
\]

\[
V''_{S=1} : \; \begin{array} {c|c|c|c}
1,1,2 & 1,0,2 & 1,-1,2 \\
\end{array}
\]

\[
V'_{S=0} : \; \begin{array} {c|c|c}
1,1,3 & 1,0,3 & 1,-1,3 \\
\end{array}
\]

\[
V''_{S=0} : \; \begin{array} {c|c|c}
0,0,1 & 0,0,2 \\
\end{array}
\]

where \( |S,M_S,i\rangle \) denotes an eigenvector of \( S^2 \) with quantum number \( S \), and of \( S \_z \) with quantum number \( M_S \); \( i \) labels the remaining multiplicity. Applying now the theorem of sec. I.4.6 we see that each of the vectors \( |2,M_S,i\rangle \), \( M_S = -2, \ldots, +2 \), spans the same 1-dimensional irrep of \( S_4 \); the 3-dimen-

sional sets: \{ \( |1,1,i\rangle \) \}, \{ \( |1,0,i\rangle \) \} and \{ \( |1,-1,i\rangle \) \} \( i = 1,2,3 \) span the same 3-dimensional irrep of \( S_4 \); and \{ \( |0,0,i\rangle \) \} \( i = 1,2 \) spans a 2-dimensional irrep of \( S_4 \).

The Casimir operator \( S^2 \) can be expressed in terms of the character projectors (sec. I.2.2) of \( S_4 \):

\[
S^2 = \sum_{\lambda} c_{\lambda} e^{[\lambda]}.
\]

Applying \( S^2 \) in this form to a projected vector \( e^{[\mu]}(v) \), \( v \in V_2 \otimes^4 \), further using \( e^{[\lambda]} e^{[\mu]} = \delta_{\lambda \mu} e^{[\lambda]} \) together with the fact that \( e^{[\mu]}(v) \) is an eigen-

vector of \( S^2 \), it follows that the expansion coefficients \( c_{\lambda} \) are eigen-

values of \( S^2 \). In ch. I.6 the irreps of \( S_N \) will be brought into correspond-

tce to the eigenvalues \( S(S+1) \) of \( S^2 \), anticipating that result we get

\[
S^2 = 2(2 + 1)e^{[4]} + 1(1 + 1)e^{[3,1]} + 0(0 + 1)e^{[2,2]}.
\]

One can also rewrite this expression in terms of class sum operators. After some manipulation one arrives at:

\[
S^2 = c^{(2,1^2)} = (12) + (13) + (14) + (23) + (24) + (34).
\]

This result could also have been obtained directly by applying the Dirac identity [33].
CHAPTER 1.5. THE DECOMPOSITION OF TENSOR SPACE

In the foregoing chapter it has been shown that the decomposition of $V_n \otimes^N$ under $S_N$ goes hand in hand with the decomposition under the Lie groups $GL(n)$, $SU(n)$, $U(n)$ and their respective Lie algebras. In this chapter it will be discussed how bases of subspaces of $V_n \otimes^N$, carrying irreps of these groups, can be constructed in practice.

In principle this purpose may be achieved in one of the following three ways:

(i) Use Lie algebraic techniques to decompose $V_n \otimes^N$ under $\Lambda(U(n))$. This decomposition may then be followed by a computation of the matrix elements of the generators of $U(n)$ over the irreducible bases thus constructed. Since every linear operator on $V_n \otimes^N$ can be expressed in terms of the generators of $U(n)$, the matrix representation of every (bisymmetric) linear operator can then be calculated. The full power of this approach has yet to be explored by quantum chemists, although recently discussions of this method, applied to $N$-electron systems, have been given by Paldus [34] and Gouyet et al. [35].

(ii) Use an approach from the point of view of the global group $U(n)$, generalizing the necessary results of finite groups to this compact group. This method requires very sophisticated mathematics.

(iii) Use the representation theory of $S_N$, developed by Young, Littlewood and others.

The last method is probably the most convenient of the three, and is in any case by far the best known among theoretical chemists. We will also follow this line of approach. To that end some of the necessary theory regarding the decomposition of the group algebra $CS_N$ of $S_N$ is introduced in the first three sections. We rely heavily on the book of Rutherford [36] for that, although not all the results to be presented can be found in that reference.

Two different basis sets corresponding to a complete decomposition of the group algebra of $CS_N$ will be discussed.

The first consists of Young units of the NP-type, sometimes called "structure projectors" [37], leading to the "Weyl-Rumer basis", the spin-free equivalent of "spin-bonded" functions [38-40]. This basis has been used in the computer calculations presented in the second half of this thesis.

The second basis of $CS_N$ to be discussed consists of Young-Yamanouchi units, which carry orthogonal representations of $S_N$. These units give rise to a basis of $V_n \otimes^N$ known as the Gelfand basis.
The definitions of Young diagrams, (standard) Young tableaux etc., especially in connection with their use for labelling the irreps of \( S_N \), are presupposed; a good reference for this is an article by Coleman [41]. We just note that the term **Young diagram** applies to a frame with empty boxes, while a **Young tableau** consists of \( N \) boxes, filled with the numbers 1 to \( N \).

### 5.1. Young units

**Definitions.**

1. All possible permutations moving numbers along the rows of a Young tableau \( T_r[^\lambda] \) form a group \( R_r[^\lambda] \): the **row group** belonging to the tableau \( T_r[^\lambda] \). The elements of \( R_r[^\lambda] \) are called horizontal permutations.

2. All possible permutations moving numbers along the columns of a Young tableau \( T_r[^\lambda] \) (vertical permutations) form a group \( C_r[^\lambda] \): the **column group** belonging to \( T_r[^\lambda] \).

3. The element
   \[
P_r[^\lambda] = \sum \pi,
   \]
   where the sum runs over all \( \pi \in R_r[^\lambda] \), is the **row symmetrizer** belonging to \( T_r[^\lambda] \).

4. The element
   \[
   p_r[^\lambda] = \sum \zeta_v \nu,
   \]
   where \( \zeta_v \) is the parity of \( \nu \) and the sum runs over all \( \nu \in C_r[^\lambda] \), is the **column antisymmetrizer** belonging to \( T_r[^\lambda] \).

**Example.**

\[
T_r[2,2] = \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} \quad \text{(a non standard Young tableau)}.
\]

\[
\]

\[
C_r[2,2] = \{(1),(34),(12),(12)(34)\}, \quad N_r[2,2] = (1) - (34) - (12) + (12)(34).
\]

**Definition.**

Let \( T_s[^\lambda] \) and \( T_t[^\lambda] \) be Young tableaux with shape \( [^\lambda] \), then we define the permutation \( c[^\lambda]_{st} \) by:
\[ \sigma_{ts} \tau_{s} = \tau_{t} \], \quad \sigma_{ts} \in S_N. \]

Example.

\[ T_s^{[\lambda]} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 5 \end{bmatrix}, \quad T_t^{[\lambda]} = \begin{bmatrix} 2 & 5 & 3 \\ 4 & 1 \end{bmatrix} \Rightarrow \sigma_{ts}^{[3,2]} = (1235). \]

(Note the permutation convention; \( \sigma_{ts} \) operates on the numbers. Because a Young tableau contains by definition \( N \) different numbers no confusion is possible).

Definition.

Let \( T_s^{[\lambda]} \) and \( T_t^{[\lambda]} \) be Young tableaux with shape \( [\lambda] \), then:

\[ Y_{ts}^{[\lambda]} = \sigma_{ts}^{[\lambda]} N_s^{[\lambda]} P_s^{[\lambda]} \]

is a Young unit of the NP-type.

Note.

The definition:

\[ Y_{ts}^{[\lambda]} = \sigma_{ts}^{[\lambda]} P_s^{[\lambda]} N_s^{[\lambda]} \]

gives a Young unit of the PN-type.

Lemma.

\[ Y_{ts}^{[\lambda]} = \sigma_{ts}^{[\lambda]} N_s^{[\lambda]} P_s^{[\lambda]} = N_t^{[\lambda]} \sigma_{ts}^{[\lambda]} P_s^{[\lambda]} = N_t^{[\lambda]} P_t^{[\lambda]} \sigma_{ts}^{[\lambda]} \]

Proof. Ref. 36, p. 16.

From here on the subscripts of the Young units will be running over standard tableaux only, unless stated differently. Recall in this connection that \( f_{[\lambda]} \) standard tableaux of shape \( [\lambda] \) can be constructed, where \( f_{[\lambda]} \) is the dimension of the irreps \( [\lambda] \) of \( S_N \).

Definition.

Since the Young units belong to \( CS_N \), they can be expressed as a linear combination of permutations:

\[ Y_{rs}^{[\lambda]} = \sum_{P \in S_N} U^{[\lambda]}(P)_{rs} P. \]

Letting \( r \) and \( s \) run over standard tableaux and keeping \( P \) fixed, the expansion coefficients \( U^{[\lambda]}(P)_{rs} \) form an \( f_{[\lambda]} \times f_{[\lambda]} \) matrix \( U^{[\lambda]}(P) \).
Notes.

1. The matrix $U^{[\lambda]}(P)$ is not a matrix representation of $P$.
2. From its definition the matrix $U^{[\lambda]}(P)$ can be calculated by writing out all $f^{2}_{[\lambda]}$ units. A shorter route is described by means of the following example.

Example.

Tableaux $\begin{pmatrix} 1 & 2 \\ 3 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 3 \\ 2 \\ 2 \end{pmatrix}$ give rise to the Young units:

$Y_{11} = (1) + (12) - (13) - (123)$ $\quad Y_{12} = (23) - (132) + (123) - (12)$

$Y_{21} = (23) + (132) - (123) - (13)$ $\quad Y_{22} = (1) - (12) + (13) - (132)$

and so for instance:

$U^{[2,1]}(123) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$, $U^{[2,1]}(132) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$

A direct way of computing the transposed matrix $U^{[\lambda]}(P)^T$ for a given permutation $P$ is the following:

Construct an $f^{\lambda}_{x} \times f^{\lambda}_{x}$ table in which the columns are labelled by standard tableaux $T^{[\lambda]}_{x}$, $x = 1, 2, \ldots, f^{\lambda}_{x}$ and the rows by the permuted tableaux $P(T^{[\lambda]}_{x})$, $x = 1, \ldots, f^{\lambda}_{x}$, thus:

\[
\begin{array}{c|cc}
(123): & 12 & 13 \\
\hline
23 & 32 & 23 \\
1 & 1 & 1 \\
21 & 12 & 0 \\
3 & 3 & 2 \\
\end{array}
\quad \begin{array}{c|cc}
(132): & 12 & 13 \\
\hline
31 & 0 & 13 \\
2 & 2 & 2 \\
32 & 32 & 23 \\
1 & 1 & 1 \\
\end{array}
\]

The tableaux contained inside the tables are obtained by application of certain horizontal permutations to the tableaux labelling the rows. These horizontal permutations are chosen such that after permutation the numbers appear in the same column as in the tableau at the top. If no such horizontal permutation exists, that is, if two numbers would have to appear in the same box, we write down a zero in the table. The tableaux thus obtained can now be transformed to the standard tableaux at the top via a vertical permutation $v$ with parity $\zeta_{v}$. The matrix $U^{[\lambda]}(P)^T$ follows finally by replacing the tableaux in the table by this parity, so:

$U^{(123)^T} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$

$U^{(132)^T} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$. 

61
The proof of this construction is given by the NP-equivalence of Rutherford's theorem 7 [36].

The Young units that belong to the same diagram $[\lambda]$ are not necessarily orthogonal. Still, by virtue of Von Neumann's theorem [36, theorem 8], they satisfy a rather simple multiplication rule. This is the content of the following theorem.

Theorem.

$$ y^{[\lambda]}_{ru} y^{[\mu]}_{vs} = \frac{N!}{\xi^{[\lambda]}_{vu}} \delta_{\lambda, \mu} y^{[\lambda]}_{rs}. $$

Here $\xi^{[\lambda]}_{vu}$ is the coefficient of the identity (1) in $y^{[\lambda]}_{vu}$.


Note.

The numbers $\xi^{[\lambda]}_{vu}$, $v, u = 1, \ldots, f^{[\lambda]}$, constitute an $f^{[\lambda]} \times f^{[\lambda]}$ matrix $\lambda^{[\lambda]}$, the "NP-structure matrix", which will play an important rôle in this and the following chapter. It is easy to show from its definition that the matrix elements of $\lambda^{[\lambda]}$ are given by the following rules. Assume to that end that the standard tableaux are in dictionary order (defined in ref. 8, ch. IV, §4) and suppress $[\lambda]$ for the moment. Then:

$$ u = v \quad \xi^{[\lambda]}_{vv} = 1 $$
$$ u > v \quad \xi^{[\lambda]}_{uv} = 0 $$

$$ \left\{ \begin{array}{l}
\xi^{[\lambda]}_{uv} = \tau_{ut}, \text{if } T_v \text{ does not have a pair of numbers in one row that } T_u \text{ has in one column.} \\
\xi^{[\lambda]}_{uv} = 0, \text{if } T_v \text{ has a pair of numbers in one row that } T_u \text{ has in one column.} 
\end{array} \right. $$

The quantity $\tau_{ut}$, which is the parity of the permutation $\sigma_{ut}$, arises from the following consideration. If $\xi^{[\lambda]}_{uv} \neq 0$ then $\sigma_{uv}$ can be written as [36, theorem 5]:

$$ \sigma_{uv} = v_u \pi_v, \quad v_u \in C_u, \quad \pi_v \in R_v. $$

Defining now the tableau $T_t$ by:

$$ T_t = \pi_v T_v $$
we see that \( \sigma_{ut} = \nu_u \), for: \( T_u = \sigma_{uv} T_v = \nu_u \pi_v T_v = \nu_u T_t \), and so the parity \( \tau_{ut} \) of \( \sigma_{ut} \) is equal to the parity of \( \nu_u \).

A mechanical way to obtain \( \frac{X}{[\lambda]} \) is the same as for \( \frac{\nu}{[\lambda]} \) (P) if we take \( P = (1) \). Label the rows of an \( f_{[\lambda]} \times f_{[\lambda]} \) array with the standard tableaux \( T_v \), \( v = 1, \ldots, f_{[\lambda]} \) and the columns with the standard tableaux \( T_u \), \( u = 1, \ldots, f_{[\lambda]} \). Apply the horizontal permutation \( \pi_v \) to \( T_v \), with \( \nu_v T_v = T_t \), so that the digits in \( T_t \) appear, if possible, in the same column as in \( T_u \). Replace the tableau \( T_t \) thus obtained by \( \zeta_{uv} = \tau_{ut} \). Transpose the result.

The NP-structure matrix is very easily determined, as can be seen in the following example.

**Example.**

\[
\begin{array}{ccccccc}
  & 1 & 2 & 1 & 2 & 1 & 3 & 1 & 3 & 1 & 4 \\
1 & 2 & 1 & 2 & 1 & 3 & 1 & 3 & 1 & 4 \\
2 & 3 & 4 & 3 & 5 & 2 & 4 & 2 & 5 & 2 & 5 \\
3 & 5 & 4 & 5 & 4 & 3 & 5 & 4 & 3 & 5 & 4 \\
\end{array}
\]

\[
\frac{X}{[\lambda]} = \begin{pmatrix}
1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

**Notes.**

1. The matrix \( \frac{X}{[\lambda]} \) is non-singular, for because it is upper triangular and has +1 along the diagonal its determinant is +1.
2. The inverse matrix \( (\frac{X}{[\lambda]} \) is also upper triangular.
3. The elements of this inverse matrix will be denoted by \( \eta_{rs} \). They also will be used frequently in the sequel.
As will be stated in the following theorem, the Young units form a basis of \( CS_N \) corresponding to a simultaneous decomposition of this semi-simple algebra into minimal left ideals, right ideals and simple invariant subalgebras. So, except for the fact that they lack orthogonality, Young units give a Wedderburn decomposition of \( CS_N \) (see sec. I.2.2).

**Theorem.**

(i) The "diagonal" elements:

\[
\frac{f[\lambda]}{N!} \psi_{rr}^{[\lambda]}, \quad \text{all } [\lambda], \ r = 1, \ldots, f[\lambda]
\]

are **primitive** idempotents.

(ii) The set:

\[
\{ \psi_{rs}^{[\lambda]} \mid s = 1, \ldots, f[\lambda] , \ [\lambda] \text{ and } r \text{ fixed} \}
\]

spans a minimal **right** ideal in \( CS_N \).

(iii) The set:

\[
\{ \psi_{rs}^{[\lambda]} \mid r = 1, \ldots, f[\lambda] , \ [\lambda] \text{ and } s \text{ fixed} \}
\]

spans a minimal **left** ideal in \( CS_N \).

(iv) The Young units afford a resolution of the identity:

\[
(1) = \sum_{[\lambda]} \frac{f[\lambda]}{N!} \sum_{r=1}^{f[\lambda]} \eta_{tr}^{[\lambda]} \psi_{rt}^{[\lambda]}. 
\]

(v) The set:

\[
\{ \psi_{rs}^{[\lambda]} \mid \text{all } [\lambda], \ r, \ s = 1, \ldots, f[\lambda] \}
\]

forms a complete linearly independent basis of \( CS_N \).

**Proof.**

(i) Translate Boerner's theorem 3.1 [8, chapter IV] into the NP-definition of a Young unit.

(ii) and (iii) Apply the manner of proof used at p. 102 of ref. 42.

(iv) Ref. 36, § 29. Note in this respect that

\[
\phi_{rr}^{[\lambda]} = \sum_{t} \eta_{rt}^{[\lambda]} \psi_{tr}^{[\lambda]}
\]

is a natural unit.

(v) Apply (iv) together with theorem 4.6 in chapter IV of ref. 8.

**Definition.**

Because of the theorem just stated the Young units span irreducible
representations of $CS_N$, with the elements of $CS_N$ acting from the right:

$$y_{rs}^\lambda(\hat{Q}) = \sum_{t=1}^f r^D(\hat{Q})_{ts} y_{rt}^\lambda, \quad \hat{Q} \in S_N$$

and also from the left:

$$Q y_{rs}^\lambda = \sum_{t=1}^f L^D(\hat{Q})_{tr} y_{ts}^\lambda, \quad \hat{Q} \in S_N.$$

The irreducible matrix $D(\hat{Q})$ defined this way can easily be computed, using the following theorem.

---

Theorem.

$$r^D(\hat{Q}) = (\chi^\lambda)^{-1} \sum_{rt}^{(Q)} y_{rt}^\lambda,$$

where $\chi^\lambda$ is the $NP$-structure matrix previously defined, and $\sum_{rt}^{(Q)} y_{rt}^\lambda$ contains the components of the Young units along $\hat{Q}$.

---

Proof. Drop $[\lambda]$.

$$\sum_{\hat{Q} \in S_N} u(P)_{rs} \hat{Q} = \sum_{t=1}^f r^D(\hat{Q})_{ts} y_{rt} = \sum_{\hat{Q} \in S_N} u(P)_{rt} r^D(\hat{Q})_{ts}.$$

Substitute $R = \hat{Q}^{-1}$ in the leftmost expression and $P = R$ in the rightmost expression:

$$\sum_{\hat{Q} \in S_N} u(R_{r}^{\hat{Q}^{-1}})_{rs} R = \sum_{t=1}^f \left( \sum_{\hat{Q} \in S_N} u(R)_{rt} r^D(\hat{Q})_{ts} \right) R.$$

Compare coefficients of $R = (1)$:

$$u(Q^{-1})_{rs} = \sum_{t}^f u(1)_{rt} r^D(\hat{Q})_{ts}.$$

By definition $u(1)_{rt}$ is the coefficient of $(1)$ in the unit $y_{rt}$, also by definition this coefficient is $\xi^\lambda_{rt}$ and so:

$$u(Q^{-1}) = \chi^\lambda r^D(\hat{Q}).$$

---

Note.

In the same way one proves for the left-irrep $L^D(\hat{Q})$

$$L^D(\hat{Q})^T = u(Q^{-1}) \chi^{-1}.$$
5.2. Young orthogonal units

Since it is often advantageous to work with an orthogonal basis of \( CS_N \) which moreover carries orthogonal irreps, Young defined an alternative basis for \( CS_N' \), satisfying these two conditions. (Note that an orthogonal basis of a group algebra does not necessarily carry orthogonal representations!)

**Definition.**

The following recursive expression for an orthogonal Young unit \( \delta_{pq}^{[\lambda]} \) is due to Young:

\[
\delta_{pq}^{[\lambda]} = \frac{f_{pq}^{[\lambda]}}{N!} \cdot A_{pq}^{[\lambda]} \cdot \delta_{pp}^{[\lambda (N-1)]} \cdot y_{pq}^{[\lambda (N)]} \cdot \delta_{qq}^{[\lambda' (N-1)]}.
\]

Here:

\( [\lambda (N)] \) is an irrep of \( S_N' \),

\( [\lambda (N-1)] \) is the irrep of \( S_{N-1} \) with a diagram obtained from \( T_p^{[\lambda (N)]} \) by removal of the digit \( N \),

\( [\lambda' (N-1)] \) is the irrep of \( S_{N-1} \) obtained by removal of the digit \( N \) from \( T_q^{[\lambda (N)]} \),

\( y_{pq}^{[\lambda (N)]} \) is a Young unit,

\( A_{pq}^{[\lambda]} \) is a normalizing constant (see Rutherford §26 [36] for the computation).

Note: \( A_{pq}^{[\lambda]} \) in the NP-definition of \( y_{pq}^{[\lambda (N)]} \) is equal to \( (A_{pq}^{[\lambda]})^{-1} \) in the PN-definition.

**Example.**

\[
T_1^{[\lambda (4)]} = \begin{bmatrix} 1 & 2 & 3 \\ 4 \end{bmatrix}, \quad T_1^{[\lambda (3)]} = \begin{bmatrix} 1 & 2 & 3 \\ 4 \end{bmatrix}, \quad T_1^{[\lambda (2)]} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad T_1^{[\lambda (1)]} = 1
\]

\[
T_2^{[\lambda (4)]} = \begin{bmatrix} 1 & 2 & 4 \\ 3 \end{bmatrix}, \quad T_1^{[\lambda' (3)]} = \begin{bmatrix} 1 & 2 \\ 3 \end{bmatrix}, \quad T_1^{[\lambda' (2)]} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad T_1^{[\lambda' (1)]} = 1
\]

So:

\[
o_{11}^{[\lambda (1)]} \equiv o_{11}^{[1]} = (1);
\]

\[
o_{11}^{[\lambda (2)]} \equiv o_{11}^{[2]} = 1/2 \cdot (1 + (12));
\]

\[
o_{11}^{[\lambda (3)]} \equiv o_{11}^{[3]} = 1/6 \cdot (1 + (12) + (13) + (23) + (123) + (132));
\]

\[
o_{11}^{[\lambda' (3)]} \equiv o_{11}^{[2,1]} = 1/3 \cdot (1 + (12)) - 1/2 \cdot ((13) + (23) + (123) + (132));
\]

\[
o_{11}^{[\lambda (4)]} \equiv o_{11}^{[3,1]} = 3/24 \cdot 3/4\sqrt{2} \quad o_{11}^{[3]} \cdot o_{12}^{[3,1]} \cdot o_{12}^{[2,1]}, \quad \text{where} \ 3/4\sqrt{2} = A_{12}^{[3,1]},
\]

66
and so:
\[ o_{12}^{\lambda(4)} = -1/3 \sqrt{2} \left( 14 + 24 - 2 \cdot 34 + 124 - 2 \cdot 134 + 142 + 143 
- 2 \cdot 234 + 243 - 2 \cdot 1234 + 1243 + 1324 - 2 \cdot 1342 
+ 1423 + 1432 - 2 \cdot 12 + 34 + 13 + 24 + 14 \right). \]

The orthogonal Young units belonging to standard tableaux form a basis of \( S^N_\mathbb{R} \) corresponding to a Wedderburn decomposition of this group algebra. The following expression, here written in terms of orthogonal Young units, can be shown to hold in general for orthogonal units of semi-simple algebras (e.g. [8], p. 66):

\[ o_{pq}^{\lambda} = \frac{f_{[\lambda]}}{N!} \sum_{P \in S^N} D_{[\lambda]}^{P^{-1}} q_{qp}. \]

Here \( D_{[\lambda]}^{P^{-1}} \) is the matrix representation of \( S^N \) carried by the orthogonal units. The orthogonal Young units have, by virtue of the specific manner in which they were defined, the property that \( D_{[\lambda]}^{P^{-1}} \) is a real orthogonal matrix [36, § 27].

Denoting the trace of \( D_{[\lambda]}^{P^{-1}} \) by \( \chi_{[\lambda]}^{P^{-1}} \) it follows that

\[ o_{pq}^{[\lambda]} = \frac{f_{[\lambda]}}{N!} \sum_{q=1}^{N} o_{qq}^{[\lambda]} = \frac{f_{[\lambda]}}{N!} \sum_{P \in S^N} \chi_{[\lambda]}^{P^{-1}} q_{qp}. \]

So: \( o_{pq}^{[\lambda]} \) is a Wigner operator and \( o^{[\lambda]} \) is a character projector.

Theorem.
The diagonal units satisfy the recursion relation:

\[ o_{qq}^{[\lambda(N)]} = o_{qq}^{[\lambda(N-1)]} o_{qq}^{[\lambda(N-1)]}, \]

where \( o_{qq}^{[\lambda(N-1)]} \) belongs to the standard tableau obtained from \( T_{qq}^{[\lambda(N)]} \) by removing the number \( N \).

Proof. (The manner of proof is inspired by chapter VII of ref. 11).
First we note that the element

\[ o_{qq}^{[\lambda(N)]} o_{qq}^{[\lambda(N-1)]} \]

is a primitive idempotent.

The element is idempotent since the factors commute and both are idempotent. (Recall here that \( o_{qq}^{[\lambda(N)]} \) belongs to the center of \( S^N_\mathbb{R} \)).

The element is primitive because of the following reason: The simple algebra obtained by the action of \( o_{qq}^{[\lambda(N-1)]} \) onto \( S^N_{N-1} \).
\[ A\{\lambda(N-1)\} \equiv o\{\lambda(N-1)\} \quad (CS_{N-1}) \]

doncurs just once in the simple algebra:
\[ A\{\lambda(N)\} \equiv o\{\lambda(N)\} \quad (CS_N),\]
because the irrep \([\lambda(N-1)]\) occurs just once in the restriction of \([\lambda(N)]\) to \(S_{N-1}\). Of course \(A_{\{\lambda(N-1)\}}\) possesses only one q-th diagonal unit, and so the element \(o_{qq}\{\lambda(N)\} \{\lambda(N-1)\}\) belongs to a one-dimensional subspace of \(CS_N\) from which follows that it is primitive.

We can now write:
\[
o_{qq}\{\lambda(N)\} = o_{qq}\{\lambda(N-1)\} = o_{qq}\{\lambda(N)\} o_{qq}\{\lambda(N)\} o_{qq}\{\lambda(N-1)\}
\]
\[
= c \cdot o_{qq}\{\lambda(N)\} \{\lambda(N-1)\}, \quad c \in C.
\]
The first equal sign follows from the facts that \(o_{qq}\{\lambda(N)\}\) is a unit of \(A\{\lambda(N)\}\) and that \(y_{qq}\{\lambda(N)\}\) belongs to this algebra. The second equal sign follows (via theorem 3.9, chapter III of ref. 8) from the fact that:
\[ o_{qq}\{\lambda(N)\} o_{qq}\{\lambda(N-1)\} \]
is a primitive idempotent. The constant \(c\) is equal to 1 since both sides in the relation to be proved are idempotent.

Notes.

1. Continuing the recursion it follows that:
\[ o_{qq}\{\lambda(N)\} = o_{qq}\{\lambda(N)\} o_{qq}\{\lambda(N-1)\} o_{qq}\{\lambda(N-2)\} \ldots o_{qq}\{1\}, \]
so there exists a 1-1 correspondence between the index \(q\) and the set of irreps \(\{\{\lambda(N)\}, \{\lambda(N-1)\}, \ldots, \{1\}\}\), which forms the genealogy of the q-th standard Young tableau with shape \(\{\lambda(N)\}\).

2. The recursive expression in the foregoing note has been given earlier by Matsen [25], who, after having defined the diagonal units in this manner, proves that they are primitive idempotents. In this manner one cannot prove, however, that the diagonal units are the same as the Young orthogonal units.

5.3. Yamanouchi units

Apparently not aware of the work of Young, Yamanouchi [43] derived a set of orthogonal irreps of \(S_N\) by using a sort of vector coupling technique.
this section it will be proved, by purely algebraic means, that Young's orthogonal and Yamanouchi's representation are identical.

**Definition.**

A *Yamanouchi unit* \( W_{r,s}^{[\lambda]} \) is defined by:

\[
W_{r,s}^{[\lambda]} = \frac{f^{[\lambda]}}{N!} \sum_{P \in S_N} O^{[\lambda]}(P)_{r,s} \frac{1}{P},
\]

where the matrix \( O^{[\lambda]}(P) \) is by Yamanouchi's construction adapted to the sequence

\[ S_N \supset S_{N-1} \supset \ldots \supset S_1. \]

Hence Yamanouchi units are orthogonal.

To prove that \( W_{r,s}^{[\lambda]} \) is in fact a Young orthogonal unit, we first note that the diagonal units can be written thus:

\[
W_{r,r}^{[\lambda(N)]} = \sum_{q=1}^{f^{[\lambda(N-1)]}} W_{r,q}^{[\lambda(N)]} W_{q,r}^{[\lambda(N-1)]}
\]

(where \( W_{r,q}^{[\lambda(N)]} \) is a "reduced" Yamanouchi unit, see the appendix of ref. 44), and therefore:

\[
W_{r,r}^{[\lambda(N)]} W_{r,r}^{[\lambda(N-1)]} = \sum_{r=1}^{f^{[\lambda(N)]}} W_{r,r}^{[\lambda(N)]} W_{r,r}^{[\lambda(N-1)]} = \sum_{r,q} W_{r,q}^{[\lambda(N)]} W_{q,r}^{[\lambda(N-1)]} W_{r,r}^{[\lambda(N-1)]}
\]

From which follows that the diagonal Yamanouchi units satisfy the same recursion relation as the diagonal Young orthogonal units; hence the two sets of diagonal units are identical.

We now show that the same holds for the off-diagonal units. Write to that end, suppressing \([\lambda]\):

\[
W_{i,j} = \sum_{k,l} c_{k,l}^{i,j} o_{k,l}^{i,j}.
\]

Multiply on the left by \( W_{i,i} = o_{i,i} \) and on the right by \( W_{j,j} = o_{j,j} \):

\[
W_{i,i} W_{i,j} W_{j,j} = W_{j,j} = \sum_{k,l} c_{k,l}^{i,j} o_{i,k}^{i,j} o_{k,l}^{j,j} = c_{i,j}^{i,j} o_{i,j}^{i,j}.
\]

Both sets of units carry real orthogonal representations:

\[
W_{i,j}^+ = W_{j,i} \Rightarrow (c_{i,j} o_{i,j})^+ = c_{j,i} o_{j,i} \Rightarrow c_{i,j} = c_{j,i} \Rightarrow c_{j,i} \in \mathbb{C}.
\]

Further:
\[ o_{ii} = w_{ii} = w_{ij} w_{ji} = c_{ij}^2 o_{ij} o_{ji} = c_{ij}^2 o_{ii} \Rightarrow c_{ij} = \pm 1. \]

Now, it has already been pointed out by Yamanouchi himself [43, p. 422] that the requirement that \([\lambda] \) must be adapted to
\[ S_N \supset S_{N-1} \supset \cdots \supset S_1, \]
only determines \([\lambda] \) up to a similarity transformation with a diagonal matrix with elements \(\pm 1\) along the diagonal. This implies that the sign of the off-diagonal units cannot be fixed unambiguously. However, the choice: \(c_{ij} = +1\) yields the convention chosen by Yamanouchi. (I am indebted to Dr. D. J. Klein for drawing my attention to this freedom in sign).

Concluding we have found that Young orthogonal units and Yamanouchi units are identical. Henceforth they will be called Young-Yamanouchi units, or briefly Y-Y units.

5.4. Graphical representation of projected tensors

In the preceding three sections two possible complete decompositions of \(CS_N\) have been given. Explicit expressions for the corresponding basis elements, Young units and Y-Y units, have been derived. From the point of view of the mathematician the problem of decomposing tensor space under \(GL(n), U(n)\) and \(SU(n)\) has now been solved completely, as the irreducible \(GL(n)\)-carrier space \(W^\lambda_i\), defined in sec. I.4.6, is simply given by:
\[ W^\lambda_i = e^\lambda_{ii} (v_n \otimes^n), \]
where \(e^\lambda_{ii}\) stands for either \(Y^\lambda_{ii}\) (a Young unit) or \(W^\lambda_{ii}\) (a Y-Y unit).

However, projection is usually a very tedious procedure, because many times a zero vector or a vector linearly dependent on the vectors already projected will be found. So, in the next two sections rules will be derived to avoid this problem. To reach that aim a graphical representation of the projected tensors due to Boerner will be employed. For two-column tableaux this representation, defined below, is easily seen to be in one-to-one correspondence with the representations by extended Rumer diagrams [39, 45], by bracket structures [38, 40] and by path diagrams [46, p. 30]. The equivalence of these representations has been discussed by Van Berkel [47].
Definitions.

1. Associate an index tableau $I^{[\lambda]}$ with a (possibly non-standard) Young tableau $T^{[\lambda]}$ and an index set $I = \{i_1, i_2, \ldots, i_N\}$, by replacing $k$ in $T^{[\lambda]}_k$ by $i_k$, for $k = 1, 2, \ldots, N$.

2. An index tableau $I^{[\lambda]}$ obeying the following two rules is standard:
   (i) In the columns of $I^{[\lambda]}$, the indices increase from top to bottom.
   (ii) In the rows of $I^{[\lambda]}$, the indices do not decrease from left to right.

3. An index tableau that is standard, except for one or more columns containing equal indices, is half-standard.

4. An index tableau that is neither standard nor half-standard is called non-standard.

Examples.

1. $I = \{1, 7, 7, 8\}$
   $T^{[2,2]} = \begin{array}{c}
   3 \\ 2 \\ 1 \\ 4 
   \end{array} \Rightarrow I^{[2,2]} = \begin{array}{c}
   7 \\ 7 \\ 1 \\ 8 
   \end{array}$ (a non-standard index tableau).

2. $I = \{1, 1, 2, 3\}$
   $T^{[2,1^2]} = \begin{array}{c}
   1 \\ 2 \\ 3 \\ 4 
   \end{array} \Rightarrow I^{[2,1^2]} = \begin{array}{c}
   1 \\ 1 \\ 2 \\ 3 
   \end{array}$ (a standard index tableau).

3. $I = \{4, 4, 5, 6\}$
   $T^{[3,1]} = \begin{array}{c}
   1 \\ 3 \\ 4 \\ 2 
   \end{array} \Rightarrow I^{[3,1]} = \begin{array}{c}
   4 \\ 5 \\ 6 \\ 4 
   \end{array}$ (a half-standard index tableau).

Definition.

Let the index tableau $I^{[\lambda]}$ belong to the pair $\{I, T^{[\lambda]}\}$, then the tableau $\pi(I^{[\lambda]})$, $\pi \in S_N$, belongs to the pair $\{\pi(I), T^{[\lambda]}\}$.

Notes.

1. The action of $\pi$ on $I^{[\lambda]}$ is given in the conventional way (sec. I.4.4) after numbering the positions of $I^{[\lambda]}$ according to the numbers in $T^{[\lambda]}$.

2. Recall that the action of a permutation on a Young tableau is defined by its action on the numbers in the tableau, not on the positions. This implies that the pair $\{I, T^{[\lambda]}\}$ yields the same index tableau as the pair $\{\pi(I), \pi(T^{[\lambda]})\}$.

3. The pair $\{I, \pi^{-1}(T^{[\lambda]})\}$ yields $\pi(I^{[\lambda]})$. 

71
Example.

\( I = \{i, j, k, \ell\}, T^{[\lambda]} = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}, \pi = (123). \)

So:

\( \pi(I) = \{k, i, j, \ell\}, \pi^{-1}(T^{[\lambda]}) = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}, \pi(T^{[\lambda]}) = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}. \)

Now:

\{I, T^{[\lambda]}\} and \{\pi(I), \pi(T^{[\lambda]}))\} give \begin{bmatrix} k & 2 \\ j & 1 \end{bmatrix}

and:

\{\pi(I), T^{[\lambda]}\} and \{I, \pi^{-1}(T^{[\lambda]}))\} give \begin{bmatrix} j & 2 \\ 1 & k \end{bmatrix}.

In the foregoing sections several operators associated with two Young tableaux were introduced. Henceforth these operators will be called tableau operators. The following tableau operators are of special importance:

(i) The partial antisymmetrizer \( N^{[\lambda]}_{pk} \equiv \sigma^{[\lambda]}_{pk} N^{[\lambda]}_k \).

(ii) The Young unit \( Y^{[\lambda]}_{pk} \).

(iii) The Y-Y unit \( W^{[\lambda]}_{pk} \).

By means of an index tableau a tensor, resulting from \( E^{[i]}_I \equiv v^{[i]}_1 \otimes \ldots \otimes v^{[i]}_N \) by the action of a tableau operator, can be represented graphically.

Definition.

Let \( I^{[\lambda]}_k \) be the index tableau belonging to \( \{I, T^{[\lambda]}_k\} \), then we define the following projected tensors.

(i) A partially antisymmetric tensor:

\[ D^{(p)}_{I^{[\lambda]}_k} \equiv N^{[\lambda]}_{pk} E^{[i]}_I. \]

(ii) A Weyl-Rumer tensor:

\[ F^{(p)}_{I^{[\lambda]}_k} \equiv Y^{[\lambda]}_{pk} E^{[i]}_I. \]

(iii) A Gelfand tensor:

\[ G^{(p)}_{I^{[\lambda]}_k} \equiv W^{[\lambda]}_{pk} E^{[i]}_I. \]

Example.

Let the index tableau \( \begin{bmatrix} j & i \\ k & \ell \end{bmatrix} \) belong to \( I = \{i, j, k, \ell\} \) and \( T^{[2,2]}_2 = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \).
Then:

\[
\begin{align*}
\mathcal{D}(p) & \equiv \frac{\mathcal{N}}{p^2} \mathcal{E}_{ijkl} = \frac{c}{p^2} \mathcal{N}^{2} \mathcal{E}_{ijkl} \\
& = \frac{c}{p^2} [2,2] \frac{(1) - (23)}{(1) - (14)} \mathcal{E}_{ijkl} \\
& = \frac{c}{p^2} (\mathcal{E}_{ijkl} - \mathcal{E}_{ikjl} - \mathcal{E}_{klij} - \mathcal{E}_{lkji})
\end{align*}
\]

for certain \( p, 1 \leq p \leq 24 \), numbering the Young tableaux, including the non-standard ones.

Notes.

1. The action of tableau operators onto tensor components can also be represented by index tableaux. Thus, for instance:

\[
\mathcal{I}_{[\lambda]}^{[\mu]}_{(p)} = \mathcal{Y}_{kp}^{[\lambda]} \mathcal{T}^{[\mu]}
\]

Here we let the first index of the tableau operator label the index tableau.

2. Note that in general:

\[
\mathcal{F}_{[\lambda]}_{(p)} \neq \mathcal{F}_{(p)}_{[\lambda]}^{[\mu]}, \quad \text{and also} \quad \mathcal{F}_{[\lambda]}_{(p)} \neq \mathcal{F}_{(p)}^{[\mu]}_{[\lambda]},
\]

but that:

\[
\mathcal{I}_{[\lambda]}^{[\mu]}_{(p)} = \mathcal{S}_{(p)}^{[\lambda]}_{[\mu]},
\]

which can be proved as follows (suppressing \([\lambda]\)):

\[
\begin{align*}
\mathcal{I}_{(p)}^{[\lambda]} & \equiv \mathcal{Y}_{kp}^{[\lambda]} \mathcal{T}^{[\mu]} = \sum_{p \in \mathcal{U}(p)_{kp}} \mathcal{F}^{[\mu]}_{(p)} \mathcal{T}^{[\mu]}_{(p)} = \sum_{p \in \mathcal{U}(p)_{kp}} \mathcal{P}^{[\mu]}_{(p)} \mathcal{T}^{[\mu]}_{(p)}, \\
& \equiv \mathcal{S}_{(p)}^{[\lambda]}_{[\mu]}.
\end{align*}
\]

5.5. Weyl-Rumer basis

In this section it will be shown that the collection of Weyl-Rumer tensors labelled by standard index tableaux forms a complete and linearly independent basis of \( V_n \otimes^n N \). It seems that no earlier general proof of this result has been given, although the equivalent case of PN-projected tensors has been treated by Boerner [8] and independently also by Seligman [48]. The case
of NP-projected tensors is more difficult, however. Since the well-known proof of Rumer et al. [45] depends essentially on the fact that only one- and two-column tableaux are considered, this manner of proof cannot be generalized to general Young tableaux. The proof given in this section is an adaptation and extension of Boerner's proof.

Lemma.

(i) The partially antisymmetric tensor:
\[ D(p) \]
\[ I_k \]

is zero if two or more indices in one column of \( I_k \) are equal.

(ii) Every non-zero partially antisymmetric tensor is equal (up to sign) to a partially antisymmetric tensor with a column-ordered index tableau (this is an index tableau with the indices increasing in the columns from top to bottom).

(iii) The set of partially antisymmetric tensors with \( p, [\lambda] \) and \( I \) fixed and \( k \) running over column-ordered index tableaux is linearly independent and contains non-zero tensors only.

Proof. Suppress superscript \([\lambda]\).

\[ D(p) \equiv N_{pk} E = N_p \sigma_{pk} E = N_p E_J = D(p) \]

where we have written \( J = \sigma_{pk}(I) \).

Note furthermore that \( \{I, T_k\} \) and \( \{\sigma_{pk}(I), \sigma_{pk}(T_k)\} \) both give \( I_k \), and that \( \{\sigma_{pk}(I), \sigma_{pk}(T_k)\} = \{J, T_p\} \), which gives \( J_p \). So \( I_k = J_p \). Because \( N_p \) is a product of antisymmetrizers, one for each column, the tensor \( D(p) \) is a product of antisymmetric tensors. Now the following points immediately follow from the standard theory of antisymmetric tensors (e.g. ref. 12, ch. VIII):

a. If, and only if, two or more equal indices occur in the same column of \( J_p = I_k \) the tensor \( D(p) \) is zero.

b. Two partially antisymmetric tensors that are connected by a vertical permutation \( \nu \in \mathcal{C}_p \) are equal (up to the parity \( \zeta_\nu \)).

c. Two tensors which are not connected by a vertical permutation are linearly independent. (To make this plausible one can assume the basis of \( V_n \) to be orthogonal. The two tensors are then also orthogonal, and hence linearly independent).

The remainder of the proof follows trivially from these points.
Corollary 1.
The tensor space $V_n \otimes^N$ does not carry irreps of $CS_N$ with more than $n$ rows.

Proof. Express a Weyl-Rumer tensor in terms of partially antisymmetric tensors:

$$ F^{(p)}_{I_k} = Y^{[\lambda]}_{pk} E_{I_k} = N^{[\lambda]}_{pk} F^{[\lambda]}_{k} E_{I_k} = \sum_{\pi \in R_k^{[\lambda]}} N^{[\lambda]}_{pk} E_{\pi(I_k)} = \sum_{\pi \in R_k^{[\lambda]}} D^{(p)}_{\pi(I_k)} \pi^{[\lambda]}_{k}.$$

If the number of different indices in $I$ is less than the number of rows of $[\lambda]$, all index tableaux $\pi(I_k), \pi \in R_k^{[\lambda]}$, have necessarily two or more equal indices in one column; in that case the Weyl-Rumer tensor at the left is zero. The maximal number of different indices in any index set $I$ labelling an element $E_I$ of $V_n \otimes^N$ is $n$. This proves the corollary.

Note.
The irrep $[1^N]$ is the irrep of $CS_N$ with the most rows in its Young diagram. Therefore, if $N > n$ not all irreps of $CS_N$ can be carried by $V_n \otimes^N$ and the representation $CS_N \rightarrow AS_N$ is non-faithful. (See page 150 of ref. 8 for a proof that the representation is faithful if $N \leq n$).

Corollary 2.
Let $[\lambda]$ have $m$ columns, and let the index tableau $I_k^{[\lambda]}$ belong to $\{I, T_k^{[\lambda]}\}$. Then $F^{(p)}_{I_k^{[\lambda]}}$ is zero whenever any index in $I$ occurs more than $m$ times.

Proof. Express $F^{(p)}_{I_k^{[\lambda]}}$ again in partially antisymmetric tensors. Since one of the indices occurs more than $m$ times, this index appears at least twice in one of the columns of all the partially antisymmetric tensors in the expansion. So they are all zero, and accordingly the W-R tensor vanishes.

Note.
These corollaries give rise to the Pauli principle in its original form, stating that an orbital may not be occupied by more than two fermions of intrinsic spin 1/2. We come back to this point in the next chapter.

Let $U$ be a subspace of $V_n \otimes^N$ belonging to a certain index set $I$; that is, $U$ is spanned by:

$$ \{PE_I | \text{all } P \in S_N \}. $$
The space $U$, being stable under $S_N'$, can be decomposed in the usual way into a direct sum of irreducible $S_N$ carrier spaces. If $Y_{pq}^{[\lambda]} E_J \neq 0$, for some $q$ and $J$, with $E_J \in U$, then:

$$\{Y_{pq}^{[\lambda]} E_J \mid p = 1, \ldots, f[\lambda]\}$$

(1)

spans an irrep of $S_N$. Note that all elements of this basis are characterized by the same index tableau $J_{q}^{[\lambda]}$. So, this index tableau can serve as a label for the space spanned by this basis, and we write accordingly $U_{J_{q}^{[\lambda]}}$ for this subspace of $U$. By virtue of the following lemma this association of an irreducible carrier space of $S_N$ with an index tableau is unique.

**Lemma.**

Let $I_{p}^{[\lambda]}$ belong to the pair $\{I_{p}^{[\lambda]}, I_{q}^{[\lambda]}\}$ to $\{I_{p}^{[\lambda]}, I_{q}^{[\lambda]}\}$, $p \neq q$. The Young tableaux $T^{[\lambda]}_{p}$ and $T^{[\lambda]}_{q}$ are not necessarily standard. Then:

$$I_{p}^{[\lambda]} = I_{q}^{[\lambda]} \implies F_{p}^{(k)} = F_{q}^{(k)}$$

for all $k, 1 \leq k \leq N!$

**Proof.** Drop $[\lambda]$.

$\{I_{p}^{[\lambda]}, \sigma_{pq}(I_{p}^{[\lambda]})\}$ gives the same index tableau $I_{q}^{[\lambda]} = I_{p}^{[\lambda]}$. Also, since $\sigma^{-1}_{pq} = \sigma_{pq}$, $\{I_{p}^{[\lambda]}, \sigma_{pq}(I_{p}^{[\lambda]})\}$ gives the same index tableau as $\{\sigma_{pq}(I_{p}^{[\lambda]}), I_{q}^{[\lambda]}\}$, (see the preceding section). So $\{I_{p}^{[\lambda]}, \sigma_{pq}(I_{p}^{[\lambda]}), I_{q}^{[\lambda]}\}$ both give $I_{p}^{[\lambda]}$; hence $I = \sigma_{pq}(I)$.

Then:

$$F_{p}^{(k)} = F_{q}^{(k)} = Y_{kp} E_{I}^{(k)} = Y_{qp} \sigma_{pq}(I) = Y_{kp} E_{I}^{(k)} = F_{q}^{(k)}$$

**Example.**

$T_{1} = \begin{bmatrix} 1 & 2 \\ 3 \end{bmatrix}$, $T_{2} = \begin{bmatrix} 1 & 3 \\ 2 \end{bmatrix}$, $I = \{1, 2, 2\}$, $I_{1} = I_{2} = \begin{bmatrix} 1 & 2 \\ 2 \end{bmatrix}$.

From the lemma:

$$Y_{11} E_{122} = Y_{12} E_{122} = F_{1}^{(1)} \begin{bmatrix} 1 & 2 \\ 2 \end{bmatrix}$$

$$Y_{21} E_{122} = Y_{22} E_{122} = F_{2}^{(2)} \begin{bmatrix} 1 & 2 \\ 2 \end{bmatrix}$$
Notes.
1. This lemma lays the mathematical foundation under the use of index tableaux for labelling projected tensors. The same proof can be given for partially antisymmetric tensors; the same result will also be shown for Gelfand tensors in the next section.
2. The association of index tableaux and projected tensors with fixed superscript $k$ is not one-to-one, as the converse of the lemma does not hold. For instance:

$$F^{(k)}_{I_p} = F^{(k)}_{\pi(I_p)'}, \quad \forall \pi \in \mathbb{P},$$

while in general: $I_p \neq \pi(I_p)$.

We are now ready to state and prove the main theorem of this section.

Theorem.

Let $U \subseteq V_n \otimes^N \Theta^N$ belong to the index set $I$. $U$ can be decomposed into a direct sum of irreducible $S^N$ carrier spaces:

$$U = \bigoplus_{[\lambda]} \bigoplus_{q} \bigoplus_{I^q_{\lambda}} U_{[\lambda]},$$

where $q$ runs over all possible standard index tableaux of shape $[\lambda]$ and index set $I$.

Proof. Without loss of generality $U$ can be taken to be generated by an ordered index set $I_o$. That is, we assume $U$ to be spanned by $\{PE^{I_o}_{I_o}\}$,

$I_o = \{i_1 < i_2 < \ldots < i_N\}$.

The set of all W-R tensors characterized by standard or half-standard index tableaux forms a (generally overcomplete) basis of $U$. This follows since for any $E^{I_o}=\sqrt{1}PE^{I_o}=\sqrt{1}E^{I_o}$ we have:

$$E^{I_o} = (1)PE^{I_o} = (\sum_{\lambda} \sum_{p=1}^{f[\lambda]} \sum_{s=1}^{f[\lambda]} \sum_{p=1}^{f[\lambda]} \frac{\lambda}{N!} \frac{\lambda}{p_s} \eta_{s_p}) PE^{I_o} = \sum_{p,q} \sum_{c_{qp}} [\lambda] E^{I_q}_{I_o},$$

where we introduced the resolution of identity afforded by Young units (sec. I.5.1) and used the fact that Young units span $f[\lambda]$-dimensional right ideals (sec. I.5.1) and where we further introduced:

$$c_{qp} = \frac{f[\lambda]}{N!} \frac{f[\lambda]}{R_{(p)}} \eta_{s_{sp}}.$$
is standard, this index tableau is either standard or half-standard.

Let us now assume that the index tableaux are ordered such that $q \leq \ell$ first runs over the $q^{[\lambda]}$ standard index tableaux and then over the $f^{[\lambda]} - q^{[\lambda]}$ half-standard index tableaux of shape $[\lambda]$ and index set $I_o$. If we can show that the set:

$$\{ P^{(p)}_I | q = 1, ..., q^{[\lambda]} ; \text{fixed } p \}$$

forms a complete and linearly independent basis for the space $Y_{pp}^{[\lambda]} (U)$, we have proved the theorem. (Note that if we had used the PN-type Young units, completeness would already have been proved at this point, because in that case all half-standard tensors vanish).

The manner of proof is the following:

(i) First it is shown that a projected tensor $Y_{pp}^{[\lambda]} T$, arbitrary $T \in U$, does not have more than $q^{[\lambda]}$ linearly independent components, from which follows that the dimension of $Y_{pp}^{[\lambda]} (U)$ is less than or equal to $q^{[\lambda]}$.

(ii) Then it is shown that the set (2), which is manifestly a subset of $Y_{pp}^{[\lambda]} (U)$, is of dimension $q^{[\lambda]}$.

An arbitrary tensor $T \in U$ can be written as:

$$T = \sum_{Q \in S} E_Q^p (I_o)$$

and the corresponding projected tensor (suppressing superscript $[\lambda]$) as:

$$Y_{pp}^{[\lambda]} (T) = \sum_{Q \in S} E_Q^p (I_o)$$

where $s^{(p)}$, the component of $Y_{pp}^{[\lambda]} (T)$ along $E_Q^p (I_o)$, is labelled by the index tableau $Q(I_o)$. This index tableau can be standard, half-standard or non-standard. We now show that every projected tensor component can be expressed as a linear combination of the $q^{[\lambda]}$ components with standard index tableaux only. Suppressing subscript $p$, we write:

$$s^{(p)} = \sum_{Q \in S} E_Q^p (I_o)$$

The index tableau $I_q$ belongs to the ordered index set $I_o$ and the standard
Young tableau $T_q$, hence $I_q$ is either standard or half-standard. We now show that $s_{q} = 0$, if $I_q$ is half-standard, that is, if $I_q$ has two equal indices in one column. Let these indices be $i_k$ and $i_k'$, and so $(k,k')I_q = I_q$.

Now:

$$s_{q} = s_{(k,k')I_q} = s_{(k,k')I_q} = -s_{I_q},$$

since $(k,k')$ belongs to $C_q$ and has parity $-1$.

Knowing that the half-standard components vanish, it follows that $s_{I_q}$ can be expressed in terms of standard components only. This proves point (i).

Remark: It may seem surprising that half-standard components vanish, whereas half-standard tensors do not. Notice that we have seen this before in an example (sec. I.4.4), where

$$Y_{11}E_{121} \equiv F_{\begin{array}{c} 1 \\ 2 \\ 1 \end{array}}$$

and $Y_{11}t_{121} \equiv s_{\begin{array}{c} 1 \\ 2 \\ 1 \end{array}}$ were explicitly worked out.

We now show that the $g_{[\lambda]}$ different standard Weyl-Rumer tensors $F_{I_q}$ are linearly independent. We do this by expanding them in the linearly independent set of partially antisymmetric tensors $D_{I_k}$, introduced in the beginning of this section. First we order the standard W-R tensors in the following manner: Read down the columns of two standard index tableaux $I_k$ and $I_l$ simultaneously, starting at the leftmost column. Let $i \in I_k$ and $j \in I_l$ be the first indices encountered that are different. Now if $i < j$, then $F_{I_k}$ comes before $F_{I_l}$, and vice versa.

Secondly we order the set $\{D_{I_k}\}$, such that the $g_{[\lambda]}$ standard tensors among them come first and are ordered in the same way as the standard W-R tensors. The ordering of the non-standard partially antisymmetric tensors is irrelevant. For the case $I_o = \{1, 2, 3, 4\}$ we have for instance the following ordering:

$$F_{\begin{array}{c} 1 \\ 4 \\ 2 \\ 3 \end{array}} < F_{\begin{array}{c} 1 \\ 3 \\ 2 \\ 4 \end{array}} < F_{\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}}$$

and

$$D_{\begin{array}{c} 1 \\ 4 \\ 2 \\ 3 \end{array}} < D_{\begin{array}{c} 1 \\ 3 \\ 2 \\ 4 \end{array}} < D_{\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}} < D_{\begin{array}{c} 2 \\ 1 \\ 3 \\ 4 \end{array}}$$

Using this ordering we shall show below that a set of linear equations of
the following general appearance holds:

\[
(F^{I_1, \ldots, I_k}_{I_1, \ldots, I_k}) = (D^{I_1, I_2}_{I_1, I_2}, \ldots) \begin{pmatrix}
1 \\
1 \\
\vdots \\
A \\
\vdots \\
0 \\
\vdots \\
B \\
- & - & - & - & - & 1
\end{pmatrix}
\]

The \(q[I] \times q[I]\) submatrix in the upper part of the matrix is upper triangular and its diagonal elements are 1. Clearly the total matrix has \(q[I]\) linearly independent columns, representing the standard W-R tensors in the basis \(\{D^{I_k}_{I_k}\}\). So, the \(q[I]\) standard W-R tensors are linearly independent, which proves the theorem.

It remains to show that the linear equations have this specific form. Expand to that end the standard W-R tensors (suppressing superscript \(p\)):

\[
F^{I_k}_{I_k} = \sum_{\pi} \sum_{\mathcal{R}_k} N^{I_o}_{I_o} E^{\pi(I_o)}_{\mathcal{R}_k} = \sum_{\mathcal{R}_k} D^{\pi(I_k)}_{\mathcal{R}_k}.
\]

Replace the possibly occurring non-column-ordered tensors in the rightmost expression by column-ordered ones via:

\[D^{\pi(I_k)}_{\mathcal{R}_k} = \xi_{\nu} D^{\nu\pi(I_k)}_{\mathcal{C}_k}, \quad \nu \in \mathcal{C}_k.\]

Look in the standard index tableau \(I_k\) for the first column that changes under the permutation \(\nu\pi\), and find in this column the highest element \(i\) that is replaced (by an element \(i''\)) under the action of \(\nu\pi\). The following reasoning shows that \(i''\) is necessarily larger than \(i\): The operation \(\pi\) replaces \(i\) by \(i'\), \(i \preceq i'\), \(\nu\) replaces \(i'\) by an element \(i''\) standing below \(i'\). There are two possibilities: either \(i''\) has not been moved by \(\nu\pi\) and then immediately: \(i'' > i\), or \(i''\) has been placed there by \(\pi\), and then it must have come from the right (remember we are looking at the leftmost column that changes) and so \(i'' \succeq i'\). Since \(i' \succeq i\), we have \(i'' \succeq i\). The tableau \(\nu\pi(I_k)\) is either standard and then comes after \(I_k\) because \(i'' > i\), or is non-standard and then comes after the last standard one. This proves the appearance of zeros below the diagonal of the upper part of the matrix. The diagonal elements finally follow from the fact that the identity (1) belongs to \(\mathcal{R}_k\), which implies that a standard W-R tensor \(F^{I_k}_{I_k}\) has the component \(+1\) along the standard partially antisymmetric tensor \(D^{I_k}_{I_k}\) with the same index tableau \(I_k\).
Example.
The index set \{1, 1, 2, 3, 4\} gives a 5!/2! dimensional space \(U\), which decomposes under \(S_N\) as follows:

\[
U = U_{11234} \oplus U_{1123} \oplus U_{1124} \oplus U_{1134}
\]

\[
\oplus U_{11} \oplus U_{23} \oplus U_{24} \oplus U_{34}
\]

\[
\oplus U_{112} \oplus U_{113} \oplus U_{114} \oplus U_{12}
\]

The space \(U\) has the dimension:

\[
1\times1 + 3\times4 + 2\times5 + 3\times6 + 3\times6 + 1\times4 = 60.
\]

The decomposition of the total tensor product space \(V_n \otimes^N\) follows easily by letting the index set \(I\) run over all essentially different sets (that is, sets which do not follow from each other by permutation), and then decomposing the associated spaces according to the theorem just given. Clearly tensors belonging to different index sets are linearly independent and so we obtain a complete reduction of \(V_n \otimes^N\) under \(S_N\) in this manner.

Example.
It is illustrative to look again at the decomposition of the four-electron spin space \(V_4 \otimes^4\). In sec. I.4.7 we discussed this problem from the point of view of SU(2), we now approach from \(S_4\).

The possible ordered index sets and the associated irreducible spaces are:

\[
\{1, 1, 1, 1\}: \quad U_{1111}
\]

\[
\{1, 1, 1, 2\}: \quad U_{1112} \oplus U_{111} \oplus U_{2}
\]

\[
\{1, 1, 2, 2\}: \quad U_{1122} \oplus U_{112} \oplus U_{11} \oplus U_{22}
\]

\[
\{1, 2, 2, 2\}: \quad U_{1222} \oplus U_{122} \oplus U_{22} \oplus U_{2}
\]

81
\{2, 2, 2, 2\}: \quad U \begin{bmatrix} 2 & 2 & 2 & 2 \end{bmatrix}

As in sec. I.4.7 we can assemble the basis vectors (standard W-R tensors) in the following table. From the theorem of sec. I.4.6 we know that these W-R tensors are eigenfunctions of the spin angular momentum operator $S^2$. Anticipating the association of the spin quantum number $S$ with the shape $[\lambda]$ (chapter I.6), the spaces are labelled by $S$.

\begin{align*}
V'_{S=2} & : \quad F^{(1)} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad F^{(1)} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \quad F^{(1)} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \\
V''_{S=1} & : \quad F^{(2)} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad F^{(2)} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \quad F^{(2)} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \\
V''_{S=1} & : \quad F^{(3)} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad F^{(3)} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \quad F^{(3)} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \\
V'_{S=0} & : \quad F^{(1)} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad F^{(1)} \begin{bmatrix} 1 & 2 \end{bmatrix} \\
V''_{S=0} & : \quad F^{(2)} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad F^{(2)} \begin{bmatrix} 1 & 2 \end{bmatrix}
\end{align*}

For instance $F^{(1)}$ reads explicitly:

\begin{align*}
F^{(1)} & = Y^{[3,1]}_{11} E_{1122} = 2(E_{1122} - E_{2121}) + 2(E_{1212} - E_{2211}) \\
& = 2(\alpha\beta\gamma - \beta\alpha\gamma) + 2(\alpha\beta\gamma - \beta\gamma\alpha),
\end{align*}

where the last line is in terms of the usual spin functions.

Notes.

1. In ref. 49, reprinted in this thesis, the graphical representation of W-R tensors is employed to classify the possible Russell-Saunders states of an atom.

2. Of course, now that we have decomposed $V_n \otimes^N$ under $S_N$, we have done the
same under $\text{GL}(n)$, $\text{U}(n)$, $\text{SU}(n)$ and their respective Lie algebras. This follows from the theorem in sec. I.4.6.

5.6. Gelfand basis

Because Young units have a very simple appearance, a Weyl-Rumer basis of $V_n \otimes^N$ is easily constructed. However, the basis has the definite drawback of being non-orthogonal. On the other hand a Gelfand basis, obtained by projection with $Y-Y$ units, is orthogonal, but has the disadvantage of being rather difficult to generate; for we have seen that a $Y-Y$ unit contains in general $N!$ permutations, all with a coefficient that cannot so easily be computed. Still, Gelfand bases find widespread application, not only in theoretical physics, but also in quantum chemistry.

The purpose of this section is to prove that the set of Gelfand tensors with standard index tableaux forms a complete linearly independent basis of $V_n \otimes^N$, just as is the case with W-R tensors.

Theorem 1.
The set Gelfand tensors:

$$\{ G_p^{(\lambda)}_{\lambda} \mid k \text{ running over standard tableaux, } p\}, [\lambda] \text{ fixed}\$$

spans an irrep $<\lambda>$ of $\text{GL}(n)$, $\text{U}(n)$ and $\text{SU}(n)$ and their respective Lie algebras. Letting furthermore $<\lambda>$ run over all diagrams with at most $n$ rows and $p$ from 1 to $f[\lambda]$, the set thus obtained forms a complete and linearly independent basis of $V_n \otimes^N$.

Proof. Suppress superscripts on the Gelfand tensors.
Following an idea of Matsen [25] we employ the recursion relation for $Y-Y$ units given in sec. I.5.2. By induction it will first be shown that the Gelfand tensor:

$$G_{\lambda}^{(\lambda)} = \sum_{\gamma} \sum_{\gamma} E_{\lambda}^{(\gamma)}$$

is non-zero if, and only if, the index tableau $I_p^{(\lambda)}$ associated with $(I_o, T_p^{(\lambda)})$ is standard. Further it will be proved that if the standard index tableau $I_p^{(\lambda)}$ associated with $(I_o, T_p^{(\lambda)})$ is equal to the index tableau $I_k^{(\lambda)}$ associated with $(I_o, T_k^{(\lambda)})$, $k \neq p$, then the corresponding Gelfand tensors are identical.
Remove, to prove these points, the parentheses from the inside out in the following expression:
\[ \begin{array}{c} \lambda \\ w_{pp}^{\lambda} \end{array} E_{I_O} = (w^{\lambda(N)} (w^{\lambda(N-1)} \ldots (w^{\lambda(1)} v_{i_1}^1) \otimes \ldots \otimes v_{i_{N-1}}^{i_{N-1}}) \otimes v_{i_N}^{i_N}). \]

Suppose that after removal of \( r \) pairs of brackets the following non-zero standard tensor has been obtained:
\[ G_1 \equiv G_{\lambda(r)}^{I}_{Q} = (w^{\lambda(r)} (w^{\lambda(1)} v_{i_1}^1) \otimes \ldots \otimes v_{i_r}^{i_r}) \]
\[ = w^{\lambda(r)}_{Qq} (v_{i_1}^q \otimes \ldots \otimes v_{i_r}^{i_r}). \]

Suppose also that the next \( s \) indices are equal:
\[ i_r < i_{r+1} = i_{r+2} = \ldots = i_{r+s} < i_{r+s+1} \]
for certain \( s, 1 \leq s \leq N-r \). (Because in this induction process all equal indices are handled simultaneously, the possibility \( i_{r+1} = i_r \) is excluded).

Removing the next \( s \) pairs of brackets we in fact consider the construction of:
\[ G_{\lambda(r+s)}^{I} = w^{\lambda(r+s)}_{Q'q'} (G_1 \otimes G_2) \]
where \( G_2 \) is the symmetric tensor \( v_{i_{r+1}}^{i_{r+1}} \otimes \ldots \otimes v_{i_{r+s}}^{i_{r+s}} \). The tensor product \( G_1 \otimes G_2 \) belongs to an irrep \( [\lambda(r)] \otimes [s] \) of \( S_r \otimes S_s \). From it and its partners one can generate a space \( \overline{V} \) by the action of the left coset generators of \( S_r \otimes S_s \) in \( S_{r+s} \). \( \overline{V} \) carries the induced representation:
\[ \Lambda = ([\lambda(r)] \otimes [s]) \uparrow S_{r+s}, \]
which can be decomposed into a direct sum of irreps of \( S_{r+s} \) (ref. 4, p. 135). From the results in the appendix of ref. 44 it follows that \( G_1 \otimes G_2 \) has a component in every one of the irreducible \( S_{r+s} \)-carrier spaces contained in \( \overline{V} \). So the tensor:
\[ G_{\lambda(r+s)}^{I} \]
is non-zero only if \( [\lambda(r+s)] \) is contained at least once in \( \Lambda \). The tensor is furthermore uniquely defined if \( [\lambda(r+s)] \) occurs just once in \( \Lambda \), which implies that if another \( X-Y \) unit \( w_{p'p}^{\lambda(r+s)} \), \( p' \neq q' \), projects also a non-zero tensor out of \( G_1 \otimes G_2 \), it gives the same result.

Invoking now Littlewood's rule (ref. 50, p. 92) we find that \( \Lambda \) contains all the irreps corresponding to Young diagrams that can be built by the addition of \( s \) boxes to \( [\lambda(r)] \), no two boxes being added to the same column. All these irreps occur with multiplicity 1. So, if \( [\lambda(r+s)] \)
originates from $[\lambda(r)]$ in this manner, then the Gelfand tensor with an index tableau $I_{q'}^\lambda(r+s)$, obtained from $I_q^\lambda(r)$ by adding $s$ boxes containing the equal indices $i_{r+1}, \ldots, i_{r+s}$ to its boundary, is non-vanishing and uniquely determined. We clarify this by an example:

Let $W$ be a Y-Y unit belonging to the standard Young tableau:

$$
\begin{array}{ccc}
1 & 2 & 5 \\
3 & 4 \\
6 & 7
\end{array}
$$

then the tensor

$$
W(G_{\begin{array}{cc}
ij \\
k\ell
\end{array}} \otimes G_{\begin{array}{cc}
aa & a \\
k\ell \\
a & a
\end{array}}) = G_{\begin{array}{cc}
ij \\
aa & a
\end{array}}
$$

arises from $E_{I_o}$, where $I_o = \{ i \leq j \leq k \leq \ell < a = a = a \}$.

Note that the Y-Y units belonging to the standard tableaux:

$$
\begin{array}{ccc}
1 & 2 & 6 \\
3 & 4 \\
5 & 7
\end{array} \quad \text{and} \quad \begin{array}{ccc}
1 & 2 & 7 \\
3 & 4 \\
5 & 6
\end{array}
$$

give identical results when acting on $E_{I_o}$.

Continuing the removal of brackets, one ends up with either a zero-tensor or a tensor with a standard index tableau. Varying $p$ from 1 to $f^{\lambda}$ and $[\lambda]$ over all possible shapes one obtains a set standard Gelfand tensors, belonging to $I_o$. The tensors with different index tableaux are linearly independent by the construction just described; the tensors with the same index tableaux are identical. The partners in irreps of $S_N$ of the tensors that are obtained in this way by the action of the diagonal Y-Y units can be generated by the corresponding off-diagonal Y-Y units. Retaining now only the $S_N$-carrier spaces characterized by different standard index tableaux we have obtained a complete and linearly independent basis for the space spanned by

$$
\{ E_{P(I_o)} \mid \text{all } P \in S_N \}.
$$

Completeness follows immediately from the fact that the Y-Y units span right ideals. Varying finally $I_o$ over all ordered index sets, a complete linearly independent basis of $V_n \otimes^N$ adapted to a decomposition under $S_N$ is obtained.

Collecting all standard Gelfand tensors transforming as the $k$-th row of $[\lambda], k = 1, \ldots, f^{\lambda}$, we obtain by the result of sec. I.4.6 a basis for the irrep $\langle \lambda \rangle$ shared by GL(n), U(n) and SU(n). This basis carries also irreps of the respective Lie algebras (sec. I.3.1).
Often Gelfand tensors are represented \([30, 31]\) in a manner which visualizes their behaviour upon subduction under the chain:

\[
\text{GL}(n) \supset \text{GL}(n-1) \supset \ldots \supset \text{GL}(1).
\]

In order to be consistent, we derive this well-known representation from the definition given earlier (sec. I.5.4), and to that end we need Weyl's branching theorem \([\text{ref. 6, p. 391}]\). It is convenient to denote from here on the irreps of \(\text{GL}(n)\) by partition symbols \(<\mu>\) in which all the \(n\) row lengths are given explicitly, even the ones of length zero.

---

**Theorem 2** (Weyl's branching theorem).

Upon subduction to \(\text{GL}(n-1)\) the irrep \(<\mu>\) of \(\text{GL}(n)\) decomposes in the following manner:

\[
<\mu>_n = \sum_{t=0}^{N} \sum_{S} \Theta (<\mu>_n-1 \otimes <t>),
\]

where:

\(<t>\)

is the symmetric and only tensor representation of \(\text{GL}(1)\).

\(<\mu>_n-1 > = \sum_{m_1, n-1', m_2, n-1', \ldots, m_n, n-1} > runs over all partitions of \(N-t, n, \text{subject to the constraint: } m_{j+1, n} < m_{j, n-1} < m_{j, n} \text{ for } j = 1, 2, \ldots, n-1\). The partition \(<\mu>_n-1 >\) denotes an irrep of \(\text{GL}(n-1)\).

**Proof.** Formula 29 of ref. 51 yields for this case:

\[
<\mu>_n = \sum_{t=0}^{N} \sum_{S} \Theta g_{\mu}^{\mu_n}_n n-1 t (<\mu>_n-1 \otimes <t>),
\]

where we have written \(g_{\mu}^{\mu_n}_n n-1 t\) for the multiplicity of \(<\mu>_n-1 > \otimes <t>\) in \(<\mu>_n >\), and \(<\mu>_n-1 >\) runs over all partitions of \(N-t\). This follows because \(<t>\) is the only tensor representation of \(\text{GL}(1)\). The multiplicity factors are the same as the ones occurring in the induction:

\[
([\mu]_n-1 \otimes [t]) \uparrow S_n = \sum_{\mu} \Theta g_{\mu}^{\mu_n}_n n-1 t [\mu]_n
\]

as also has been discussed in ref. 51.

Invoking Littlewood's rule \([50, \text{p. 92}]\) it follows that:

\(g_{\mu}^{\mu_n}_n n-1 t = 1\) if \([\mu]_n\) can be constructed from \([\mu]_n-1\) by addition of \(t\) boxes to \([\mu]_n-1\) without adding more than one box to a column.

\(g_{\mu}^{\mu_n}_n n-1 t = 0\) in all other cases.

It is now easy to see that \(g_{\mu}^{\mu_n}_n n-1 t = 1\) if:

\[
\sum_{k=1}^{t} m_k = n - t
\]
\[ m_{j+1,n} \leq m_{j,n-1} \leq m_{jn} \]

for all \( j = 1, \ldots, n-1 \).

**Example.**

Subduction to GL(3) of the irrep \(<3,2,1,0>\) of GL(4) carried by \( V_4 \otimes 6 \):

\[
<3,2,1,0> = (\langle 3,2,1 \rangle \otimes <0> \oplus (\langle 2,2,1 \rangle \oplus \langle 3,1,1 \rangle \oplus \langle 3,2,0 \rangle) \otimes <1> \\
\oplus (\langle 2,1,1 \rangle \oplus \langle 2,2,0 \rangle \oplus \langle 3,1,0 \rangle) \otimes <2> \oplus (\langle 2,1,0 \rangle) \otimes <3>.
\]

In ref. 44 it is exhibited how to construct an element in a basis for \( \langle \nu_n \rangle \) that is simultaneously an element in a basis for \( \langle \nu_{n-1} \rangle \otimes <t> \). The procedure is the following: first adapt a tensor \( \mathbf{E}_{i_1 i_2 \cdots i_{N-t}} \) of \( S_{N-t}^\nu_n \) to \( [\nu_{n-1}] \) of \( S_{N-t}^\nu_{n-1} \), where \( 1 \leq i_j \leq N-1 \) and \( j = 1, \ldots, N-t \). This yields \( G_i \), associated with \( I_q^\nu_{n-1} \). Then adapt:

\[
G_i \otimes v_n \otimes v_n \otimes \cdots \otimes v_n
\]

t factors

\[
[\nu_n] \quad [\nu_{n-1}]
\]

to \( S_n \), yielding a Gelfand tensor \( G \), characterized by \( I_q^\nu_n \). But this is exactly the manner in which the Gelfand tensors were constructed during the proof of theorem 1, and hence Gelfand tensors are sequence-adapted to GL(n) \( \supseteq \) GL(n-1), that is, they carry simultaneously \( \langle \nu_n \rangle \) and \( \langle \nu_{n-1} \rangle \otimes <t> \).

The partition \( [\nu_{n-1}] \) is obtained by stripping all \( t \) boxes containing the index \( n \) off \( I_q^\nu_n \). Repeating this process for \( n, n-1, n-2, \ldots, 1 \), noting that all accompanying subductions are multiplicity free and that the tail group GL(1) has only one-dimensional tensor irreps, that is, the chain GL(n) \( \supseteq \) GL(n-1) \( \supseteq \) \( \ldots \) \( \supseteq \) GL(1) is canonical, we find that Gelfand tensors can be uniquely labelled by the successive partitions:

\( \langle \nu_n \rangle, \langle \nu_{n-1} \rangle, \ldots, \langle \nu_1 \rangle \). It is common to write these partitions in the form of a Gelfand pattern:

\[
G \equiv W_{qq}^\nu_n \mathbf{E}_i \equiv \\
\begin{pmatrix}
m_{1n} & m_{2n} & m_{3n} & \ldots & m_{nn} \\
m_{1,n-1} & m_{2,n-1} & \ldots & m_{n-1,n-1} \\
m_{1,n-2} & \ldots \\
& & & & m_{11}
\end{pmatrix}
\]

Recall again that Weyl's branching theorem states: \( m_{i+1,j+1} \leq m_{ij} \leq m_{i,j+1} \).
Example.

Element in $V_n \otimes^9$:

$$G = \begin{pmatrix}
3 & 2 & 2 & 1 & 1 & 0 \\
3 & 2 & 2 & 1 & 0 \\
3 & 2 & 1 & 0 \\
3 & 1 & 0 \\
3 & 0 \\
1
\end{pmatrix}$$

Notes.
1. Clearly a Gelfand tensor is labelled **uniquely** by the genealogy $\langle \mu_n \rangle, \langle \mu_{n-1} \rangle, \ldots, \langle \mu_1 \rangle$. This follows directly from the chain $GL(n) \supset GL(n-1) \supset \ldots \supset GL(1)$ being canonical. As pointed out in sec. I.4.7 an irrep $\langle \mu_k \rangle$ can alternatively be labelled by the eigenvalues of $k$ Casimir invariants. If we do this for $k = n, n-1, \ldots, 1$, then every Gelfand tensor can also be labelled uniquely by the eigenvalues of the $\lambda_{n(n+1)}$ invariants associated with this canonical chain, or in other words: these Casimir invariants form a complete set of commuting operators.

2. The fact that projection with Y-Y units, which themselves are adapted to the canonical chain $S_N \supset S_{N-1} \supset \ldots \supset S_1$, yields tensors adapted to the chain $GL(n) \supset GL(n-1) \supset \ldots \supset GL(1)$ is of course well-known. For example it encompasses also the result that genealogically constructed spin functions are identical with Y-Y projected spin functions [46, app. 2].

5.7. Explicit construction of matrix representations of $GL(n)$

In the computer calculations, reported on in the second part of this dissertation, actual irreducible tensor representations of $GL(n)$ have been employed to study the effect of a basis transformation in the one-electron space $V_n$ onto the Weyl-Rumer tensors. In this section it will be discussed how the irreducible matrices representing the elements of $GL(n)$ may be constructed. Naturally these matrices depend on the choice of basis for $V_n \otimes^N$.

Let us first consider the basis projected by the Y-Y units $\nu_{kk}^{[\lambda]}$, because in this case the computation is straightforward. The following notation will be used throughout this section:
\[ E'_I = (av_{i_1} \otimes (av_{i_2}) \otimes \ldots \otimes (av_{i_N})) , \quad \alpha \in \text{GL}(n) . \]

\( (a^I_j) \) is the matrix of \( \alpha \).

\[ A^I_J = a^I_1 a^I_2 \ldots a^I_N . \]

\( I_o, J_o \) are ordered index sets (indices not decreasing from left to right).

\[ C(I) = 1/(N_1! N_2! \ldots N_N!) , \]

where \( N_k \) is the number of times \( i_k \) occurs in \( I , \) \( k = 1, \ldots , N \).

Furthermore the superscript \( [\lambda] \) will be omitted everywhere in this section.

The rows and columns of the matrix \( T(\alpha) \), representing \( \alpha \) on an irreducible subspace of \( V_n \otimes^N \), are labelled by standard index tableaux, just as the Gelfand basis itself. If \( I_k \) is a standard index tableau belonging to the index set \( I_o \) and the standard Young tableau \( T_k \) (see sec. I.5.4), then one can write:

\[ G^I_k = W_{kk} (E^I_O) = W_{kk} (\sum J o C(J_O) \sum J o A^J_I ) P(J)^P_O . \]

where the summation over all possible index sets \( J \) has been broken up into one over ordered index sets \( J_o \) and a sum over all the permutations of \( J_o \).

The factor \( C(J_o) \) corrects for the presence of equal indices in \( J_o \).

Using:

\[ W_{kk} E^P(J)^P_O = W_{kk} P^P O(J) = \sum Q k q W_{kk} E^J_q = \sum Q k q G^J_q , \]

where \( O(P) \) is the orthogonal Young-Yamanouchi matrix (sec. I.5.3), (recall that the Y-Y units span minimal right ideals), and:

\[ A^I_O = P^{-1} A^I_O , \]

(see I.4.4), one gets:

\[ G^I_k = \sum J_o C(J_o) \sum J_o O(P)_{kj} P^{-1}(J) A^I_O G^J_{q} . \]

Further defining:

\[ A^I_r = \sum J_o O(P)_{kj} A^I_O = \frac{N!}{f[\lambda]} W_{kk} A^I_O \]

one gets:

\[ G^I_k = \sum J_o C(J_o) \sum J_o A^I_o G^J_{q} . \]

The summation over \( q \) in this expression is still running over all \( f[\lambda] \).
standard Young tableaux, and so \( J_q \) is either a standard or a half-standard index tableau. However, it has been shown in the preceding section that the half-standard Gelfand tensors are zero and that the ones with the same standard index tableau are identical. So, we may let \( q \) run over the different standard index tableaux only, if we introduce a factor \( m_q(J_o) \), which is the number of times \( J_q \) occurs in the original sum over \( q \). Finally one arrives at:

\[
G_k^i = \sum J_q \sum C(J_o) m_q(J_o) A_q^i G_k^q.
\]

The prescription for the calculation of \( T(\alpha) \) is now transparent: The columns of \( T(\alpha) \) are labelled by the different standard index tableaux \( I_k \), which are obtained by letting \( I_o \) run over all ordered index sets and \( k \) over all the standard index tableaux associated with \( I_o \). The rows, characterized by \( J_q \), are similarly obtained from letting \( J_o \) run over all ordered index sets and \( q \) over all different standard index tableaux associated with \( J_o \).

The matrix element

\[
T(\alpha)_{J_q}^i = C(J_o) m_q(J_o) A_q^i
\]

follows by projection with \( \frac{N!}{F[\lambda]} \) \( W_k \) onto \( A_q^i \).

Note.
This prescription is equivalent to the one given by Littlewood [50, p. 183 ff] for the construction of "invariant matrices". Littlewood employs to this end natural units of the PN-type.

The tensor representations carried by the W-R basis are more difficult to derive. When trying to follow the procedure above, one finds that the half-standard tensors do not drop out of the expressions, because they are in general non-zero, as we have seen in sec. I.5.5. So, we must first be able to express the half-standard tensors in terms of the standard ones. Although it is not possible to give explicit relations, one can derive sets of linear equations from which the half-standard tensors can be obtained in terms of the standard tensors.

Let \( (p,q) \) be the permutation that interchanges the two equal indices that are by definition present in one of the columns of the half-standard index tableau \( I_j \). Then we get the equation:
\[ F_{I_j} = F_{(p,q)} I_j = Y_{jj} (p,q) E_I = \sum_k R^{(p,q)}_{kj} Y_{jk} E_I = \sum_k R^{(p,q)}_{kj} F_{I_k}, \]

where \( R^{(p,q)} \) is the matrix representation of \((p,q)\), introduced in sec. I.5.1. As we have seen there, this matrix has a very simple appearance; and consequently the half-standard tensors can be easily eliminated from equations of this type.

**Example.**

Consider \([\lambda] = [2^2, 1]\) and \(I = \{1, 1, 2, 3, 4\}\).

The linear equations generated by \((p,q) = (12)\) read:

\[
\begin{pmatrix}
1 & 2 & 1 & 2 & 1 & 3 & 1 & 3
3 & 1 & 4 & 1 & 4 & 1 & 4 & 2
4 & 3 & 2
\end{pmatrix}
= \begin{pmatrix}
1 & 2 & 1 & 2 & 1 & 3 & 1 & 3
3 & 1 & 4 & 1 & 4 & 1 & 4 & 2
4 & 3 & 2
\end{pmatrix}
\begin{pmatrix}
-1 & 0 & 1
0 & -1 & 1
-1 & 0 & 0
0 & -1 & 0
0 & 0 & -1
\end{pmatrix}
\]

These equations yield the following expressions for the half-standard tensors:

\[
F_{12} = -\frac{1}{2} F_{12}, \quad F_{13} = -\frac{1}{2} F_{13}, \quad F_{14} = \frac{1}{2} F_{14} + \frac{1}{2} F_{14}
\]

We are dealing with sets \(\{F_{I_k} | k = 1, \ldots, f[\lambda]\}\) in which only the tensors labelled by different standard index tableaux are linearly independent. Tensors characterized by the same (standard) index tableau are identical, half-standard tensors in \(\{F_{I_k}\}\) depend linearly on the standard tensors. We now choose a maximal linearly independent subset of \(\{F_{I_k}\}\), and label its elements by \(r, s, t, \ldots\). The remaining tensors in \(\{F_{I_k}\}\) are labelled by \(a, b, c, \ldots\); the whole set is labelled by \(i, j, k, \ldots\). We can then summarize the foregoing considerations in the one formula:

\[
F_{I_a} = \sum_r S(I)_{ra} F_{I_r}.
\]

The matrix \(S(I)\) can be computed in the manner just described and is in the sequel assumed to be known.

The computation of the tensor representation carried by the W-R basis is now straightforward:
\[ F'_i = \sum_{J_0} C(J_0) \sum_{i} \left( \sum_{P \in S_{N}} D(P)_{ir} A_{I_0} \right) P(J_0) F_{J_i}. \]

Recall (sec. I.5.1):
\[ R^{-1}(P) = X^{-1} U(P^{-1}) \]
and so the expression between square brackets becomes:
\[ \sum_{P} \sum_{j} \eta_{ij} U(P^{-1})_{jr} A_{I_0}^{J_0} = \sum_{j} \eta_{ij} Y_{jr} A_{I_0}^{J_0}, \]
where \( \eta_{ij} \) is the \((i,j)\)-matrix element of \( X^{-1} \). Now, \( Y_{jr} A_{I_0}^{J_0} = 0 \) if \( J_j \) is half-standard (see the proof of the main theorem of sec. I.5.5) and the sum over \( j \) can accordingly be replaced by one over \( s \), if we correct for the \( m_s(j) \)-fold occurrence of the standard index tableau \( J_s \).

Defining:
\[ A_{I_r}^{J_i} = \sum_{s} m_s(J_0) \eta_{is} Y_{sr} A_{I_0}^{J_0}, \]
we get:
\[ F'_i = \sum_{J_0} C(J_0) A_{I_r}^{J_i} F_{J_i}. \]

The summation over \( i \) in this expression runs still over all standard and half-standard index tableaux. \( J_0 \) runs over all ordered index sets that have no indices occurring more often than either the number of rows or the number of columns of \( \lambda \), because if any index appears with a frequency higher than either of these two numbers \( F_{J_i} \) is zero.

Inserting the expressions derived earlier, one finally arrives at:
\[ F'_i = \sum_{J_0} C(J_0) \sum_{t} \left( A_{I_r}^{J_t} + \sum_{a} S(J_0)_{ta} A_{I_r}^{J_a} \right) F_{J_t}, \]
from which one concludes that the \((J_0^t, I_r)\)-matrix element is:
\[ T(a)_{I_r}^{J_t} = C(J_0) \left[ \sum_{s} m_s(J_0) (\eta_{ts} + \sum_{a} S(J_0)_{ta} \eta_{as}) Y_{sr} \right] A_{I_0}^{J_0}. \]

This expression looks rather complicated, but usually simplifies drastically in practice, because most of the times \( \eta_{as} \) is equal to zero for all values of its row and column index, which run over the half-standard and standard index tableaux respectively. Indeed, I have not been able to find a one- or two-columned representation with an \( \eta_{as} \)-value not equal to zero for any index set. Note further that \( \eta_{ts} \) is almost always equal to \( \delta_{ts} \).
and we see that the calculation of $T(u)_{I_{tr}}^J$ usually amounts to not much more than projection with $Y_{tr}$.

To clarify these remarks we turn to the example $[2^2, 1]$, treated in sec. I.5.1. The matrix $X^1$ is here:

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
$$

The first row is labelled by:

$$
T_1 = \begin{array}{c}
1 \\
2 \\
3 \\
4 \\
5 \\
\end{array}
$$

The only off-diagonal element not equal to zero is $n_{15}$. The ordered index sets $J_o = \{ j_1 \leq j_2 \leq j_3 \leq j_4 \leq j_5 \}$ that give half-standard index tableaux in cooperation with $T_1$ satisfy:

$j_1 = j_3$ and/or $j_2 = j_4$ and/or $j_3 = j_5$ and/or $j_1 = j_5$.

But then, because of the ordering:

$j_1 = j_2 = j_3$ and/or $j_2 = j_3 = j_4$ and/or $j_3 = j_4 = j_5$ and/or $j_1 = j_2 = j_3 = j_4 = j_5$.

And so $J_o$ has at least one index triply occurring, and hence does not contribute to tensor representations with two-columned diagrams. Therefore in this example $n_{15}$ does not contribute to $T(u)$ and $S(J)$ does not arise for any index set $J$.

Note.

Very recently another formula for the tensor representation carried by the W-R basis has been given by Brunet and Seligman [52]. Their derivation is completely different from the one just given. Also their final formula has an entirely different appearance.

5.8. A note on the tensor irreps of the Lie algebra of $GL(n)$

In this section a few of the Lie algebraic connotations of the decomposition of tensor space will be discussed. First we introduce some general definitions concerning the representations of Lie algebras. Since our main
interest is in GL(n), these definitions are all given for the Lie algebra AL(n) of GL(n). However, these definitions also hold for arbitrary Lie algebras; see ref. 16 for further details.

Definitions.

1. Let \( E^{ij} \), \( i, j = 1, \ldots, n \) be the linear operator on \( V_n \) with the matrix:

\[
(E^{ij})_{kl} = \delta^i_k \delta^j_l.
\]

As discussed in sec. I.3.2, these \( n^2 \) operators span AL(n), the Lie algebra of GL(n).

2. Let \( \mathcal{H}_n \) be the \( n \)-dimensional subalgebra of AL(n) spanned by:

\[
\{ E^{ii} \mid i = 1, \ldots, n \}.
\]

It is easy to show that this algebra is the maximal commutative subalgebra of AL(n): the Cartan subalgebra of AL(n).

3. Let D be an \( m \)-dimensional linear Lie algebra representation of AL(n).

Let D be carried by \( V_m \). Since the elements of the Cartan subalgebra \( \mathcal{H}_n \) commute, all the linear operators \( D(H), H \in \mathcal{H}_n \), can be simultaneously diagonalized on \( V_m \). Let \( v_o \in V_m \) be a common eigenvector of the elements of \( D(\mathcal{H}_n) \), that is:

\[
D(H) v_o = \alpha(H) v_o, \quad \forall H \in \mathcal{H}_n, \; 0 \neq v_o \in V_m, \; \alpha(H) \in \mathbb{C}.
\]

The set of complex eigenvalues \( \alpha(H) \) can be considered as a mapping of \( \mathcal{H}_n \) into \( \mathbb{C} \). This mapping \( \alpha \) is linear, for:

\[
D(\lambda H + \lambda' H') v_o = \lambda D(H) v_o + \lambda' D(H') v_o = \lambda \alpha(H) v_o + \lambda' \alpha(H') v_o,
\]

and also:

\[
D(\lambda H + \lambda' H') v_o = \alpha(\lambda H + \lambda' H') v_o.
\]

Hence:

\[
\alpha(\lambda H + \lambda' H') = \lambda \alpha(H) + \lambda' \alpha(H')
\]

with \( \lambda, \lambda' \in \mathbb{C}, H, H' \in \mathcal{H}_n \).

The linear functional \( \alpha: \mathcal{H}_n \to \mathbb{C} \), associated with the common eigenvector \( v_o \in V_m \), is called a weight of the representation D.

4. Let \( H = \Sigma h_i E^{ii} \) belong to \( \mathcal{H}_n \), then:

\[
\alpha(H) = \Sigma h_i \alpha(E^{ii}) = \Sigma h_i a_i
\]

with \( a_i \equiv \alpha(E^{ii}) \).

So, every eigenvalue \( \alpha(H) \) can be expressed in terms of \( a_1, a_2, \ldots, a_n \). These scalars constitute a weight vector, associated with the
representation $D$ and the eigenvector $v_o \in V_m$. (The word vector is often omitted in this definition).

5. As we will see below, the components of the weight vectors associated with tensor representations of $\text{AL}(n)$ are natural numbers. It is convenient to order these weight vectors in the following manner:

The weight vector $(\alpha_1', \alpha_2', \ldots, \alpha_n')$ comes before the weight vector $(\alpha_1', \alpha_2', \ldots, \alpha_n')$, if the first non-zero difference $\alpha_i - \alpha_i' (i = 1, \ldots, n)$ is positive. The weight $(\alpha_1', \alpha_2', \ldots, \alpha_n')$ is then called higher than the weight $(\alpha_1', \alpha_2', \ldots, \alpha_n')$; so it makes sense to speak of the highest weight of a tensor representation.

We proceed by showing that an arbitrary Weyl-Rumer or Gelfand tensor is a simultaneous eigenvector of all elements in the Cartan subalgebra of $\text{AL}(n)$.

In sec. I.4.3 the following tensor representation $T$ of $\text{AL}(n)$ carried by $V \otimes V$ has been defined:

$$T: E^{ij} \rightarrow \sum_{k=1}^{N} E^{ij}(k).$$

The operator $E^{ij}$ acts on a basis vector $v_p$ of $V_n$ in the following manner:

$$E^{ij} v_p = \delta^j_p v_i,$$

where $v_i$ is also a basis vector of $V_n$. Now:

$$T(E^{ij}) E_I \equiv \sum_{k=1}^{N} v_{i_1} \otimes \cdots \otimes v_{i_k-1} \otimes (E^{ij}) v_{i_k} \otimes \cdots \otimes v_{i_N} = N_j E_I,$$

where the "occupation number" $N_j$ is the number of times the index $j$ occurs in the index set $I$ which labels the tensor $E_I$. Physically, the operator $T(E^{ij})$ represents the numbers of particles in orbital $v_j$ and is therefore called a number operator. The element $E^{ij}$ is by definition a basis element of $\mathcal{H}_n \subset \text{AL}(n)$, and so the set of occupation numbers $(N_1, N_2, \ldots, N_n)$ is a weight vector associated with the tensor representation $T$ and the tensor $E_I \in V_n \otimes V$. Every tensor $E_I$ is a common eigenvector of the $n$ number operators $T(E^{ij})$, $j = 1, \ldots, n$. Clearly, all elements of the subspace $U$ of $V_n \otimes V$ spanned by:

$$\{E_P(1) \mid \text{WPS}_N, \text{fixed } I\}$$

also belong to this same weight (or in more physical terms: belong to the same configuration $I$). Adaptation of $U$ to $S_N$ by means of either Young units or $Y-Y$ units does not change the weight of the elements of $U$, and
hence we conclude that all W-R and Gelfand tensors, arising from $E_+$, belong to the weight vector $(N_1, N_2, \ldots, N_n)$.

Consider the tensor irrep $<\lambda_1, \lambda_2, \ldots, \lambda_n>$ of AL(n). A W-R or Gelfand tensor with the following standard index tableau:

\[
\begin{array}{cccccc}
1 & 1 & 1 & \cdots & \cdots & 1 \\
2 & 2 & \cdots & \cdots & \cdots & 2 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
n & \vdots & \cdots & \cdots & \cdots & n \\
\end{array}
\]

where the $i$-th row has $\lambda_i$ boxes ($i = 1, \ldots, n$), is associated with weight $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. It can easily be verified that no standard index tableau of this shape can be constructed that is associated with a weight higher than $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ (see sec. I.5.5 or I.5.6). So, the row lengths of the Young diagram are the components of the highest weight vector associated with the irrep $<\lambda_1, \lambda_2, \ldots, \lambda_n>$ of AL(n). Since a tensor irrep of AL(n) is uniquely characterized by a Young diagram it is also determined unambiguously by the components of the highest weight vector.

The elements $T(E_{ij}^{\dagger})$, $i < j$, change a tensor of weight:

$$(N_1, N_2, \ldots, N_1, \ldots, N_j, \ldots, N_n)$$

into one of weight:

$$(N_1, N_2, \ldots, N_{i+1}, \ldots, N_{j-1}, \ldots, N_n).$$

Because the new weight is higher, $T(E_{ij}^{\dagger})$ is often called a raising operator. Similarly $T(E_{ij})$, $i > j$, is a lowering operator.

One can prove that $E_{ij}^i = T(E_{ij}^{\dagger})E_{ij}$ ($i \neq j$) is of weight $N_i+1$ either by inspection or by using: $T(E_{ij}^{\dagger}) = T(E_{ii}^{\dagger})T(E_{ij}^{\dagger}) - T(E_{i+1}^{\dagger})T(E_{i+1}^{\dagger})$ from which follows $T(E_{ij}^{\dagger})E_{ij}^i = (N_i+1)E_{ij}^i$. Similarly one can prove that $E_{ij}^i$ is of weight $N_j-1$ by using: $T(E_{ij}^{\dagger}) = T(E_{ij}^{\dagger})T(E_{ii}^{\dagger}) - T(E_{jj}^{\dagger})T(E_{jj}^{\dagger})$.

Example:

The Gelfand tensor $G$ is an element in a basis for $<4,2>$ of AL(2),

$$\begin{pmatrix}
1 & 1 & 1 & 2 \\
2 & 2 & \end{pmatrix}$$

$\Lambda(U(2))$ and $\Lambda(SU(2))$. It has weight $(3,3)$. The raising operator $T(E_{12}^{\dagger})$ has the following effect:

$$T(E_{12}^{\dagger})G = G + G + G = G.$$
The resulting tensor is of weight \((4,2)\).

The Casimir invariants of \(U(2)\) and \(SU(2)\) can easily be expressed in terms of the raising and lowering operators. For instance in the case of \(SU(2)\) one first goes over to a traceless basis of \(\Lambda(SU(2))\), via:

\[
\begin{align*}
\overline{E}^{12} &\equiv \mathbf{E}^{12} \equiv \frac{\sigma_+}{2} \\
\overline{E}^{21} &\equiv \mathbf{E}^{21} \equiv \frac{\sigma_-}{2} \\
\overline{E}^{11} &\equiv \frac{(\mathbf{E}^{11} - \mathbf{E}^{22})}{2} \equiv \frac{\sigma_z}{2} \\
\overline{E}^{22} &\equiv -\frac{(\mathbf{E}^{11} - \mathbf{E}^{22})}{2} = -\frac{\sigma_z}{2}.
\end{align*}
\]

Note that this basis differs by a simple linear transformation from the one introduced in the example of sec. I.3.4.

The Casimir invariants are then [31]:

\[
\begin{align*}
I^{(2)}_1 &\equiv \overline{E}^{11} + \overline{E}^{22} = 0 \\
I^{(2)}_2 &\equiv \overline{E}^{11}\overline{E}^{11} + \overline{E}^{22}\overline{E}^{22} + \overline{E}^{11}\overline{E}^{22} + \overline{E}^{22}\overline{E}^{11} = 2s^2.
\end{align*}
\]

From the theory of chapter I.4 we know that Gelfand tensors are eigenvectors of the Casimir operators (the representations of the Casimir invariants). Using these expressions for the invariants one can easily calculate the eigenvalues:

\[
T(I^{(2)}_2)G = 2(G + G) = 2(1(1+1))G
\]

The reader is referred to the work of Biedenharn et al. [53], Louck [31], Moshinsky [54] and others for the representation theory of \(SU(n)\) in this manner, which is in fact a generalization to \(SU(n)\) of the Racah-Wigner calculus of \(SU(2)\).
CHAPTER I.6. PHYSICAL APPLICATIONS

In this chapter the mathematics introduced in the preceding chapters will be applied to the physics of N-particle systems. In particular it will be shown that $GL(n)$ is a dynamical group and $S_N$, a symmetry group of a model system satisfying the following conditions:

- The system contains $N$ identical particles, all moving in a stationary outer electric field (e.g. electrons moving in the field of one or more clamped nuclei).
- Magnetic and relativistic (including spin) contributions to the energy are absent. (These contributions can be introduced in a later stage as perturbations).
- The Hilbert space of the system is $[(2\sigma+1) \times n]^N$ dimensional, where $\sigma$ is the spin of the particles and $n$ is finite.

The great majority of the quantum chemical calculations performed to date satisfy these conditions.

In order to separate the approximating assumptions about the form of the wave function from the customary exact quantum mechanical postulates, the first two sections of this chapter will be devoted to a brief review of the axioms regarding the form of exact N-particle wave functions.

. 6.1. N-particle state vectors

The chief purpose of this section is to establish some terminology and to introduce the notation used in later sections.

Definitions.

1. A Euclidean space $V$ is a vector space furnished with a positive definite inner product (see e.g. ref. 55). Henceforth this inner product will be denoted by $\langle f | g \rangle$, $f, g \in V$. A positive definite inner product induces a metric [55] on $V$, and so every Euclidean space is a metric space.

   Euclidean spaces are also called pre-Hilbert, unitary or inner product spaces.

2. A metric space is complete if every Cauchy sequence [20, p. 45] converges to an element of the space.

3. A Hilbert space is a complete Euclidean space.

4. A function $f: \mathbb{R}^m \to \mathbb{C}$ is square integrable if the integral:
\[
\int_{\mathbb{R}^m} f^*(x)f(x)dx
\]
is well-defined and finite.

5. The class of all square integrable functions on \(\mathbb{R}^m\) will be denoted by \(L^2(\mathbb{R}^m)\).

**Examples.**

1. A Euclidean space of finite dimension is complete and hence a Hilbert space. (Completeness follows almost immediately from the completeness of the real line [56, p. 55]).

2. The parameter space of \(U(n)\) is a complete (non-linear) subspace of \(\mathbb{R}^{n^2}\) with respect to the usual metric of \(\mathbb{R}^{n^2}\).

3. Endowing \(L^2(\mathbb{R}^m)\) with the inner product:

\[
\langle f|g \rangle \equiv \int_{\mathbb{R}^m} f^*(x)g(x)dx
\]

the set becomes a Hilbert space (see theorem II.4.3 and II.4.4 of ref. 55).

**Notes.**

1. Defining the norm of \(f \in L^2(\mathbb{R}^m)\) by

\[
|f| = \langle f|f \rangle^{1/2}
\]

it follows from the definition that all elements of \(L^2(\mathbb{R}^m)\) have a finite norm.

2. All Hilbert spaces of interest to us have a countable orthonormal basis that is complete (in the vector space sense of the term). Hilbert spaces having such a basis are called separable.

**Definitions.**

1. Let \(V\) be a Euclidean space. One can define an inner product on \(V \otimes^N\) thus:

\[
\langle f_1 \otimes \ldots \otimes f_N | q_1 \otimes \ldots \otimes q_N \rangle = \langle f_1|q_1 \rangle \ldots \langle f_N|q_N \rangle.
\]

It is easily seen that this inner product is positive definite, and so \(V \otimes^N\) is a Euclidean space.

2. Let \(V\) be a Hilbert space. If \(V\) is of infinite dimension, \(V \otimes^N\) is not necessarily complete. However, the completion of \(V \otimes^N\) is uniquely determined [55, p. 144] and will also be denoted by \(V \otimes^N\). From here on the definition of the tensor product is assumed to be extended in such a
manner that $V \otimes^N$ is always a Hilbert space.

Note.
Interpreting a tensor product of functions as a product function in the ordinary sense, the Hilbert space $L^2(\mathbb{R}^3) \otimes^N$ is imbedded in $L^2(\mathbb{R}^{3N})$. It can be proved that:

$$L^2(\mathbb{R}^3) \otimes^N = L^2(\mathbb{R}^{3N}).$$

(See ref. 55, p. 144).

Definition.
Let $v$ be an arbitrary fixed vector in the linear space $V$. The set

$$\{av \mid \text{all } a \in \mathbb{C}\}$$

is a ray in $V$.

After having introduced this necessary mathematical terminology we are now ready to introduce a physics postulate.

Postulate.
The state of a system consisting of $N$ identical particles of spin $\sigma$ is uniquely characterized by a ray in the Hilbert space:

$$\mathcal{H} \equiv L^2(\mathbb{R}^{3N}) \otimes (L^2(s) \otimes^N).$$

Here:

$L^2(\mathbb{R}^{3N})$ is the spatial Hilbert space of square integrable functions on $\mathbb{R}^{3N}$.

$L^2(s)$ is the spin Hilbert space of dimension $2\sigma+1$.

Notes:
1. Instead of considering rays it is customary to take a representative of the ray by normalizing the state to unity. Even so a phase factor of absolute magnitude 1 is left unspecified.
2. The system Hilbert space can be written as the tensor product of one-particle Hilbert spaces:

$$\mathcal{H} = [L^2(\mathbb{R}^3) \otimes L^2(s)] \otimes^N.$$

3. A vector in $L^2(\mathbb{R}^{3N})$ will be written as $\phi(r)$ with 

$$r = (r_1, r_2, \ldots, r_N) \in \mathbb{R}^{3N}.$$

Analogously a vector in $L^2(s) \otimes^N$ will be written as $\theta(s_1, s_2, \ldots, s_N)$, where $s_i$ is the spin coordinate of particle $i$, $i = 1, \ldots, N$.  

100
4. Note that $L^2 (\mathbb{R}^3) \otimes L^2 (\mathbb{S}^2)$ is isomorphic with the space of $(2\sigma+1)$-component spinors (ref. 4, p. 192). This isomorphism also preserves inner products. So, the postulate above is equivalent to one which expresses that a state is characterized by a ray in a tensor product of the Hilbert spaces of $(2\sigma+1)$-component spinors.

5. Although every state of a system is characterized by a ray in a Hilbert space, the converse is not true; that is, not every ray in a Hilbert space, associated with a system, represents a realizable state of this system.

6.2. Symmetry of N-particle systems

Before turning our attention to systems consisting of N identical particles; we first review briefly the case of one particle.

Definitions.
1. The set of all unitary and antiunitary operators [4, p. 203] on an n-dimensional space is denoted by $U^\pm (n)$. It is easy to show that $U^+ (n)$ forms a group.

2. Let $V_n = V_{n_1} \oplus V_{n_2}$. The set of all unitary and antiunitary operators on $W$ that leave both $V_{n_1}$ and $V_{n_2}$ invariant is denoted by $U^+(n_1) \times U^+(n_2)$. This set forms a group.

Consider a one-particle system with configuration space $\mathbb{R}^3$ and Hamiltonian $h(\mathbf{r})$, $\mathbf{r} \in \mathbb{R}^3$. Assume that $L^2 (\mathbb{R}^3)$ is completely decomposed into eigenspaces of $h(\mathbf{r})$:

$$h(\mathbf{r}) | \epsilon_{k_1} \rangle = \epsilon_{k_1} | \epsilon_{k_1} \rangle, \quad i = 1, 2, \ldots; \quad k_1 = 1, \ldots, n_1.$$ 

Here $k_1$ runs over the degenerate states of energy $\epsilon_{k_1}$. In this equation as in the rest of this chapter, it is ignored that part of the spectrum of $h(\mathbf{r})$ may be continuous. All operators on $L^2 (\mathbb{R}^3)$ that commute with $h(\mathbf{r})$ and that are either unitary or antiunitary form a group $G^f_h$: the full commutator group of $h(\mathbf{r})$. Knowing the degeneracies $n_1$, $i = 1, 2, \ldots$, of $h(\mathbf{r})$, it can easily be derived [57] that:

$$G^f_h = U^+(n_1) \times U^+(n_2) \times U^+(n_3) \times \ldots$$

And so each eigenspace of $h(\mathbf{r})$ is invariant and irreducible under $G^f_h$.

Clearly the full commutator group of $h(\mathbf{r})$ lacks physical meaning and is
not helpful in any way to the simplification of practical calculations. The group is only known after all degeneracies of \( \hat{h}(\vec{r}) \) have been given and hence cannot contribute to an a priori knowledge of the splitting of energy levels.

However, if physical arguments, e.g. of geometric or Lie algebraic nature, make it possible to find subgroups of \( G_{h}' \), then group theory will become useful. If a subgroup of \( G_{h}' \) can be found that shares with \( G_{h}' \) all the irreps carried by \( L^2(\mathbb{R}^3) \), or in other words if a subgroup can be found that predicts the spectrum of \( h(\vec{r}) \), then group theory contributes to the solution of quantum mechanical problems in the maximal possible way. In general \( G_{h}' \) can have several such spectrum predicting subgroups, the smallest of these will be called the symmetry group of \( h(\vec{r}) \), denoted by \( G_{h}^s \).

Simple geometric arguments enable us almost always to define at least one physically meaningful subgroup of \( G_{h}' \): the invariance group \( G_{h}^i \) of \( h(\vec{r}) \). This group is associated with the maximal group of isometric operators on \( \mathbb{R}^3 \) that leave \( h(\vec{r}) \) invariant. (See e.g. ref. 4, p. 214 for the association of a Hilbert space operator with an \( \mathbb{R}^3 \) operator). Since isometric operators on \( \mathbb{R}^3 \) give rise to unitary Hilbert space operators and since the elements of \( G_{h}^i \) commute with \( h(\vec{r}) \) [4, p. 214], it follows that the invariance group is a subgroup of \( G_{h}' \). Also, because the eigenspaces of \( h(\vec{r}) \) are stable under \( G_{h}^i \), it follows that the invariance group is a (proper or improper) subgroup of the symmetry group.

**Example.**
To illustrate these points we consider the hydrogen atom. The energy levels are labelled by \( n = 1, 2, \ldots \), the \( n \)-th level is \( n^2 \)-fold degenerate, and so the full commutator group is \( U^\pm((1)^2) \times U^\pm((2)^2) \times \ldots \). It has been found by Fock [58] that the 4-dimensional rotation-reflection group \( O(4) \) predicts the (discrete part of the) spectrum of the H-atom, and so \( O(4) \) is the symmetry group. Finally, the invariance group is \( O(3) \) (this group does of course not predict the degeneracy of levels with the same \( n \)- but different \( l \)-quantum numbers). Summarizing:

\[
U^\pm(1) \times U^\pm(4) \times U^\pm(9) \times \ldots \supset O(4) \supset O(3).
\]

In this discussion of one-particle symmetry the definition of the symmetry group of the Hamiltonian \( h(\vec{r}) \) lacked preciseness; and indeed one can only be assured that \( G_{h}^s \) really predicts the spectrum when this is known
completely. In most of the few cases where the spectrum is known exactly one finds degeneracies that are higher than those that can be expected on the grounds of the invariance group alone. Besides the H-atom, other well-known cases showing "hidden" symmetry are the 3-dimensional harmonic oscillator and the spherical top [59]. Also many model Hamiltonians have this feature [60].

Nevertheless, the one particle symmetry group $G^S_h$ is assumed to be known in the now following discussion of N-particle symmetry.

---

**Definition.**

Consider a system of N identical particles with Hamiltonian

$$H = \sum_{i=1}^{N} h(i) + \sum_{i<j}^{N} h(i,j)$$

acting on the Hilbert space $\mathcal{H}$. Let $G^S_h$ be the symmetry group of $h(i)$, and let $G^S_h$ be the maximal (proper or improper) subgroup of $G^S_h$ which is such that the elements of the N-fold inner direct product group (sec. I.4.2) $[G^S_h]^N$ commute with $H$. The symmetry group $G_H$ of $H$ is defined by:

$$G_H \equiv [G^S_h]^N \otimes S_N'$$

where $S_N$ permutes the spatial and spin coordinates of the particles, i.e. the particle labels.

---

**Notes.**

1. The elements of $[G^S_h]^N$ commute with those of $S_N$; this justifies the use of the direct product symbol in the definition of $G_H$.

2. Elements of $G_H$ are unitary or antiunitary. This follows because the elements of $S_N$ are unitary and the elements of $G^S_h$ are either unitary or antiunitary.

3. Unitarity or antiunitarity of an element $U \in G_H$ assures that the transformed eigenstates $U|E\rangle$ of $H$ cannot be distinguished by any measurement from the original states $|E\rangle$.

4. The symmetry group $G_H$ contains operators which effect the same symmetry transformation on each particle separately, followed by a permutation of the identical particles among themselves. (Such a permutation is evidently an N-particle symmetry operator).

5. The occurrence of an inner direct product in the definition of $G_H$ is necessary if we wish to interpret the elements of Lie algebras of
symmetry groups as observables. For if this specific form of \( G_H \) were not stipulated we would be led to consider Lie algebras of outer direct product (or perhaps even larger) groups. The elements of such algebras do not commute with \( S^N_N \), so they either do not correspond to observables, or they can distinguish between the different particles in contradiction with the assumption that they are identical.

6. The symmetry group \( G_H \) may or may not be identical with \( [G_H^S]^N_N \otimes S^N_N \). A spinless \( N \)-electron atom yields an example of a proper embedding: the \( N \)-electron symmetry group is \( S^N_N \otimes [O(3)]^N \), while the one-particle symmetry group is \( O(4) \). The appearance of \( O(3) \) rather than \( O(4) \) in the inner product is due to the presence of electron-electron interaction terms in \( H \). In general, the interaction term \( h(i,j) \) may or may not lead to a symmetry group in which \( G_H \) is a proper subgroup of \( G_H^S \).

The following postulate gives \( G_H \) quantum mechanical meaning.

**Irreducibility Postulate.**

The eigenstates of \( H \) belonging to the same eigenvalue of \( H \) carry an irrep of \( G_H \).

Knowledge of \( G_H \) and its Hilbert space irreps enables the prediction of a good deal about the spectral decomposition of \( H \). All degeneracies can be predicted and labelled. The transformation properties of the eigenstates of \( H \) can be derived, even if the states themselves are not known. From these one can derive e.g. selection rules, and rules to describe the splittings of levels in perturbing fields, etc. All this and much more has been worked out in detail in many text books on group theory. So, we forgo any discussion about these general aspects of the irreducibility postulate.

The symmetry group \( G_H \) is a direct product of the two commuting groups \( [G_H^S]^N_N \) and \( S^N_N \). To obtain a decomposition of the system Hilbert space

\[ \mathcal{H} = [L^2(R^3) \otimes L^2(s)]^N \]

under \( G_H \) one can use this fact by proceeding in the following manner:

(i) First decompose \( \mathcal{H} \) into irreducible \( S^N_N \)-carrier spaces. All elements transforming according to the \( k \)-th row of the irrep \( [\lambda] \) span a space \( \mathcal{H}_k^{[\lambda]} \subset \mathcal{H} \). This space is invariant under \( [G_H^S]^N_N \). (Apply, to prove this, Schur's lemma in the manner of the proof given in sec. I.2.3).
(ii) Decompose the spaces \( \mathcal{H}_k^{[\lambda]} \), \( k = 1, \ldots, f[/\lambda] \) under \( [G_h]^N \), applying within each space the same basis transformation.

Let

\[
\mathcal{H}^{(\lambda, \mu)}_{k,p} \subset \mathcal{H}_k^{[\lambda]}
\]

carry the representation \((\mu)\) of \( [G_h]^N \). The index \( p \) labels the multiple occurrence of \((\mu)\) in \( \mathcal{H}_k^{[\lambda]} \). It now immediately follows that

\[
V^{(\lambda, \mu)} = \sum_{k=1}^{f[/\lambda]} \otimes \mathcal{H}^{(\lambda, \mu)}_{k,p}
\]

carries an irrep of \( G_H \equiv S_N \otimes [G_h]^N \).

If one would wish to arrive at a complete decomposition of \( \mathcal{H} \) into eigenspaces of \( H \), then this is as far as group theory can go. To obtain a complete spectral decomposition of \( H \) one can continue as follows: Take from every space \( V^{(\lambda, \mu)}_p \), \( p = 1, 2, \ldots \), the basis element that transforms to the same row of the irrep \((\lambda, \mu)\) of \( G_H \), say the \( q \)-th row. The set:

\[
\{ |[\lambda], (\mu), q, p \rangle | p = 1, 2, \ldots \}
\]

then spans a subspace of \( \mathcal{H} \) invariant under \( H \). Diagonalize \( H \) over this set. The irreducibility postulate states that no eigenvalue of \( H \) occurs more than once in this space, and so these eigenvalues may replace \( p \) to index uniquely the basis vectors. The index \( q \) labels the degenerate states.

The preceding discussion exhibits that the eigenstates of \( H \) are labelled by \((\mu)\) and \([\lambda]\). The following postulate, however, shows that \([\lambda]\) is superfluous for physically realizable states.

**Pauli Postulate.**

Only states belonging to one-dimensional irreps of \( S_N \) are physically realizable. Systems of identical particles with integral spin belong to the symmetric representation, systems of identical particles with half-integral spin belong to the antisymmetric representation.

**Notes.**

1. The Pauli principle has essentially the form of a superselection rule [61, p. 32], with for instance the Hermitean class sum operators of \( S_N \) (which are non-observable constants of motion), playing the rôle of superselection operators.

2. The link between spin and permutation symmetry, here introduced as a
postulate, has been shown by Pauli [62] to be unavoidable in a consistent quantum field theory.

3. The Pauli postulate can be replaced by another postulate, viz. the permutations of identical particles cannot be observed. See ref. 63 for a discussion of this alternative.

6.3. A model for N-electron systems

In this section some restrictions and approximations will be introduced leading to a widely used model for N-electron systems.

One of the assumptions defining this model is:

M1. There are no spin-dependent contributions to the energy. Consequently the Hamiltonian has the form:

\[ H = H(r) \otimes 1(s), \]

where \( H(r) \) is an energy operator on \( L^2(\mathbb{R}^3) \otimes \mathbb{N} \) and \( 1(s) \) is the unit operator on \( L^2(s) \otimes \mathbb{N} \).

Notes.
1. From here on we will ignore that \( H \) may have antiunitary symmetry operators. This non-essential simplification is introduced as it would go beyond the scope of this work to explain where the presence of antiunitary symmetry operators causes doubling of degeneracies. See ref. 4 for a careful discussion of the cases where this occurs.

2. The assumption M1 and the simplification introduced in the foregoing note imply that the symmetry group \( G_H \) has the form:

\[ G_H = [G_h^{(r)} \otimes SU(2)]^N \otimes S_N, \]

where \( G_h^{(r)} \otimes SU(2) \) is a subgroup of the one-particle space-spin symmetry group \( G_h^s \), introduced in the foregoing section. \( SU(2) \) arises here because every unitary operator on the 2-dimensional space \( L^2(s) \) commutes with the one-electron part of \( H \). It has been pointed out in sec. I.4.6 that it makes no difference whether we consider \( U(2) \) or \( SU(2) \). According to common practice we have chosen the group with the smallest number of parameters.

The second assumption is in general a rather strong approximation, but is usually necessary, for calculations on systems with more than one electron.
and other systems where the Schrödinger equation cannot be solved exactly.

M2a. The Hilbert space \( \mathcal{H}_M \) of the model system is finite dimensional:

\[
\mathcal{H}_M = V_m \otimes (L^2(s) \otimes \mathbb{R}^{3N}),
\]

where \( V_m \) is an appropriately chosen \( m \)-dimensional subspace of \( L^2(\mathbb{R}^{3N}) \). This assumption can be formulated alternatively thus:

M2b. Replace everywhere the unit operator on \( \mathcal{H} \) by the projector onto \( \mathcal{H}_M \).

Let \( \{|k\rangle, k = 1, 2, \ldots, \infty\} \) be an orthonormal basis of \( \mathcal{H} \), and let the first \( M \) vectors span \( \mathcal{H}_M = \mathcal{H}_M' = M \times 2^N \). The assumption M2b then reads:

\[
1_\infty = \sum_{k=1}^{\infty} |k\rangle\langle k| \sim 1_M = \sum_{k=1}^{M} |k\rangle\langle k|.
\]

We obtain a representation \( \Omega_M \) of an observable \( \Omega \) in the following manner:

\[
\Omega = 1_\infty \Omega 1_\infty \sim 1_M \Omega 1_M = \sum_{k=1}^{M} |k\rangle\langle k| \Omega |\ell\rangle\langle \ell| \equiv \Omega_M.
\]

The homomorphism condition for the multiplication of observables is only approximately fulfilled:

\[
(PQ)_M \equiv 1_M (PQ) 1_M \sim (1_M P 1_\infty)(1_\infty \Omega 1_M) \sim (1_M P 1_M)(1_M \Omega 1_M) = P_M \Omega_M.
\]

So strictly speaking \( \Omega_M \) is not a representation of \( \Omega \).

Further it is easy to show that \( \Omega = \Omega^\dagger \) implies \( \Omega_M = \Omega_M^\dagger \).

In practical calculations one often solves the eigenvalue problem of \( H_M \) in order to obtain approximations to eigenstates of the exact Hamiltonian \( H \). The following theorem provides a theoretical foundation for this procedure.

Theorem.

Let the following eigenvalue equations hold:

\[
H |E_k\rangle = E_k |E_k\rangle \quad \text{and} \quad H_M |E_k^{(M)}\rangle = E_k^{(M)} |E_k^{(M)}\rangle.
\]

Then:

(i) \( \lim_{M \to \infty} E_k^{(M)} = E_k \), and the sequence \( \{E_k^{(M)}|M = 1, 2, \ldots\} \) is monotonically decreasing.

(ii) \( \lim_{M \to \infty} |E_k^{(M)} - E_k| = 0 \).

Proof. First replace the eigenvalue equations by variational equations and then invoke the variational principle. See ref. 64, p. 415 ff. for the details of the proof.
How good an approximation one gets by diagonalizing $H_M$ is solely determined by the choice of the model Hilbert space $\mathcal{H}_M$. This choice has been the subject of many computational experiments. In these experiments one nearly always chooses $\mathcal{H}_M$ to be of the form:

$$\mathcal{H}_M = [V_n \otimes L^2(s)] \otimes N,$$

where $V_n$ is an $n$-dimensional subspace of the one-particle Hilbert space (orbital space) $L^2(\mathbb{R}^3)$. Of course $V_n$ must satisfy some basic requirements. For instance it must belong to the domain of the observables of interest. Thus, in order to be able to compute the kinetic energy of the electrons, the elements of $V_n$ must be at least twice differentiable. If, for instance, the eigenstates of $H_M$ have to reflect the symmetry of the exact solutions, $V_n$ must be invariant under the one-particle symmetry group $G^{(r)}_n$, etc.

A model which uses a tensor product of finite-dimensional orbital spaces as the system Hilbert space is often called a multi-configuration model. Examples are provided by such well-known quantum chemical methods as the "limited" configuration interaction method or the "full" configuration interaction method. In both cases $V_n$ is spanned by self-consistent field molecular orbitals; in the limited case only part of the basis of $\mathcal{H}_M$ is employed, whereas in full configuration interaction the complete basis of $\mathcal{H}_M$ is used. Also the valence bond method ($V_n$ is spanned by atomic orbitals) and the multi-configuration self-consistent field method (orbitals and N-electron states are optimized simultaneously) belong to the category of multi-configuration methods.

From here on we will only consider model Hilbert spaces of the form just introduced:

$$\mathcal{H}_M = [V_n \otimes L^2(s)] \otimes N.$$ 

Let $H_M$ be a spin-free operator on this space, it is then easy to show that:

$$H_M = H_M^{(r)} \otimes 1^{(s)},$$

where $H_M^{(r)}$ is an energy operator on $V_n \otimes N$ and $1^{(s)}$ is the unit operator on the spin space. $H_M^{(r)}$ contains the coordinates of the N identical electrons constituting the system and commutes with all operators permuting these coordinates, so $H_M^{(r)}$ is bisymmetric. In sec. I.4.5 it has been proved that the algebra of bisymmetric operators on $V_n \otimes N$ is generated by $\Lambda([GL(n)]^N)$, therefore $H_M^{(r)}$ is expressible as a polynomial in the generators of this algebra. The following formula, in this form due to Matsen
[25], gives this relation explicitly:

$$H_M(x) = \sum_{i,j=1}^{n} \sum_{k=1}^{N} H_{ij}^{(1)} \left[ \sum E^{ij}(k) \right] + \sum_{i,j=1}^{n} \sum_{p,q=1}^{N} H_{ij}^{(2)} \left[ \sum E^{ip}(k) E^{jq}(z) - \delta_{pq} \sum E^{iq}(k) \right].$$

To explain the meaning of the symbols we sketch part of the derivation.

Let \( V_n \) be spanned by the orthogonal basis \( \{|i\rangle | i = 1, \ldots, n \} \), and let the spatial part of the exact Hamiltonian be:

$$H(x) = \sum_{k=1}^{N} H^{(1)}(x_k) + \frac{1}{2} \sum_{k,l=1}^{N} \{ H^{(2)}(x_k, x_l) - \delta_{kl} H^{(2)}(x_k, x_k) \}.$$ 

Write the unit operator on \( V_n \otimes \cdots \otimes V_n \) as:

$$1_M = \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} |i_1 \otimes i_2 \otimes \cdots \otimes i_N \rangle \langle i_1 \otimes i_2 \otimes \cdots \otimes i_N |,$$

where \(|i_1 \otimes i_2 \otimes \cdots \otimes i_N \rangle\) stands for \(|i_1 \rangle \otimes |i_2 \rangle \otimes \cdots \otimes |i_N \rangle\).

Then the one-particle part becomes:

$$H^{(1)}(x) = 1_M H^{(1)}(x) 1_M = \sum_{k=1}^{N} \left( \sum_{i_1=1}^{n} |i_1 \rangle \langle i_1 | \right) \otimes \left( \sum_{i_2=1}^{n} |i_2 \rangle \langle i_2 | \right) \otimes \cdots \otimes \left( \sum_{i_N=1}^{n} |i_N \rangle \langle i_N | \right)$$

$$= \sum_{i,j=1}^{n} H_{ij}^{(1)} \sum_{k=1}^{N} 1 \otimes 1 \otimes \cdots \otimes (|i\rangle \langle j|) \otimes \cdots \otimes 1.$$ 

We have used:

$$\langle i|j \rangle = \delta_{ij}$$

$$\sum_{i=1}^{n} |i\rangle \langle i| = 1$$

$$H^{(1)}_{ij} = \langle i|H^{(1)}|j \rangle$$

and:

$$E^{ij}(k) \equiv 1 \otimes 1 \otimes \cdots \otimes |i\rangle \langle j| \otimes \cdots \otimes 1.$$

Clearly the operator \(|i\rangle \langle j|\) has the matrix \(E^{ij}\), defined in sec. I.2.2 as a basis element of the algebra \(A(n,C)\), which is the Lie algebra of \(GL(n)\), as we have seen in sec. I.3.2. In sec. I.4.3 it has been exhibited that \(A([GL(n)]^N)\) is spanned by elements of the form:

109
\[ \sum_{k=1}^{N} E^{ij}(k). \]

Hence the one-electron part of \(H_M(\hat{r})\) is a linear function of the generators of the Lie algebra of \([GL(n)]^N\).

In exactly the same way one proceeds for the two-electron part of \(H_M(\hat{r})\). Defining: \(H^{(2)}_{ij;pq} = \langle ij | \frac{1}{r_{12}} | pq \rangle\) one gets the final result. Note that the two-electron part is a second order polynomial in the generators of \(\Lambda([GL(n)]^N)\). The fact that \(H_M(\hat{r})\) is such a simple expression follows from the fact that the basis of \(V_n^N\) was taken to be orthogonal.

The model operator \(H_M(\hat{r})\) being bisymmetric has also the interesting consequence that \(GL(n)\) is a dynamical group of the model system.

Definition.
A dynamical group of a system is a Lie group with a Lie algebra containing sufficient generators to express the Hamiltonian of the system as a function of the generators.

Well-known systems with a dynamical group are the three-dimensional harmonic oscillator [13, ch. 10] and the hydrogen atom [65]. In these two cases the dynamical group happens to be also the symmetry group of the system. This has the interesting consequence that the eigenvalues and eigenvectors of these systems can be determined from group theoretical arguments alone, without ever having to solve the Schrödinger equation explicitly.

To exhibit how the knowledge can help us that \(GL(n)\) is a dynamical group (not a symmetry group!) of our model system, we assume that \(V_n^N\) has been decomposed into orthogonal irreducible \(GL(n)\)-carrier spaces. Assemble again all basis vectors carrying the irrep \(\lambda\) in one array:

\[
\begin{bmatrix}
\phi^{\langle \lambda \rangle}_{11} & \phi^{\langle \lambda \rangle}_{12} & \ldots \ldots & \phi^{\langle \lambda \rangle}_{1n} \\
\phi^{\langle \lambda \rangle}_{21} & \ldots \ldots & \ldots & \ldots \\
\phi^{\langle \lambda \rangle}_{f_1^{[\lambda]}} & \ldots \ldots & \phi^{\langle \lambda \rangle}_{f_1^{[\lambda]}} & \ldots \ldots \\
\phi^{\langle \lambda \rangle}_{f_2^{[\lambda]}} & \ldots \ldots & \phi^{\langle \lambda \rangle}_{f_2^{[\lambda]}} & \ldots \ldots \\
\phi^{\langle \lambda \rangle}_{f_3^{[\lambda]}} & \ldots \ldots & \phi^{\langle \lambda \rangle}_{f_3^{[\lambda]}} & \ldots \ldots \\
\phi^{\langle \lambda \rangle}_{f^n_{\lambda}^{[\lambda]}} & \ldots \ldots & \phi^{\langle \lambda \rangle}_{f^n_{\lambda}^{[\lambda]}} & \ldots \ldots \\
\end{bmatrix}
\]

By definition the rows span identical irreps of \(GL(n)\). (From the result of sec. I.4.6 it follows that the columns span identical irreps of \(S_n^N\).) Each row spans also an irrep of the algebra of bisymmetric operators (sec. I.4.5)
and is hence invariant under $H_M(\hat{r})$. Since the different rows are orthogonal, the spaces are non-interacting under $H_M(\hat{r})$, and so this model operator can be diagonalized within each row separately. If we know the explicit representation of the generators of $GL(n)$ carried by the subspaces, we can also write down the explicit representation of the model Hamiltonian on basis of each of the rows. This is the approach to the multi-configuration problem described in ref. 34 and 35.

Notes.

1. We are not yet in a position to consider which irreducible representations of $GL(n)$ are physically realizable. This will be discussed in the next section, where we will find that only irreps with not more than $(2\sigma+1)$ columns are allowed for half-integral $\sigma$.

2. Although, as we have seen, the labelling of eigenstates of $H$ according to irreps of $S_N$ is superfluous since only antisymmetric representations of $S_N$ are physically realizable, the labelling according to irreps of $S_N^{(r)}$, the permutation group acting on spatial coordinates alone, is non-trivial. This justifies the attention paid to the representation theory of $S_N$ in the preceding chapter.

3. Much work [31, 53, 54] has been done to generalize the Wigner-Eckart theorem, so that the matrix elements of the raising and lowering operators $E_{ij}$ can be expressed in terms of the vector coupling coefficients of $GL(n)$. This seems a promising approach to the $N$-electron problem.

6.4. Symmetry restrictions imposed by spin

In this section the same physical model will be considered as in the preceding section, but in a somewhat different manner. Whereas in the second part of the foregoing section the attention was restricted to the spatial part $H_M(\hat{r})$ of the total spin-free Hamiltonian $H_M \equiv H_M(\hat{r}) \otimes 1(\hat{s})$, and accordingly only the spatial part of the model Hilbert space was considered, spin will remain included in the derivations presented subsequently. This will enable us to show that, although spin does not contribute to the dynamics of the model system, it still plays an important rôle by imposing certain symmetry restrictions. In particular we will be able to identify the physically realizable irreps of $S_N^{(r)}$. 

111
For the following exposition the spatial one-particle symmetry is inessential; for reasons of clarity we assume it to be absent. The symmetry group of the system is then:

\[ G_H = [1(\pi)]^N \otimes [SU(2)]^N \otimes S_N \]

where \( 1(\pi) \) is the unit operator on \( L^2(\mathbb{R}^3) \). Since group theory states that inequivalent irreducible carrier spaces of the symmetry group are non-interacting under the Hamiltonian of the system, we do not introduce an extra approximation, when \( H_M \) is diagonalized with respect to a basis of an irreducible subspace of the model Hilbert space

\[ \mathcal{H}_M = (V_n \otimes \ldots \otimes V_n) \otimes (L^2(s) \otimes \ldots \otimes L^2(s)) \].

To construct such a basis we first adapt \( L^2(s) \otimes \ldots \otimes L^2(s) \) to SU(2) and then adapt the whole spatial-spin Hilbert space to \( S_N \); that is, we subsequently anti-symmetrize.

In the foregoing chapters it has been discussed how to adapt a tensor product space to SU(n). We recapitulate the main findings, specialized to spin space:

- \( L^2(s) \otimes \ldots \otimes L^2(s) \) can be decomposed employing Lie algebraic techniques, e.g. ladder operators or vector coupling (sec. I.4.7). This decomposition under \( \Lambda(SU(2)) \) induces a complete decomposition under \( S_N^{(s)} \), the permutation group acting on the spin coordinates (sec. I.4.6).

- Alternatively, \( L^2(s) \otimes \ldots \otimes L^2(s) \) can be decomposed under \( S_N^{(s)} \). This induces a decomposition into irreducible SU(2)-carrier spaces (sec. I.5.5 and I.5.6).

- Each irreducible SU(2)-carrier space \( V_{<\lambda>} \) in \( L^2(s) \otimes \ldots \otimes L^2(s) \) can be labelled by a partition \( <\lambda> \) of \( N \).

- Each space \( V_{<\lambda>} \) occurs \( f_[\lambda] \) times, where \( f_[\lambda] \) is the dimension of the corresponding irrep \( [\lambda] \) of \( S_N^{(s)} \).

- The direct sum \( V_{<\lambda>} \oplus \ldots \oplus V_{<\lambda>} \) (\( f_[\lambda] \) terms) is a maximal eigenspace of the Casimir operator \( S^2 \), and belongs to eigenvalue \( S(S+1) \). So this space may be labelled alternatively by the spin quantum number \( S \) (sec. I.4.7).

- Since \( L^2(s) \) is 2-dimensional, only irreps with not more than two rows are afforded by this Hilbert space (sec. I.5.5).

- Each weight (uniquely characterized by an eigenvalue of \( S_z \)) occurs once in \( V_{<\lambda>} \), for only one standard Weyl-Rumer or Gelfand tensor can be associated with a certain weight (sec. I.5.5 and I.5.6).
It remains to associate the spin quantum number $S$ defined by:

$$S^2 \theta^{\langle \lambda \rangle} (s_1, \ldots, s_N) = S(S+1) \theta^{\langle \lambda \rangle} (s_1, \ldots, s_N)$$

with the corresponding partition $\langle \lambda \rangle$. From angular momentum theory [ref. 17, sec. 2.3] it is known that an eigen-space of $S^2$ with quantum number $S$ is $(2S+1)$-dimensional. A subspace $V^{\langle \lambda \rangle}$ with Young diagram:

```
  ****
  ****
(.N-p)
```

is $(p+1)$-dimensional. (Apply to prove this formula (14) given in ref. 51).

Comparison of dimensions shows that $p = 2S$, so $S$ corresponds to $
\langle \frac{1}{2} N + S, \frac{1}{2} N - S \rangle$.

Notes.

1. The irreducibility postulate states: "Diagonalization of a spin-free Hamiltonian operator $H = H(\vec{r}) \otimes 1(\vec{s})$ induces a decomposition of the spin part of the Hilbert space into irreducible SU(2)-carrier spaces. Each SU(2)-carrier space belongs to one eigenvalue of $H$." (Therefore the dimension of such carrier spaces is called the "spin degeneracy" of the state). This group theoretical statement is equivalent to the quantum mechanical assertion: "The eigenstates of a spin-free Hamiltonian $H$ are simultaneously eigenstates of the constants of motion $S^2$ and $S_z$. The spin-degenerate states are connected via ladder operators."

2. At this point it may seem surprising that the spin Hilbert space $L^2(s) \otimes^N$ will be decomposed by diagonalization of $H$, for $H$ has the unit operator as its spin part. Subsequently we will see that the adaptation of $H_M$ to $S^N_N$ is responsible for this effect.

3. It is transparent how to generalize the theory to arbitrary spin $\sigma$. The Hilbert space is in general $(2\sigma+1)$-dimensional, the spin symmetry group is in general SU$(2\sigma+1)$ and the maximum number of rows allowed in the Young diagram is then $(2\sigma+1)$.

The next step in the adaptation of the model Hilbert space

$$\mathcal{H}_M \equiv (V_n \otimes^N) \otimes (L^2(s) \otimes^N)$$

to the symmetry group $G_H$ is the projection onto the antisymmetric subspace of $H_M$. Doing this we will obviously profit by $L^2(s) \otimes^N$ already being
decomposed under SU(2), and hence under $S_N^{(s)}$.

At this point the choice must be made which of the possible bases of $L^2(s) \otimes N$ is to be considered. Because the actual computations, reported on in the second half of this thesis, were performed on a basis of Weyl-Rumer tensors, we proceed in that representation. (See e.g. ref. 66 or ref. 15 for the corresponding derivations in terms of Gelfand tensors).

As we will see, the NP-structure matrix $X^{\lambda}_{\bar{\lambda}}$ (defined in sec. I.5.1) will arise in the derivation below. It seems that the relevance of this matrix for the antisymmetrization problem has not been recognized before.

In the following theorem the operators on $J_M$ are all factorized into two factors, the first acting on $V_n \otimes N$ and the second on $L^2(s) \otimes N$. The antisymmetrizer $A$ reads in this notation:

$$A = 1/N! \sum_{P \in S_N} \zeta_p \otimes P,$$

Theorem.

$$A(1 \otimes Y_{k\ell}^{[\lambda]}) = 1/N! \sum_{r=1}^{f^\lambda} \tau_{kr} Y_{r\ell}^{[\lambda]} \otimes g_{rl}^{[\lambda]},$$

where:

$Y_{r\ell}^{[\lambda]}$ is the Young unit belonging to the standard tableaux $T_r^{[\lambda]}$ and $T_k^{[\lambda]}$, the associated tableaux of $T_r^{[\lambda]}$ and $T_k^{[\lambda]}$, respectively.

$\tau_{kr}^{[\lambda]}$ is the parity of $T_{kr}^{[\lambda]}$, the permutation which converts $T_r^{[\lambda]}$ into $T_k^{[\lambda]}$.

$g_{rl}^{[\lambda]} \equiv \sum_j \eta_{rj}^{[\lambda]} Y_{jl}^{[\lambda]}$ (a natural unit).

$\eta_{rj}^{[\lambda]}$ is the $(r,j)$-element of the matrix $(X^{[\lambda]}_{\bar{\lambda}})^{-1}$ defined in sec. I.5.1.

Proof.

$$A(1 \otimes Y_{k\ell}^{[\lambda]}) = 1/N! \sum_{P \in S_N} \zeta_p \otimes (PY_{k\ell}^{[\lambda]}).$$

Recall (sec. I.5.1) that

$$L_{\alpha}^{[\lambda]}(\pi)T = U_{\alpha}^{[\lambda]}(\pi^{-1}) (X^{[\lambda]}_{\bar{\lambda}})^{-1},$$

where $U_{\alpha}^{[\lambda]}(\pi^{-1})_{k\ell}$ is the coefficient of $\pi^{-1}$ in $Y_{k\ell}^{[\lambda]}$. It can be seen from the definition of $U_{\alpha}^{[\lambda]}(\pi)$ that:

$$\zeta_p U_{\alpha}^{[\lambda]}(\pi^{-1})_{kr} = \tau_{kr}^{[\lambda]} U_{\alpha}^{[\lambda]}(\pi)_{\bar{k}\ell}$$

and so:
\[ A(1 \otimes \gamma[y^{(\lambda)}]) = \frac{1}{N!} \sum_{r=1}^{f[\lambda]} \sum_{\pi \in \mathcal{P}_{N}} \left( \sum_{x=1}^{d[\lambda]} \eta[x^{(\lambda)}] \otimes \left( \sum_{j=1}^{f[\lambda]} \gamma[j^{(\lambda)}] \right) \right) \]

Notes.

1. This theorem shows that the adaptation of spin space to SU(2) by the Young unit \( \gamma[y^{(\lambda)}] \), followed by the antisymmetrization of \( \mathcal{H}_M \), can also be performed in the following manner:

   (i) Adapt \( V_n \otimes \mathcal{N} \) to \( S_N^{(r)} \) by \( \gamma[x^{(\lambda)}] \), \( x = 1, \ldots, f[\lambda] \).

   (ii) Adapt \( \mathcal{L}^2(s) \otimes \mathcal{N} \) to SU(2), using the natural units \( g[x^{(\lambda)}] \). This gives also adaptation to \( S_N^{(s)} \).

   (iii) Form linear combinations of the tensors so constructed. This amounts to the construction of the Clebsch-Gordan series \( [\lambda] \otimes [\lambda] + S_N \) associated with the restriction of \( S_N^{(r)} \otimes S_N^{(s)} \) to \( S_N \).

So, adaptation of the spin Hilbert space to SU(2) imposes, in cooperation with antisymmetrization, \( S_N^{(r)} \)-symmetry on the spatial part of the wavefunction. See e.g. ref. 63 for a discussion of some of the physical consequences of this observation.

2. In the preceding section it has been shown that diagonalization of \( H_M \) on \( V_n \otimes \mathcal{N} \) decomposes this space into irreducible GL(n)-carrier spaces and hence, in the absence of spatial symmetry, to \( S_N^{(r)} \). Using:

\[ A(\gamma[k^{(\lambda)},\bar{x}] \otimes 1) = A(1 \otimes \gamma[x^{(\lambda)}]) \]

and the theorem above it follows that the eigenstates of \( H_M = H_\tau \otimes 1(\bar{s}) \) are adapted to \( S_N^{(s)} \) and SU(2)-symmetry, i.e. they are eigenfunctions of \( S^2 \). So, antisymmetrization ensures that SU(2)-symmetry does not violate the irreducibility postulate. In more general terms it is the presence of \( S_N \) in the definition of \( G_H \) that takes care that the irreducibility postulate is not violated, even in the case of a spin-free Hamiltonian.

3. Since irreps \( [\lambda] \) of \( S_N^{(s)} \) with more than \((2\sigma+1)\) rows are not afforded by a general spin Hilbert space \( \mathcal{L}^2(s) \otimes \mathcal{N} \), it follows from the theorem above that the only physically realizable states of spin \( \sigma \) fermions are those with not more than \((2\sigma+1)\) columns in their Young diagrams.
6.5. A note on the calculation of matrix elements

The main purpose of this review: the presentation of a self-contained account of the group theoretical properties of N-particle systems has at this point been fulfilled. The second stage in a consistent and logical treatment of such systems would be the calculation of matrix elements, employing to this end all the simplifications that group theory can offer.

A discussion in the manner of the preceding chapters would certainly require more pages than the foregoing exposition. Also, this problem is still the subject of much current research, and therefore it would go far beyond the scope of this thesis to give an account including such exciting new developments as for example the use of double cosets in $S_N$ or $U(n)$-shift operators for the evaluation of matrix elements.

In this section we will just briefly mention some of the existing methods.

If one wishes to employ the Gelfand basis a possible approach is from the side of $\Lambda(\text{GL}(n))$, writing the model Hamiltonian in terms of the generators of this Lie algebra (sec. I.6.3). One then calculates the matrix elements of the generators in this basis [34, 35]. Alternatively, one can use permutation group techniques, reducing the group theoretical part of the problem to the calculation of spin-free fractional parentage coefficients. This has been worked out in detail in ref. 67. Both methods ultimately end up with having to calculate $3n-j$ symbols of $\text{SU}(2)$. (See also ref. 68 for a discussion of the permutation group method).

Using the Weyl-Rumer basis, consisting of canonical valence bond functions, which in our terminology are called "NP-projected tensors characterized by standard index tableaux" (sec. I.5.5), one can proceed in two different ways: "spin-free" or "conventional". To explain the difference we need the following theorem.

Theorem.

$$\langle A(\phi' \otimes y^{[\lambda]} \phi) | H | A(\phi \otimes y^{[\lambda]} \phi) \rangle = C t_{ij}^{[\lambda]} \langle y^{[\lambda]} \phi' | H | y^{[\lambda]} \phi \rangle,$$

where: $a, 1 \leq a \leq f_{[\lambda]}$, is arbitrary;

$C$ is a constant which vanishes after normalization;
\[ \tau^{[\lambda]}_{ij} \] is the parity of \( \phi_{ij}^{[\lambda]} \);
\( \phi, \phi' \in L^2(R^3) \otimes N \);
\( \theta \in L^2(s) \otimes N \) is chosen such that \( Y^{[\lambda]}_{ljg} \theta \) and \( Y^{[\lambda]}_{lg} \theta \) are represented by a standard index tableau.

**Proof.** (Suppress \([\lambda]\), put a tilde over the operator to denote association). First we show:
\[ y^{[\lambda]}_{klpq} = c_{kp} y^{[\lambda]}_{alpq}, \]
where \( c_{kp} \) is a scalar independent of \( l \) and \( q \) and \( a \) is arbitrary. Write:
\[
y^{[\lambda]}_{klpq} = P_k N_l \sigma^l \lambda_k \sigma N^l p q q \] (definition)
\[ = c_{kp} P_k N_l \sigma^l \lambda_q \sigma N^l p q q \] (Von Neumann's theorem [36, p. 18])
\[ = (c_{kp}^{\sigma^l \lambda_q}) P_k N_l \sigma^l \lambda_q \sigma N^l p q q \] (because \( N_k^2 = \sigma^l \lambda_k \sigma N^l p q q \))
\[ = c_{kp} P_k N_l \sigma^l \lambda_q \sigma N^l p q q \] (\( \sigma^l \lambda_q \sigma = 1 \), \( c_{kp} = c_{kp}^{\sigma^l \lambda_q} \))
\[ = c_{kp} y^{[\lambda]}_{alpq}. \]

Using the theorem of the foregoing section, the matrix element becomes:
\[
\frac{1}{N!} \sum_{k,l} \sum_{p,q} \tau_{kj} \tau_{pi} \eta_{kl} \eta_{pq} \langle \tilde{y}^{[\lambda]}_{kj} \phi \mid H \mid \tilde{y}^{[\lambda]}_{pi} \phi \rangle \langle \tilde{y}^{[\lambda]}_{lg} \theta \mid \tilde{y}^{[\lambda]}_{qg} \theta \rangle
\]
Using:
\[ \tau_{kj} \tau_{pi} = (\tau_{kb} \tau_{bj})(\tau_{pb} \tau_{bi}) = (\tau_{kb} \tau_{bp})(\tau_{ib} \tau_{bj}) = \tau_{kp} \tau_{ij} \]
and:
\[ y^{[\lambda]}_{kj} \tau_{pi} = c_{kp} y^{[\lambda]}_{al} \]
we get the matrix element:
\[
\tau_{ij} \langle \tilde{y}^{[\lambda]}_{aj} \phi \mid H \mid \tilde{y}^{[\lambda]}_{al} \phi \rangle \left( \frac{1}{N!} \sum_{k,l} \sum_{p,q} c_{kp} \tau_{kp} \eta_{kl} \eta_{pq} \langle \tilde{y}^{[\lambda]}_{lg} \theta \mid \tilde{y}^{[\lambda]}_{qg} \theta \rangle \right)
\]

The expression between square brackets is independent of \( a, i \) and \( j \), call it \( C \) and the theorem has been proved.

**Notes.**
1. This theorem gives a relation for all matrix elements which do not vanish because of symmetry reasons, for recall:
   - \( y^{[\lambda]}_{ljg} \theta \) and \( y^{[\lambda]}_{lg} \theta \) are of definite weight (i.e. are eigenfunctions of \( S_z \)).
   - Each weight occurs only once in \( \langle \lambda \rangle \).
   - Vectors of different weight are orthogonal.
   - Vectors belonging to different \( \langle \lambda \rangle \) are orthogonal.
   - The spin part of \( H \) is the unit operator.
Hence all other matrix elements of $H$ are zero.

2. The spin-dependent eigenvalue problem gives the same result as the spin-free eigenvalue problem defined in the space:

$$V^{\lambda}_a \equiv Y^{[\lambda]}_{aa} (V_n \otimes N)$$

for arbitrary $a, 1 \leq a \leq f(\lambda)$. The basis of this space consists of Weyl-Rumer tensors of the NP-type (sec. I.5.5).

The "conventional" way of calculating matrix elements proceeds by evaluating the antisymmetrizers in the left hand side of the expression in the theorem above. This yields:

$$\sum_{PCS_n} \zeta \, \langle \phi' | HP | \phi \rangle \langle Y^{[\lambda]}_{ijg} | p | Y^{[\lambda]}_{i'j'g} \rangle.$$ 

Pauling [69] was the first to consider the evaluation of the spin matrix element arising in this evaluation. Or more precisely, he considered the special case of eigenfunctions of $S^2$ which are of the form:

$$(a \beta - \beta a)(a \beta - \beta a) \cdots (a \beta - \beta a).$$

These functions are obtained by the action of the Young unit belonging to the tableau:

$$\begin{array}{cccccc}
1 & 3 & 5 & 7 & \cdots & \\
2 & 4 & 6 & \cdots & \\
\end{array}$$

onto the function:

$$(a \beta)(a \beta) \cdots (a \beta).$$

Clearly these projected functions have the eigenvalue $S(S+1) = 0$. Pauling imposed the Rumer diagram of the bra on the Rumer diagram of the permuted ket. This gives a superposition pattern, the analysis of which enables the formulation of very simple rules for the calculation of the spin matrix elements. Pauling's method has been extended to general spin multiplicities by, among others, Reeves [38], Cooper and McWeeny [39] and Sutcliffe [40]. Reeves published his results in the form of a computer algorithm which forms an important part of our programs. The reader is referred to the work of Van Berkel [47], who, following the Cooper and McWeeny paper, has described in detail the calculation of the spin matrix elements and the manner in which Reeves' algorithm has been implemented in our programs.

The "spin-free" method on the other hand starts working on the right hand side of the expression in the theorem above. This yields:
\( \langle \phi' | H (Y^{\lambda}_{\omega j})^\dagger (Y^{\lambda}_{\omega i}) | \phi \rangle \propto \langle \phi' | H P_j [\tilde{\lambda}] N_j [\tilde{\lambda}] P_j [\tilde{\lambda}] 0_{\tilde{\lambda} \tilde{i}} | \phi \rangle \equiv \sum_{P \in S_N} [P]_{j \tilde{i}} [\tilde{\lambda}] \langle \phi' | H P | \phi \rangle. \)

The scalars \([P]_{j \tilde{i}} [\tilde{\lambda}]\) are sometimes called Pauling numbers \([70]\), they clearly satisfy:

\[ [P]_{j \tilde{i}} [\tilde{\lambda}] \propto \zeta_P \langle Y^{\lambda}_{j g} \Theta | P | Y^{\lambda}_{i g} \Theta \rangle. \]

It is possible to evaluate Pauling numbers by purely group algebraic arguments, see Roël's dissertation \([71]\). Roël uses for the evaluation a double coset \([68, 70]\) decomposition of \(S_N\) with respect to the normalizer \([4, \text{p. 15}]\) of the row group \(R_j^{[\lambda]}\). He then proves that all elements belonging to one double coset have the same Pauling number and that elements belonging to different double cosets have different Pauling numbers. These double cosets can be uniquely characterized by diagrams that are in one-to-one correspondence with Pauling's superposition patterns and their generalization to arbitrary spin multiplicities.

Notes.
1. All that has been said in this section about the calculation of matrix elements of the spin-free Hamiltonian holds for other spin-free operators too.
2. In the case of spin-dependent operators it is often advantageous to reduce the total Hilbert space first with respect to \(S^{(r)}_N\) and \(S^{(s)}_N\) and to apply the Wigner-Eckart theorem to space and spin matrix elements separately \([72, 73]\). However, not all spin-dependent operators are adapted to \(S^{(r)}_N\) and \(S^{(s)}_N\), although every physical operator is of course symmetric under \(S^{(r)}_N\). In ref. \(74\) it has been discussed how operators can be classified according to irreps of \(S^{(r)}_N\) and \(S^{(s)}_N\) separately.
REFERENCES PART I

1. B.R. Judd,
   Operator Techniques in Atomic Spectroscopy.

2. B.G. Wybourne,
   Symmetry Principles and Atomic Spectroscopy.

3. J. Tropfke,
   Geschichte der Elementar-Mathematik, Zw. Band,
   Dritte Auflage, Walter de Gruyter, Berlin, 1933.

4. L. Jansen, M. Boon,
   Theory of Finite Groups.
   North Holland, Amsterdam, 1967.

5. G.C. Shephard,
   Vector Spaces of Finite Dimension.

6. H. Weyl,
   The Theory of Groups and Quantum Mechanics.
   Translated by H.P. Robertson.

7. H. Weyl,
   The Classical Groups.

8. H. Boerner,
   Darstellungen von Gruppen.

9. C.W. Curtis, I. Reiner,
   Representation Theory of Finite Groups and Associative Algebras.

10. P.E.S. Worner,
    Representatie Theorie van Eindige Groepen en Eindig Dimensionale
    K-algebra's.

11. D.J. Klein,
    Finite Groups and Semi-Simple Algebras in Quantum Mechanics.
    To appear in: Group Theory and its Applications III.

12. W.H. Greub,
    Linear Algebra, 2nd edition.

13. W. Miller Jr.,
    Symmetry Groups and their Applications.
14. P.O. Löwdin,
Rev. Mod. Phys. 39, 259 (1967).

15. P.E.S. Wormer,
Different Orbitals for Different Spins.
Afstudeerverslag, Technological University Delft, The Netherlands,
June 1969.

16. M. Hausner, J.T. Schwartz,
Lie Groups; Lie Algebras.

17. A.R. Edmonds,
Angular Momentum in Quantum Mechanics.

18. F.D. Murnaghan,
The Theory of Group Representations.

19. F.D. Murnaghan,
The Unitary and Rotation Groups.

20. W. Rudin,

21. A. Messiah,
Quantum Mechanics.
North Holland, Amsterdam, 1965.

22. J.J. de Swart,
Group Theory, Part B.
Lecture Notes, Institute for Theoretical Physics,
University of Nijmegen, The Netherlands.

23. E.T. Whittaker,
Analytical Dynamics of Particles and Rigid Bodies, 4th edition.

24. K.J. Lezuo,

25. F.A. Matsen,
The Unitary Group Formulations of the Many-Electron Problem.
Preprint.

26. M. Hamermesh,
Group Theory.
Addison-Wesley, Reading, 1962.

27. D.E. Knuth,
Addison-Wesley, Reading, 1968.
28. S. Lang,
   Algebra.
   Addison-Wesley, Reading, 1965.

29. N. Jacobson,
   Lie Algebras.

30. L.C. Biedenharn,

31. J.D. Louck,

32. R.L. Hudson,

33. P.A.M. Dirac,

34. J. Paldus,

35. J.F. Gouyet, R. Schranner, T.H. Seligman,

36. D.E. Rutherford,
   Substitutional Analysis.

37. F.A. Matsen, A.A. Cantu, R.D. Poshusta,

38. C.M. Reeves,

39. I.L. Cooper, R. McWeeny,

40. B.T. Sutcliffe,

41. A.J. Coleman,

42. F.A. Matsen,

43. T. Yamanouchi,

44. P.E.S. Wormer, A. van der Avoird,
   (Reprinted in this thesis).
45. G. Rumer, E. Teller, H. Weyl,

46. R. Pauncz,
   Alternant Molecular Orbital Method.

47. T. van Berkel,
   Ab Initio Berekening van de Van der Waals Wisselwerking tussen
   Twee He-Atomen.
   Hoofdvakverslag, University of Nijmegen, 1974.

48. T.H. Seligman,

49. P.E.S. Wormer,
   (Reprinted in this thesis).

50. D.E. Littlewood,

51. P.E.S. Wormer, A. van der Avoird,
   (Reprinted in this thesis).

52. M. Brunet, T.H. Seligman,
   (To be published).

53. W.J. Holman III, L.C. Biedenharn, in:
   Group Theory and Its Applications.
   Ed. E.M. Loebl.

54. M. Moshinsky,
   Group Theory and the Many-Body Problem.

55. E. Prugovecki,
   Quantum Mechanics in Hilbert Space.

56. E.T. Copson,
   Metric Spaces.

57. M. Boon,
   Private communication.

58. V. Pock,
59. H.V. McIntosh, in: 
    Group Theory and Its Applications II. 
    Ed. E.M. Loebel. 

60. D.J. Klein, 
    Private communication.

61. P. Roman, 
    Advanced Quantum Theory. 
    Addison-Wesley, Reading, 1965.

62. W. Pauli, 
    Phys. Rev. 58, 716 (1940).

63. A. van der Avoird, 
    Identiteit. 

64. E.C. Kemble, 
    The Fundamental Principles of Quantum Mechanics. 

65. C.E. Wulfman, in: 
    Group Theory and Its Applications II. 
    Ed. E.M. Loebel. 

66. A. van der Avoird, 
    Perturbation Theory for Intermolecular Forces. 

67. I.G. Kaplan, 
    Symmetry of Many-Electron Systems. 

68. P. Kramer, T.H. Seligman, 

69. L. Pauling, 

70. D.J. Klein, C.H. Carlisle, F.A. Matsen, 

71. R.W.J. Roël, 
    University of Amsterdam.

72. I.L. Cooper, J.I. Musher, 

73. I.L. Cooper, J.I. Musher, 

74. A. van der Avoird, P.E.S. Wormer, 
    (Reprinted in this thesis).