Integral and non-integral dimension of modules over non-commutative rings.

Sejal K. Dharia
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation
DOI: https://doi.org/10.31979/etd.evmm-y6dp
https://scholarworks.sjsu.edu/etd_theses/3988

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.
INTEGRAL AND NON-INTEGRAL DIMENSION OF MODULES OVER NON-COMMUTATIVE RINGS

A Thesis

Presented to

The Faculty of the Department of Mathematics
San José State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

by
Sejal K. Dharia
December 2009
SAN JOSÉ STATE UNIVERSITY

The Undersigned Thesis Committee Approves the Thesis Titled

INTEGRAL AND NON-INTEGRAL DIMENSION OF MODULES OVER
NON-COMMUTATIVE RINGS

by
Sejal K. Dharia

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

Dr. Timothy Hsu, Department of Mathematics  9/16/09

Dr. Jane Day, Department of Mathematics  9-16-09

Dr. Brian Peterson, Department of Mathematics  9-16-09

APPROVED FOR THE UNIVERSITY

David L. Bush  12/4/09

Associate Dean  Office of Graduate Studies and Research  Date
ABSTRACT

INTEGRAL AND NON-INTEGRAL DIMENSION OF MODULES OVER
NON-COMMUTATIVE RINGS

by Sejal K. Dharia

This thesis studies the conditions under which certain $U(G)$-modules have integral dimension. Specifically, the paper explores the circumstances under which certain $W(G)$-modules are free modules. Two main theorems are proven for this endeavor. The first theorem takes the matrix representation $A$ of a submodule $N$ of a free left module $M$ over an arbitrary ring $R$. The theorem states that if $A$ can be put into reduced row echelon form, then $N$ and $M/N$ are both free left $R$-modules.

The second theorem takes the matrix representation $B$ of an arbitrary left $R$-module homomorphism $T$. If $B$ can be put into reduced column echelon form, then $\text{Ker}(T)$ is a free left $R$-module as well. The results of these theorems are well-known if $R$ is a field or division ring. The results are also probably known for an arbitrary ring $R$, but there does not seem to be a readily accessible source in the standard literature.

Recall that the group von Neumann algebra, denoted $\mathcal{N}(G)$, is the ring of all bounded $C[G]$-module endomorphisms on $\ell^2(G)$. Von Neumann dimension is defined for Hilbert $\mathcal{N}(G)$-modules. Using results from Lück [Lück02], this paper demonstrates that a Hilbert $\mathcal{N}(G)$-module can have fractional von Neumann dimension. Lück shows there is a dimension-preserving bijection between Hilbert $\mathcal{N}(G)$-modules and finitely generated projective left $\mathcal{N}(G)$-modules, and thus there exists a projective left $\mathcal{N}(G)$-module, $P$, of of fractional dimension. Furthermore, per Lück, the tensor product $U(G) \otimes_{\mathcal{N}(G)} P$, which is itself a left $U(G)$-module, has the same fractional dimension as $P$. It is therefore non-trivial to show that some $U(G)$-modules have integral dimension.
First of all, I would like to thank my thesis advisor, Dr. Timothy Hsu, for his guidance, patience, humor, and most importantly, his tireless red pen! I would also like to thank Dr. Brian Peterson and Dr. Jane Day for their participation in this project and their helpful suggestions.

Thank you to my family, especially my husband Karl, for their support throughout the entire process of completing my master's degree. Finally, I would like to acknowledge and thank all of my math professors at SJSU for all of their care and encouragement. My time at SJSU was very rewarding and it can mostly be attributed to their efforts.
# TABLE OF CONTENTS

## CHAPTER

1. **INTRODUCTION**

2. **MODULES**
   - 2.1 Basic Definitions and Examples
   - 2.2 $R$-module Homomorphisms
   - 2.3 Submodules
   - 2.4 Direct Product and Direct Sum
   - 2.5 Free Modules
   - 2.6 Modules over Division Rings
   - 2.7 Group Algebras

3. **LINEAR ALGEBRA OVER NON-COMMUTATIVE RINGS**
   - 3.1 Conventions
   - 3.2 Matrix Multiplication
   - 3.3 Coordinate Matrices
   - 3.4 Row Operations
   - 3.5 Column Operations
   - 3.6 Some Sufficient Conditions for Freeness
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>(\ell^2(G))</td>
<td>66</td>
</tr>
<tr>
<td>4.1</td>
<td>Review of Analysis/ Topology</td>
<td>66</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Basic Topology of Metric Spaces</td>
<td>66</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Sequences, Limits, and Convergence</td>
<td>71</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Series</td>
<td>76</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Continuity</td>
<td>77</td>
</tr>
<tr>
<td>4.2</td>
<td>Topological Vector Spaces</td>
<td>79</td>
</tr>
<tr>
<td>4.3</td>
<td>Convergence Theorems</td>
<td>87</td>
</tr>
<tr>
<td>4.4</td>
<td>(\ell^2(G))</td>
<td>90</td>
</tr>
<tr>
<td>4.5</td>
<td>Complete Orthonormal Sets</td>
<td>96</td>
</tr>
<tr>
<td>4.6</td>
<td>Projections</td>
<td>99</td>
</tr>
<tr>
<td>5</td>
<td>GROUP VON NEUMANN ALGEBRAS</td>
<td>106</td>
</tr>
<tr>
<td>5.1</td>
<td>Bounded Linear Operators</td>
<td>106</td>
</tr>
<tr>
<td>5.2</td>
<td>(\ell^2(G)) is a (C[G])-bimodule</td>
<td>109</td>
</tr>
<tr>
<td>5.3</td>
<td>Group von Neumann Algebras</td>
<td>113</td>
</tr>
<tr>
<td>5.4</td>
<td>Examples when (G) is Finite</td>
<td>124</td>
</tr>
<tr>
<td>5.5</td>
<td>Applications of Chapter 2</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>BIBLIOGRAPHY</td>
<td>138</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

The “traditional” definition of dimension signifies the number of elements in a basis for a vector space $V$ over a field $F$. For example, the dimension of the complex numbers over the reals is 2, that is, $\dim_{\mathbb{R}} \mathbb{C} = 2$, or the dimension of the vector space of $m \times n$ matrices over a field $F$ is $mn$, that is, $\dim_F M_{m \times n}(F) = mn$. It would be “odd” to think of dimension as having a fractional or real-number value. However, as we shall see, there are alternate notions of dimension of a module that allow for fractional/real number values, specifically, von Neumann dimension and $\mathcal{U}(G)$-dimension (where $G$ is a countable group, and $\mathcal{U}(G)$ is a ring to be described later).

The motivation behind this thesis is as follows. The results of this thesis will be used in [BDH] which studies modules over a ring $\mathcal{U}(G)$. The $\mathcal{U}(G)$-dimensions of these modules are called $\ell^2$-Betti numbers. There are certain topology and algebra problems that may be solved if it can be shown that these $\ell^2$-Betti numbers have an integral (i.e., whole number) value. The conclusion of this thesis discusses the conditions under which certain $\mathcal{U}(G)$-modules have integral $\mathcal{U}(G)$-dimension.

This paper is divided into four main chapters (after the Introduction). Chapter 2 offers background information on (left and right) modules over a ring $R$. Examples of modules are given and the notions of $R$-module homomorphisms, submodules,
and direct sums/products are examined. The latter sections of Chapter 2 define free $R$-modules, modules over division rings, and group algebras, and discuss theorems associated with these objects. The discussion on free modules is particularly important, as it is needed to understand the theorems at the end of Chapter 3 and Chapter 5. In particular, it is important to note that free $R$-modules always have integral dimension. The theorems related to group algebras are also essential to follow some of the discussion in Chapter 5.

Chapter 3 looks at standard linear algebra theory for matrices over an arbitrary ring $R$. The topics that are discussed include matrix multiplication, coordinate matrices, and row/column operations. The final section of the chapter uses matrix operations to examine the conditions under which certain $R$-modules are actually free $R$-modules. The two theorems at the end of Chapter 3 are important for proving the results at the end of Chapter 5.

Chapter 4 offers a review of topics from analysis and topology including metric spaces, sequences, series, limits, and convergence. The notions of normed linear space, inner product space, and Hilbert space are defined, and important theorems related to these spaces are proven as well. The space $\ell^2(G) \subseteq \mathbb{C}^G$ is defined, and its relationship with the group algebra, $\mathbb{C}[G]$ is established. The final sections of Chapter 4 explore complete orthonormal sets, closed subspaces, and orthogonal projections. The latter two concepts are used in definitions and theorems in Chapter 5.

Chapter 5 defines the notions of bounded operators, group von Neumann algebra (denoted $\mathcal{N}(G)$), Hilbert $\mathcal{N}(G)$-modules, and von Neumann dimension. Section 5.4 demonstrates a particular circumstance under which von Neumann dimension may take on a fractional value. Section 5.5 relates the notion of von Neumann dimension to $\mathcal{U}(G)$-dimension, and uses the results of Chapter 3, section 3.6, to demonstrate the conditions under which certain $\mathcal{U}(G)$-modules have integral dimension.
CHAPTER 2

MODULES

This chapter offers a brief overview of module theory and assumes some familiarity with basic concepts from abstract and linear algebra. We use Hungerford [Hun74, Ch. IV] as our standard reference.

2.1 Basic Definitions and Examples

Definition 2.1.1. Let $R$ be a ring with (multiplicative) identity $1_R$. A **left $R$-module** is an additive abelian group $A$ together with a function $R \times A \rightarrow A$ (given by $(r, a) \mapsto ra$) such that for all $r, s \in R$ and $a, b \in A$:

1. $r(a + b) = ra + rb$
2. $(r + s)a = ra + sa$
3. $r(sa) = (rs)a$
4. $1_Ra = a$

Definition 2.1.2. Let $R$ be a ring with (multiplicative) identity $1_R$. A **right $R$-module** is an additive abelian group $A$ together with a function $A \times R \rightarrow A$ (given by $(a, r) \mapsto ar$) such that for all $r, s \in R$ and $a, b \in A$:

1. $(a + b)r = ar + br$
\( a(r + s) = ar + as \)

\( (ar)s = a(rs) \)

\( a1_R = a \)

Note that a module over a division ring \( D \) is referred to as a \textbf{vector space}.

Additionally, the function \( R \times A \to A \) [or \( A \times R \to A \)] in Definitions 2.1.1 and 2.1.2 is also referred to as a \textbf{left [resp. right] action} of \( R \) on \( A \).

Further note that all future references to "ring" will mean "ring with identity", unless otherwise noted.

**Definition 2.1.3.** Let \( R \) and \( S \) be rings. An abelian group \( A \) is an \textbf{\( R-S \) bimodule} if \( A \) is both a left \( R \)-module and a right \( S \)-module, and \( r(as) = (ra)s \) for all \( a \in A, r \in R, s \in S \). If \( R = S \), then \( A \) is called an \textbf{\( R \)-bimodule}.

**Theorem 2.1.4.** Let \( R \) be a commutative ring and let \( A \) be a left \( R \)-module. Then \( A \) can be given the structure of a right \( R \)-module with right action \( ar = ra \) for all \( r \in R \) and \( a \in A \).

\textbf{Proof.} Let \( a, b \in A \) and let \( r, s \in R \). First of all,

\( (a + b)r = r(a + b) = ra + rb = ar + br. \) \hspace{1cm} (2.1)

Furthermore,

\( a(r + s) = (r + s)a = ra + sa = ar + as. \) \hspace{1cm} (2.2)

Additionally,

\( (ar)s = (ra)s = s(ra) \) \hspace{1cm} (*)

\( = (sr)a \)

\( = (rs)a \)

\( = a(rs), \) \hspace{1cm} (2.3)
where (*) follows because \( ra \in A \). It follows that \( A \) is a right \( R \)-module.

Note that Theorem 2.1.4 also holds, *mutatis mutandis*, when \( A \) is a right \( R \)-module.

**Theorem 2.1.5.** If \( R \) is a ring with additive identity \( 0_R \) and \( A \) is a left \( R \)-module with additive identity \( 0_A \), then for all \( a \in A \) and \( r \in R \), we have:

\[
(1) \ r0_A = 0_A, \text{ and} \\
(2) \ 0_RA = 0_A.
\]

**Proof.**

(1) Since \( A \) is an \( R \)-module, \( r0_A + 0_A = r0_A = r(0_A + 0_A) = r0_A + r0_A. \) By cancellation in the additive group \( A \), it follows that \( r0_A = 0_A. \)

(2) Since \( A \) is an \( R \)-module, \( 0_RA + 0_A = 0_RA = (0_R + 0_R)a = 0_RA + 0_RA. \) By cancellation in the additive group \( A \), it follows that \( 0_RA = 0_A. \)

**Theorem 2.1.6.** Let \( R \) be a ring and \( A \) a left \( R \)-module. Then for all \( r \in R \) and \( a \in A \), \( (-r)a = -(ra) = r(-a). \)

**Proof.** The element \( -(ra) \in A \) is the (unique) additive inverse of \( ra \) in \( A \). Thus, if we can show that \( (-r)a + ra = 0_A \), then we have \( (-r)a = -(ra) \). By property (2) of a left \( R \)-module and Theorem 2.1.5, \( (-r)a + ra = (-r + r)a = 0_RA = 0_A. \) It follows that \( (-r)a = -(ra). \) In the same way, by property (1) of a left \( R \)-module, \( r(-a) + ra = r(-a + a) = r0_A = 0_A. \) It follows that \( r(-a) = -(ra). \)

Note: Theorems 2.1.5 and 2.1.6 also hold, *mutatis mutandis*, for right \( R \)-modules and \( R \)-\( S \) bimodules.

Examples of Modules:
(1) Every additive abelian group \( G \) is a \( \mathbb{Z} \)-module, with \( na \ (n \in \mathbb{Z}, a \in G) \) given by \( a + a + \cdots + a \) \((n \text{ times})\).

(2) If \( R \) is a ring, then its subrings \( \{0\} \) and \( R \) are \( R \)-bimodules. More generally, if \( I \) is a left ideal of a ring \( R \), then \( I \) is a left \( R \)-module with \( ra \ (r \in R, a \in I) \) being the ordinary product in \( R \).

(3) If \( R \) is a ring and \( S \) is a subring, then \( R \) is an \( S \)-bimodule where \( ar \) and \( ra \) \((a \in S, r \in R)\) are multiplication in \( R \).

### 2.2 \( R \)-module Homomorphisms

**Definition 2.2.1.** Let \( A \) and \( B \) be left modules over a ring \( R \). A function \( f : A \to B \) is a **left \( R \)-module homomorphism** provided that for all \( a, c \in A \) and \( r \in R \), we have

1. \( f(a + c) = f(a) + f(c) \) and
2. \( f(ra) = rf(a) \).

If \( f \) satisfies the conditions above and is a bijective map, then \( f \) is a **left \( R \)-module isomorphism** and we say that \( A \) is isomorphic to \( B \) (denoted \( A \cong B \)). A **right \( R \)-module homomorphism** is defined as above, except that \( A \) and \( B \) are right \( R \)-modules, and instead of (2) we have \( f(ar) = f(a)r \). An **\( R-S \) bimodule homomorphism** is both a left \( R \)-module and right \( S \)-module homomorphism.

Henceforth we will write 0 to represent the additive identity of a group or ring when the meaning is clear from the context.

Examples of \( R \)-module homomorphisms:
(1) Let $A$ and $B$ be modules over a ring $R$. The zero map $0 : A \to B$ given by $a \mapsto 0$ ($a \in A$) is a left and right $R$-module homomorphism as well as an $R$-bimodule homomorphism.

Proof. Let $a, c \in A$ and $r \in R$. Then $f(a + c) = 0 = 0 + 0 = f(a) + f(c)$ and $f(ra) = 0 = r0 = rf(a)$. □

(2) Fix $x \in R$. Let $f : R \to R$ be given by $f(a) = ax$. Then $f$ is a left $R$-module homomorphism of the left $R$-module $R$.

Proof. Let $a, b, \text{ and } r \in R$. Then $f(a + b) = (a + b)x = ax + bx = f(a) + f(b)$ and $f(ra) = (ra)x = r(ax) = rf(a)$. (Both results use the definition of a ring). □

(3) Let $A$ be an $R$-bimodule. Fix $x \in R$ and define $f : A \to A$ by $f(a) = ax$ for all $a \in A$. Then $f$ is a left $R$-module homomorphism but not necessarily a right $R$-module homomorphism.

Proof. Let $a, b \in A$, and $r \in R$. Then $f(a + b) = (a + b)x = ax + bx = f(a) + f(b)$. Additionally, $f(ra) = (ra)x = r(ax) = rf(a)$, but $f(ar) = (ar)x$ which is not necessarily equal to $(ax)r$. Thus, $f$ may not be a right $R$-module homomorphism. □

Definition 2.2.2. Let $A$ and $B$ be left $R$-modules over a ring $R$ and let $f : A \to B$ be a left $R$-module homomorphism. The kernel of $f$ (denoted $\text{Ker } f$) is the set $\{a \in A | f(a) = 0\}$. The image of $f$ (denoted $\text{Im } f$) is the set $\{f(a) | a \in A\}$.

Theorem 2.2.3. Let $R$ be a ring, and let $f : A \to B$ be a left $R$-module isomorphism. Then if $g : B \to A$ is the inverse function of $f$, it follows that $g$ is a left $R$-module isomorphism.
Proof. Since \( g \) is the inverse function of \( f \), it is bijective. It remains to show that \( g \) is an \( R \)-module homomorphism. Suppose \( b_1, b_2 \in B \) and \( r \in R \). Then \( b_1 = f(a_1) \) and \( b_2 = f(a_2) \) for some \( a_1, a_2 \in A \), and \( g(b_1) = a_1 \) and \( g(b_2) = a_2 \) (since \( g \) is the inverse function of \( f \)). So

\[
g(b_1 + b_2) = g(f(a_1) + f(a_2))
\]

\[= g(f(a_1 + a_2)) \quad (\ast)\]

\[= a_1 + a_2\]

\[= g(b_1) + g(b_2)\]

Additionally,

\[
g(rb_1) = g(r(f(a_1)))
\]

\[= g(f(ra_1)) \quad (**)
\]

\[= ra_1
\]

\[= rg(b_1)
\]

where \((\ast)\) and \((**\)) follow from the fact that \( f \) is a left \( R \)-module homomorphism. \( \square \)

**Definition 2.2.4.** Let \( R \) be a ring and let \( V \) be a left \( R \)-module. Then a left \( R \)-module homomorphism from \( V \) to \( V \) is called a **left \( R \)-module endomorphism**, and the set of all such functions is denoted \( \text{End}(R^V) \). Similarly, if \( V \) is a right \( R \)-module, then the set of all right \( R \)-module endomorphisms from \( V \) to \( V \) is denoted \( \text{End}(V_R) \).

Note that per Lam [Lam01, Ch. 1, Sect. 3], elements \( f \) of \( \text{End}(R^V) \) operate on the right, that is, \((ra)f = r(a)f\) for all \( a \in V \), \( r \in R \). Similarly, we adopt the convention that elements \( f \) of \( \text{End}(V_R) \) operate on the left, with \( f(ar) = f(a)r \).

**Theorem 2.2.5.** Let \( R \) be a ring and let \( V \) be a right \( R \)-module. Then the set \( \text{End}(V_R) \) is a ring with addition given by

\[
(f + g)(a) = f(a) + g(a),
\]

\((2.6)\)
and multiplication given by
\[(fg)(a) = f(g(a))\] (2.7)
for \(f, g \in \text{End}(V_R)\) and \(a \in V_R\).

Proof. Suppose \(f, g,\) and \(h \in \text{End}(V),\) \(a, b \in V,\) and \(r \in R.\) Then
\[(f + g)(ar + b) = f(ar + b) + g(ar + b)\]
\[= [f(a)r + f(b)] + [g(a)r + g(b)]\] (2.8)
\[= [f(a)r + g(a)r] + [f(b) + g(b)] \quad (*)\]
\[= (f + g)(ar + (f + g)(b)).\]

where (*) follows by associativity and commutativity of addition in (the additive group) \(V.\) Furthermore,
\[(f + g)(a) = f(a) + g(a) = g(a) + f(a) = (g + f)(a),\] (2.9)
by commutativity of addition in (the additive group) \(V.\) Additionally,
\[((f + g) + h)(a) = (f + g)(a) + h(a)\]
\[= [f(a) + g(a)] + h(a)\]
\[= f(a) + [g(a) + h(a)] \quad (*)\]
\[= [f + (g + h)](a),\]
where (*) follows by associativity of addition in (the additive group) \(V.\)

Let 0 be the zero homomorphism. Then 0 \(\in \text{End}(V),\) and
\[(f + 0)(a) = f(a) + 0(a) = f(a)\]
\[= 0 + f(a) = 0(a) + f(a).\] (2.11)
Finally, define \((-f)(a) = -f(a)\). Then

\[
(-f)(ar + b) = -f(ar + b)
\]

\[
= -[f(a)r + f(b)]
\]

\[
= -f(a)r - f(b)
\]

\[
= (-f)(a)r + (-f)(b),
\]

and

\[
[(-f) + f](a) = (-f)(a) + f(a) = -f(a) + f(a) = 0.
\]

It follows that \(\text{End}(V_R)\) is an abelian group.

In terms of multiplication, we have that

\[
(fg)(ar + b) = f(g(ar + b)) = f(g(a)r + g(b))
\]

\[
= f(g(a)r) + f(g(b)) = (fg)(a)r + (fg)(b).
\]

Furthermore, \(f(gh) = (fg)h\) by the associativity of function composition.

Additionally, if we let \(\text{id}\) denote the identity map in \(\text{End}(V_R)\), then since \((f \text{id}) = f = (\text{id} f)\), it follows that \(\text{id}\) is the identity element of \(\text{End}(V_R)\).

It remains to prove the left and right distributive laws. We have

\[
f(g + h)(a) = f[(g + h)(a)] = f[g(a) + h(a)] = f(g(a)) + f(h(a))
\]

\[
= (fg)(a) + (fh)(a) = (fg + fh)(a).
\]

The right distributive law is proved similarly. It follows that \(\text{End}(V_R)\) is a ring. \(\square\)

**Theorem 2.2.6.** Let \(R\) be a commutative ring, \(V\) a right \(R\)-module, and \(f \in \text{End}(V_R)\). Let \(c \in R\) and \(a \in V_R\). Then the function \((cf)\) given by \((cf)(a) = cf(a)\) is an element of \(\text{End}(V_R)\).

**Proof.** Let \(a, b \in V_R\) and let \(\text{We have that}\)

\[
(cf)(ar + b) = c[f(ar + b)] = c[f(a)r + f(b)] = cf(a)r + cf(b) = (cf)(a)r + (cf)(b),
\]

\[
(2.16)
\]
where $cf(a)$ and $cf(b) \in A$ by Theorem 2.1.4, since $R$ is commutative.

Note that Theorems 2.2.5 and 2.2.6 also hold, *mutatis mutandis*, for $\text{End}(R_V)$.

**Definition 2.2.7.** Let $V$ be a left $R$-module. Then we define $GL(R_V)$ to be the set of all left $R$-module isomorphisms from $V$ to $V$. Similarly, if $V$ is a right $R$-module, then $GL(V_R)$ is the set of all right $R$-module isomorphisms from $V$ to $V$.

Note that although $GL(R_V)$ is a subset of $\text{End}(R_V)$, it is not a subring of $\text{End}(R_V)$ because the sum of two elements of $GL(R_V)$ may not necessarily be in $GL(R_V)$. Specifically, if $f, -f \in GL(R_V)$, then the sum $(f + (-f)) = 0$ (the zero map) will not be in $GL(R_V)$, as this map is not bijective. However, $GL(R_V)$ is the group of units of $\text{End}(R_V)$.

### 2.3 Submodules

**Definition 2.3.1.** Let $R$ be a ring, $A$ a left $R$-module and $B$ a non-empty subset of $A$. We say that $B$ is a (left) submodule of $A$ provided that $B$ is an additive subgroup of $A$ and $rb \in B$ for all $r \in R$, $b \in B$.

**Theorem 2.3.2.** If $B$ is a submodule of a left $R$-module $A$, then $B$ is a left $R$-module.

**Proof.** Since $B$ is a submodule of $A$, it is an additive (abelian) subgroup of $A$. Additionally, since $rb \in B$ for all $r \in R$, $b \in B$, the operation $R \times B \to B$ given by $(r, b) \mapsto rb$ is well-defined. Finally, since $A$ is an $R$-module and $B \subseteq A$, the elements of $B$ satisfy properties (1)-(4) in the definition of an $R$-module.

Note that Definition 2.3.1 and Theorem 2.3.2 also hold, *mutatis mutandis*, for right $R$-modules and $R$-$S$ bimodules.

Examples of submodules:
(1) If \( R \) is a ring and \( A \) is a left \( R \)-module, then \( A \) and \( \{0\} \) are both submodules of \( A \).

(2) If \( R \) is a ring and \( f : A \to B \) is a left \( R \)-module homomorphism, then \( \ker f \) is a left submodule of \( A \) and \( \im f \) is a left submodule of \( B \).

**Proof.** (a) Since \( f \) is a homomorphism of additive groups, \( \ker f \) is an additive subgroup of \( A \). Suppose \( r \in R \) and \( a \in \ker f \). Then \( f(ra) = rf(a) = r0 = 0 \), since \( f \) is an \( R \)-module homomorphism and \( a \in \ker f \). It follows that \( ra \in \ker f \).

(b) Since \( f \) is a homomorphism of abelian groups, \( \im f \) is an additive subgroup of \( B \). Suppose \( r \in R \) and \( b \in \im f \). Then there exists \( a \in A \) such that \( f(a) = b \) and \( f(ra) = rf(a) = rb \). It follows that \( rb \in \im f \). \( \square \)

(3) If \( R \) is a ring, \( f : A \to B \) is a left \( R \)-module homomorphism, and \( C \) is any left submodule of \( B \), then \( f^{-1}(C) = \{ a \in A | f(a) \in C \} \) is a left submodule of \( A \).

**Proof.** Since \( f \) is a homomorphism of abelian groups, \( f^{-1}(C) \) is an additive subgroup of \( A \). Suppose \( r \in R \) and \( a \in f^{-1}(C) \). Then \( f(a) \in C \) and \( f(ra) = rf(a) \in C \) because \( C \) is a submodule of \( B \). So, \( ra \in f^{-1}(C) \). \( \square \)

(4) Let \( R \) be a ring and \( A \) a left \( R \)-module. Let \( a \) be in \( A \). Then \( Ra = \{ ra | r \in R \} \) is a submodule of \( A \).

**Proof.** (a) We first show that \( Ra \) is an additive subgroup of \( A \).

(i) Let \( x, y \in Ra \). Then \( x = r_1a \) and \( y = r_2a \) for \( r_1, r_2 \in R \), and \( x + y = r_1a + r_2a = (r_1 + r_2)a \in Ra \).
(ii) $0_{Ra} = 0_A \in Ra$.

(iii) Let $x \in Ra$. Then $x = ra$ for some $r \in R$. By Theorem 2.1.6,
$$-x = -(ra) = (-r)a \in Ra.$$  

(b) Let $r \in R$ and $b \in Ra$. Then $b = r'a$ for some $r' \in R$. By property (3) of a left $R$-module $A$, it follows that $rb = r(r'a) = (rr')a \in Ra$. □

Note that if $A$ is an $R$-$S$ bimodule, then $RaS = \{ras|r \in R, s \in S\}$ is not necessarily a submodule of $A$, as it may not be an additive subgroup of $A$. Specifically, if we let $x, y \in RaS$, then $x = r_1as_1$ and $y = r_2as_2$ for $r_1, r_2 \in R$, and $s_1, s_2 \in S$. So, $x + y = r_1as_1 + r_2as_2$, which is not necessarily in $RaS$.

(5) If $\{B_i|i \in I\}$ is a family of submodules of a left $R$-module $A$, then $B = \bigcap_{i \in I} B_i$ is a submodule of $A$.

Proof. Since each $B_i$ is a submodule of $A$, it follows that each $B_i$ is an additive subgroup of $A$. The arbitrary intersection of subgroups is itself a subgroup; therefore $B$ is an additive subgroup of $A$. Suppose $r \in R$ and $b \in B$. Then $b \in B_i$ for all $i$. So, $rb \in B_i$ for all $i$ (because each $B_i$ is a submodule). It follows that $rb \in B$. Thus, $B$ is a left submodule of $A$. □

**Definition 2.3.3.** If $X$ is a subset of a left $R$-module $A$, then the intersection of all submodules of $A$ containing $X$ is called the **submodule generated by** $X$ and is denoted $\langle X \rangle$ or $\langle x_i|i \in I \rangle$ for $x_i \in X$. Note that $\langle \emptyset \rangle = \{0\}$ (the zero submodule). If $X$ is finite, and $X$ generates the module $B$, then $B$ is said to be **finitely generated**. If $X$ consists of a single element, e.g. if $X = \{a\}$, then the submodule generated by $X$ is called the **cyclic (sub)module** generated by $a$ (see example (4) above). If
\{B_i|i \in I\} is a family of submodules of a module \(A\), then the submodule generated by \(X = \bigcup_{i \in I} B_i\) is called the sum of the modules \(B_i\) and is denoted \(B_1 + B_2 + \ldots\).

**Definition 2.3.4.** Let \(A\) be a left \(R\)-module and \(X\) an arbitrary subset of \(A\). A left \(R\)-linear combination of elements of \(X\) is defined to be a sum of the form
\[
r_1 x_{i_1} + r_2 x_{i_2} + \cdots + r_k x_{i_k}, r_t \in R, x_{i_t} \in X, 1 \leq t \leq k, k \in \mathbb{N} \text{ (where } i_t \neq i_s \text{ for } t \neq s).\]
(Note that, by definition, left \(R\)-linear combinations are finite even if \(X\) is infinite.)

The span of \(X\) is the set of all left \(R\)-linear combinations of elements in \(X\). (Similar definitions hold, mutatis mutandis, for right \(R\)-modules).

**Theorem 2.3.5.** Let \(R\) be a ring with identity, \(A\) a left \(R\)-module, and \(x_i \in A\) for each \(i \in I\). Then \(\langle x_i|i \in I \rangle = \text{span}\{x_i|i \in I\} \).

**Proof.**

(1) Suppose \(y \in \text{span}\{x_i|i \in I\}\). Then \(y = a_1 x_{i_1} + a_2 x_{i_2} + \cdots + a_k x_{i_k}\), for some \(k \in \mathbb{N}, i_j \in I, a_j \in R\). Let \(B\) be a submodule of \(A\) such that \(\{x_i|i \in I\} \subseteq B\). Then \(a_1 x_{i_1}, a_2 x_{i_2}, \ldots, a_k x_{i_k} \in B\) for \(a_j \in R\). Since \(B\) is an additive subgroup, it follows that \(y = a_1 x_{i_1} + a_2 x_{i_2} + \cdots + a_k x_{i_k} \in B\).

(2) It remains to show that \(\text{span}\{x_i|i \in I\}\) is a submodule of \(A\).

(a) Let \(y, z \in \text{span}\{x_i|i \in I\}\). Then \(y\) and \(z\) can be expressed as left \(R\)-linear combinations of elements of \(\{x_i|i \in I\}\). Let \(\{i_1, i_2, \ldots, i_\ell\}\) represent the union of the indices appearing in \(y\) and \(z\). Then \(y = a_1 x_{i_1} + a_2 x_{i_2} + \cdots + a_\ell x_{i_\ell}\) and \(z = b_1 x_{i_1} + b_2 x_{i_2} + \cdots + b_\ell x_{i_\ell}\), for some \(a_j, b_j \in R\). Since \(x_i \in A, a_j, b_j \in R, \) and \(A\) is an \(R\)-module, it follows that \(y + z = (a_1 + b_1)x_{i_1} + (a_2 + b_2)x_{i_2} + \cdots + (a_\ell + b_\ell)x_{i_\ell}\). Thus, \(y + z \in \text{span}\{x_i|i \in I\}\).

(b) We see that \(0_A = 0_R x_{i_1} \in \text{span}\{x_i|i \in I\}\).

(c) Suppose \(y \in \text{span}\{x_i|i \in I\}\). Then \(y = a_1 x_{i_1} + a_2 x_{i_2} + \cdots + a_k x_{i_k}, k \in \mathbb{N}\),
\[ a_j \in R. \text{ It follows that } -y = (-1)y = (-1)[a_1x_{i_1} + a_2x_{i_2} + \cdots + a_kx_{i_k}] = -a_1x_{i_1} - a_2x_{i_2} - \cdots - a_kx_{i_k} \in \text{span}\{x_i| i \in I\}. \]

(d) Finally, suppose \( r \in R \) and \( y \in \text{span}\{x_i|i \in I\} \). Then \( y = a_1x_{i_1} + a_2x_{i_2} + \cdots + a_kx_{i_k}, \) for some \( k \in \mathbb{N}, i_j \in I, a_j \in R. \) It follows that

\[
ry = r[a_1x_{i_1} + a_2x_{i_2} + \cdots + a_kx_{i_k}]
= r(a_1x_{i_1}) + r(a_2x_{i_2}) + \cdots + r(a_kx_{i_k}) \tag{2.17}
= (ra_1)x_{i_1} + (ra_2)x_{i_2} + \cdots + (ra_k)x_{i_k},
\]

an element of \( \text{span}\{x_i|i \in I\}. \)

Note: Theorem 2.3.5 holds for right \( R \)-modules, but may not necessarily hold for an \( R-S \) bimodule \( A, \) since for \( a \in A, RaS \) may not be a submodule of \( A. \) (See examples of submodules, note after item (4).)

**Theorem 2.3.6.** Let \( B \) be a submodule of a left \( R \)-module \( A. \) Then the **quotient module** \( A/B \) is a left \( R \)-module with elements \( a + B \in A/B, (a \in A), \) and the action of \( R \) on \( A/B \) is given by \( r(a + B) = ra + B \) for all \( a \in A \) and \( r \in R. \)

**Proof.** Since \( B \) is a submodule of \( A, A/B \) is an additive abelian group with elements \( a + B, (a \in A). \) Let \( a, a' \in A \) and \( r \in R. \) Suppose \( a + B = a' + B. \) Then \( a - a' \in B. \) Since \( B \) is a submodule of \( A, r(a-a') = ra-ra' \in B. \) It follows that \( ra+ B = ra'+ B. \)

Thus, the operation \( r(a + B) = ra + B \) for all \( a \in A \) and \( r \in R \) is well-defined.

It remains to show that \( A/B \) satisfies the four properties of a left \( R \) module:

1. Suppose \( a_1 + B \) and \( a_2 + B \in A/B \) and \( r \in R. \) Then by definition of \( A/B, \)
\[(a_1 + B) + (a_2 + B) = (a_1 + a_2) + B.\] It follows that:

\[
r[(a_1 + B) + (a_2 + B)] = r[(a_1 + a_2) + B] = r(a_1 + a_2) + B = (ra_1 + ra_2) + B = (ra_1 + B) + (ra_2 + B) = r(a_1 + B) + r(a_2 + B).
\]

(2) Let \(r, s \in R\) and \(a + B \in A/B\). Then

\[
(r + s)(a + B) = (r + s)a + B = (ra + sa) + B = (ra + B) + (sa + B) = r(a + B) + s(a + B).
\]

(3) Let \(r, s \in R\) and \(a + B \in A/B\). Then \(r[s(a + B)] = r(sa + B) = r(sa) + B = (rs)a + B = (rs)(a + B)\).

(4) Let 1\(_R\) be the identity element of \(R\). Then 1\(_R\)(a + B) = 1\(_R\)a + B = a + B. □

**Theorem 2.3.7.** (The First Isomorphism Theorem for Modules). Let \(A\) and \(B\) be \(R\)-modules, \(f : A \to B\) a left \(R\)-module homomorphism, and \(K = \text{Ker} f\). Then \(A/K \cong f(A)\).

**Proof.** Define \( \bar{f} : A/K \to f(A) \) by \( \bar{f}(a + K) = f(a) \) for all \( a + K \in A/K \).

The function \( \bar{f} \) is well-defined because if \( a + K = b + K \), then \( a = b + k \) for
some $k \in K$, and
\[
\tilde{f}(a + K) = f(a)
\]
\[
= f(b + k)
\]
\[
= f(b) + f(k)
\]
\[
= f(b) + 0
\]
\[
= f(b)
\]
\[
= \tilde{f}(b + K).
\]
Furthermore, $\tilde{f}$ is a left $R$-module homomorphism because
\[
\tilde{f}[(a + K) + (b + K)] = \tilde{f}[(a + b) + K]
\]
\[
= f(a + b)
\]
\[
= f(a) + f(b)
\]
\[
= \tilde{f}(a + K) + \tilde{f}(b + K)
\]
and
\[
\tilde{f}[r(a + K)] = \tilde{f}(ra + K)
\]
\[
= f(ra)
\]
\[
= rf(a)
\]
\[
= r\tilde{f}(a + K).
\]
Additionally, $\tilde{f} : A/K \to f(A)$ given by $\tilde{f}(a + K) = f(a)$ is clearly surjective. Finally, suppose $\tilde{f}(a + K) = \tilde{f}(b + K)$. Then $f(a) = f(b)$, so, $f(a) - f(b) = f(a - b) = 0$. Hence $a - b \in K$ which implies that $a + K = b + K$. Thus, $\tilde{f}$ is injective.

**Theorem 2.3.8.** (The Second Isomorphism Theorem for Modules). Let $B$ and $C$ be submodules of a left $R$-module $A$. Then $B/(B \cap C) \cong (B + C)/C$.

**Proof.** First note that $C$ is a submodule of $B + C$, hence $(B + C)/C$ is a left $R$-module.

Define $f : B \to (B + C)/C$ by $f(b) = b + C$. 

First, we show that \( f \) is a left \( R \)-module homomorphism. Suppose \( b_1, b_2 \in B \) and \( r \in R \). Then

\[
f(b_1 + b_2) = (b_1 + b_2) + C
\]

\[
= (b_1 + C) + (b_2 + C)
\]

\[
= f(b_1) + f(b_2).
\]

Additionally,

\[
f(rb_1) = rb_1 + C
\]

\[
= r(b_1 + C)
\]

\[
= rf(b_1).
\]

We see that \( \ker(f) = B \cap C \), since

\[
x \in \ker(f) \iff x \in B \text{ and } f(x) = C
\]

\[
\iff x \in B \text{ and } x + C = C
\]

\[
\iff x \in B \text{ and } x \in C
\]

\[
\iff x \in B \cap C.
\]

Finally, we show that \( f \) is surjective. Suppose \( x \in (B + C)/C \). Then for some \( b \in B \), \( c \in C \), it follows that

\[
x = (b + c) + C
\]

\[
= b + C
\]

\[
= f(b).
\]

By Theorem 2.3.7, it follows that \( B/(B \cap C) \cong (B + C)/C \).

**Theorem 2.3.9.** (The Third Isomorphism Theorem for Modules). Let \( B \) and \( C \) be submodules of a left \( R \)-module \( A \). If \( C \subseteq B \), then \( B/C \) is a submodule of \( A/C \) and \( (A/C)/(B/C) \cong A/B \).
Proof. Suppose \( C \subseteq B \) where \( B \) and \( C \) are submodules of a left \( R \)-module \( A \).

We first show that \( B/C \) is a submodule of \( A/C \).

(1) By the third isomorphism theorem for groups, we know that \( B/C \) is an additive subgroup of \( A/C \).

(2) Let \( r \in R \) and \( b + C \in B/C \) (where \( b \in B \)). Then since \( B \) is a submodule of \( A \), it follows that \( rb \in B \). Therefore \( r(b + C) = rb + C \in B/C \).

Define \( f : A/C \rightarrow A/B \) by \( f(a + C) = a + B \), where \( a \in A \).

(1) We claim that \( f \) is a well-defined function. Suppose \( a_1, a_2 \in A \) and \( a_1 + C = a_2 + C \). Then \( a_1 = a_2 + c \) for some \( c \in C \). Furthermore, \( a_1 + B = (a_2 + c) + B = a_2 + B \) (since \( c \in C \subseteq B \)). It follows that \( f(a_1 + C) = f(a_2 + C) \).

(2) We further claim that \( f \) is a left \( R \)-module homomorphism. Suppose \( a_1, a_2 \in A \). Then

\[
\begin{align*}
[f((a_1 + C) + (a_2 + C))] &= f[(a_1 + a_2) + C] \\
&= (a_1 + a_2) + B \\
&= (a_1 + B) + (a_2 + B) \\
&= f(a_1 + C) + f(a_2 + C). \\
\end{align*}
\]

Additionally,

\[
\begin{align*}
f(ra_1 + C) &= ra_1 + B \\
&= r(a_1 + B) \\
&= rf(a_1 + C). \\
\end{align*}
\]

(3) It also follows that \( \text{Ker}(f) = \{a + C | a \in B\} = B/C \), since

\[
\begin{align*}
a + C \in \text{Ker}(f) &\iff f(a + C) = B \\
&\iff a + B = B \\
&\iff a \in B. \\
\end{align*}
\]
(4) Finally, we show that \( f \) is surjective. Suppose \( a + B \in A/B \) for some \( a \in A \).

Then \( a + C \in A/C \). It follows that \( f(a + C) = a + B \).

By Theorem 2.3.7, it follows that \( (A/C)/(B/C) \cong A/B \). \( \square \)

Note: Theorems 2.3.7, 2.3.8, and 2.3.9 also hold, mutatis mutandis, for right \( R \)-modules and \( R-S \) bimodules.

### 2.4 Direct Product and Direct Sum

**Definition 2.4.1.** Let \( \{A_i \mid i \in I\} \) be a family of sets. The **Cartesian product** of the sets \( A_i \) is the set of all functions \( a : I \to \bigcup_{i \in I} A_i \) such that \( a(i) = a_i \in A_i \) for every \( i \in I \). The Cartesian product is denoted by \( \prod_{i \in I} A_i \) and elements of \( \prod_{i \in I} A_i \) are denoted by \( (a_i) \).

**Definition 2.4.2.** Let \( R \) be a ring and \( \{A_i \mid i \in I\} \) a family of left \( R \)-modules. The **direct product** of the family of left \( R \)-modules \( \{A_i \mid i \in I\} \) (denoted \( \prod_{i \in I} A_i \)) is the Cartesian product of the sets \( A_i \) together with the operation given by \( (a_i) + (b_i) = (a_i + b_i) \) for \( (a_i), (b_i) \in \prod_{i \in I} A_i \), and the action of \( R \) on \( \prod_{i \in I} A_i \) given by \( r(a_i) = (ra_i) \).

The **support** of \( (a_i) \) (denoted \( \text{supp}(a_i) \)) is the set of all \( i \in I \) such that \( a_i \neq 0 \). The **direct sum** of the family of left \( R \)-modules \( \{A_i \mid i \in I\} \) denoted by \( \bigoplus_{i \in I} A_i \), is the subset of the direct product consisting of all elements of finite support.

**Theorem 2.4.3.** Let \( R \) be a ring with identity \( 1_R \), and let \( \{A_i \mid i \in I\} \) be a family of left \( R \)-modules. Then

1. the direct product, \( \prod_{i \in I} A_i \), is a left \( R \)-module, and

2. the direct sum, \( \bigoplus_{i \in I} A_i \), is a left submodule of \( \prod_{i \in I} A_i \).
Proof. (1) Since $A_i$ is an additive abelian group for each $i \in I$, $\prod_{i \in I} A_i$ is an additive abelian group [Hungerford p. 59, Thm 8.1]. Additionally, if $(a_i), (b_i) \in \prod_{i \in I} A_i$ and $r, s \in R$, then:

(a) $r[(a_i) + (b_i)] = r(a_i + b_i) = (ra_i + rb_i) = (ra_i) + (rb_i) = r(a_i) + r(b_i)$.

(b) $(r + s)(a_i) = ((r + s)a_i) = (ra_i + sa_i) = (ra_i) + (sa_i) = r(a_i) + s(a_i)$.

(c) $r[s(a_i)] = r(sa_i) = ((rs)a_i) = (rs)(a_i)$.

(d) $1_R(a_i) = (1_R a_i) = (a_i)$. 

(2) (a) Claim: $\bigoplus_{i \in I} A_i$ is a subgroup of $\prod_{i \in I} A_i$.

Suppose $(a_i), (b_i) \in \bigoplus_{i \in I} A_i$. Then $\text{supp}(a_i)$ and $\text{supp}(b_i)$ are both finite.

Therefore, $\text{supp}(a_i) \cup \text{supp}(b_i)$ is finite. Observe that $\text{supp}(a_i + b_i) \subseteq \text{supp}(a_i) \cup \text{supp}(b_i)$. It follows that $\text{supp}(a_i + b_i)$ is finite and therefore $(a_i) + (b_i) = (a_i + b_i) \in \bigoplus_{i \in I} A_i$. By definition of direct sum, the identity element, $(0)$, of $\prod_{i \in I} A_i$ is also the identity element of $\bigoplus_{i \in I} A_i$. Finally, if $(a_i) \in \bigoplus_{i \in I} A_i$, then $-(a_i) = (-a_i) \in \bigoplus_{i \in I} A_i$, since $\text{supp}(-a_i) = \text{supp}(a_i)$.

(b) Suppose $r \in R$ and $(a_i) \in \bigoplus_{i \in I} A_i$. Then by definition, the support of $(a_i)$ is finite. Also, $r(a_i) = (ra_i)$ has finite support because if $a_i = 0$, then $ra_i = r0 = 0$. Thus, $r(a_i) \in \bigoplus_{i \in I} A_i$. \(\square\)

Definition 2.4.4. Let $R$ be a ring and let $\{B_i|i \in I\}$ be a family of submodules of a left $R$-module $A$. Then the sum of the family $\{B_i|i \in I\}$ (denoted $\sum_{i \in I} B_i$) is the set consisting of all finite sums of the form $b_{i_1} + b_{i_2} + \cdots + b_{i_n}$, where $i_k \in I$, $b_{i_k} \in B_{i_k}$, $n \in \mathbb{N}$.
Theorem 2.4.5. Let \( R \) be a ring, and let \( \{ B_i \mid i \in I \} \) be a family of submodules of a left \( R \)-module \( A \). Then \( \bigcup_{i \in I} B_i = \sum_{i \in I} B_i \).

Proof. Suppose \( C \) is a submodule of a left \( R \)-module \( A \) such that \( B_i \subseteq C \) for all \( i \in I \). Then since \( C \) is a submodule, it follows that \( \sum_{i \in I} B_i \subseteq C \). It remains to show that \( \sum_{i \in I} B_i \) is a submodule of \( A \). Suppose \( x, y \in \sum_{i \in I} B_i \) and \( r \in R \).

1. Let \( J \subseteq I \) be the set whose elements are all of the \( i_t \) which are used to describe \( x \) and \( y \). Then \( x = a_{i_t} + a_{i_2} + \cdots + a_{i_n} \) and \( y = b_{i_t} + b_{i_2} + \cdots + b_{i_m} \), for some \( n \in \mathbb{N} \), \( i_t \in J \), \( a_{i_t}, b_{i_t} \in B_{i_t} \). It follows that \( x + y = (a_{i_t} + b_{i_t}) + \cdots + (a_{i_n} + b_{i_m}) \).
   Since each \( (a_{i_t} + b_{i_t}) \in B_{i_t} \), it follows that \( x + y \in \sum_{i \in I} B_i \).

2. Since \( 0 \in B_i \) for all \( i \in I \), it follows that \( 0 \in \sum_{i \in I} B_i \).

3. For \( x = a_{i_t} + a_{i_2} + \cdots + a_{i_n} \), \( a_{i_t} \in B_{i_t} \), we have \( -x = -(a_{i_t} + a_{i_2} + \cdots + a_{i_n}) = -a_{i_t} - a_{i_2} - \cdots - a_{i_n} \in \sum_{i \in I} B_i \).

4. Finally, for \( x = a_{i_t} + a_{i_2} + \cdots + a_{i_n} \), \( a_{i_t} \in B_{i_t} \), we have \( rx = r(a_{i_t} + a_{i_2} + \cdots + a_{i_n}) = ra_{i_t} + ra_{i_2} + \cdots + ra_{i_n} \in \sum_{i \in I} B_i \).

The result follows. \( \square \)

Definition 2.4.6. Let \( R \) be a ring and let \( \{ A_i \mid i \in I \} \) be a family of left submodules of an \( R \)-module \( A \) such that:

1. \( A = \sum_{i \in I} A_i \); and

2. For every \( k \in I \), \( A_k \cap (\sum_{i \neq k} A_i) = 0 \).

Then \( A \) is said to be the internal direct sum of the family of left submodules \( \{ A_i \mid i \in I \} \).
Lemma 2.4.7. Let $R$ be a ring and let a left $R$-module $A$ be the internal direct sum of the family of left submodules $\{A_i|i \in I\}$. Then if $a_{i_1} + a_{i_2} + \cdots + a_{i_k} = 0$, where $a_{i_k} \in A_{i_k}$ and $i_k \in I$, it follows that $a_{i_t} = 0$ for all $1 \leq t \leq k$.

Proof. Suppose $a_{i_1} + a_{i_2} + \cdots + a_{i_k} = 0$, where $k \in \mathbb{N}$, $a_{i_k} \in A_{i_k}$ and $i_k \in I$. Let $1 \leq t \leq k$. Then $a_{i_t} = -(a_{i_1} + \cdots + a_{i_{t-1}} + a_{i_{t+1}} + a_{i_k}) \in A_{i_t} \cap \left( \sum_{i_t \neq i_k} A_i \right)$. By Definition 2.4.6 (2), it follows that $a_{i_t} = 0$. \hfill \Box

Theorem 2.4.8. Let $R$ be a ring and $\{A_i|i \in I\}$ a family of submodules of a left $R$-module $A$ such that $A$ is the internal direct sum of $\{A_i|i \in I\}$. Then $\bigoplus A_i \cong A$.

Proof. Define $f : \bigoplus_{i \in I} A_i \to A$ by $f((a_i)) = a_{i_1} + a_{i_2} + \cdots + a_{i_k}$ where $a_{i_t} \in A_{i_t}$, and $i_t \in \text{supp}(a_i)$. Suppose $a = (a_i), b = (b_i) \in \bigoplus_{i \in I} A_i$ and $r \in R$, where $\{i_1, \ldots, i_k\} = \text{supp}(a) \cup \text{supp}(b)$. Then

\begin{equation}
(1) \quad f(ra) = f(r(a_i)) = f((ra_i)) = ra_{i_1} + ra_{i_2} + \cdots + ra_{i_k} = r(a_{i_1} + a_{i_2} + \cdots + a_{i_k}) = rf((a_i)), \tag{2.30}
\end{equation}

and

\begin{equation}
f(a + b) = (a_{i_1} + b_{i_1}) + \cdots + (a_{i_k} + b_{i_k}), i_t \in \text{supp}(a + b)
\begin{align*}
&= (a_{i_1} + \cdots + a_{i_k}) + (b_{i_1} + \cdots + b_{i_k}) \\
&= f(a) + f(b). \tag{2.31}
\end{align*}
\end{equation}

It follows that $f$ is a left $R$-module homomorphism.
(2) Suppose \((a_i) \in \text{Ker } f\). Suppose further that \(\text{supp}(a_i) \neq \emptyset\). Let \(\{i_1, \ldots, i_k\} = \text{supp}(a_i)\). Then \(f((a_i)) = 0\) implies \(a_{i_1} + a_{i_2} + \cdots + a_{i_k} = 0\), where \(i_t \in \text{supp}(a_i)\).

By Lemma 2.4.7, it follows that \(a_{i_t} = 0\) for all \(1 \leq t \leq k\). Thus, we must have that \(\text{supp}(a_i) = \emptyset\), and therefore, \((a_i) = (0)\). It follows that \(f\) is injective.

(3) Suppose \(y \in A\). Then \(y = a_{i_1} + a_{i_2} + \cdots + a_{i_k}, i_k \in I,\) and \(|k| < \infty\). Define \(a = (a_i)\) by

\[
a_{i_t} = \begin{cases} 
a_{i_m} & \text{if } i = i_m \in \{i_1, \ldots, i_k\} \\
0 & \text{otherwise}
\end{cases}
\]

Then \((a_i) \in \bigoplus_{i \in I} A_i\) and \(f((a)) = a_{i_1} + a_{i_2} + \cdots + a_{i_k} = y\). It follows that \(f\) is surjective.

\[\square\]

2.5 Free Modules

The following definitions build from Definition 2.3.4.

**Definition 2.5.1.** Let \(R\) be a ring, and let \(X\) be a subset of a left \(R\)-module \(A\). The set \(X\) is said to be **linearly independent** over \(R\) if whenever an \(R\)-linear combination of distinct elements of \(X\) is set equal to zero, all of its coefficients must be equal to zero. In other words, to say \(X\) is linearly independent means that if \(r_1x_{i_1} + r_2x_{i_2} + \cdots + r_kx_{i_k} = 0, r_t \in R, x_{i_t} \in X,\) and \(k \in \mathbb{N},\) then \(r_1 = r_2 = \cdots = r_k = 0\).

**Definition 2.5.2.** Let \(R\) be a ring, and let \(X\) be a subset of a left \(R\)-module \(A\). Then \(X\) is a **basis** for \(A\) if \(X\) generates \(A\) and \(X\) is linearly independent.

**Theorem 2.5.3.** Let \(R\) be a ring, and let \(A\) be a left \(R\)-module with basis \(X \subseteq A\). Then every element of \(A\) can be expressed uniquely as a left \(R\)-linear combination of the elements of \(X\). (Specifically, the non-zero coefficients of the left \(R\)-linear combination will be unique.)
Proof. Suppose \( y \in A \) such that \( y = r_1u_1 + \cdots + r_nu_n = s_1v_1 + \cdots + s_mv_m \), where \( r_i, s_j \in R \) and \( u_i, v_j \in X \). Let \( \{w_1, w_2, \ldots, w_k\} = \{u_i|1 \leq i \leq n\} \cup \{v_j|1 \leq j \leq m\} \). Then \( y = a_1w_1 + \cdots + a_kw_k = b_1w_1 + \cdots + b_kw_k \) for some \( a_i, b_i \in R \), and \( y - y = 0 = (a_1 - b_1)w_1 + \cdots + (a_k - b_k)w_k \). Now since the \( w_i \in X \), they are linearly independent, and it follows that \( (a_1 - b_1) = \cdots = (a_k - b_k) = 0 \), which implies that \( a_i = b_i \) for all \( i \). The result follows.

**Theorem 2.5.4.** Let \( R \) be a ring and let \( X \) be a set. The following conditions on a left \( R \)-module \( F \) are equivalent:

1. \( F \) is left \( R \)-module isomorphic to a direct sum of copies of the left \( R \)-module \( R \), indexed by \( X \).
2. \( F \) has a non-empty basis indexed by \( X \).
3. There exists a function \( \iota : X \to F \) with the following property: given any left \( R \)-module \( A \) and function \( f : X \to A \), there exists a unique left \( R \)-module homomorphism \( \bar{f} : F \to A \) such that \( \bar{f}\iota = f \).

Note that item (3) in Theorem 2.5.4, essentially says that, if \( F \) and \( A \) are two left \( R \)-modules such that \( F \) is a free left \( R \)-module, then there is a unique left \( R \)-module homomorphism, \( \bar{f} : F \to A \), that will be completely determined by where the generators of \( F \) are sent in \( A \).

**Proof.** (1) \( \Rightarrow \) (2): Let \( F = \bigoplus_{x \in X} R \). Consider the set \( B = \{e_x|x \in X\} \subseteq F \), where \( e_x(x) = 1, e_x(y) = 0 \) for \( x \neq y \). We claim that \( B \) is a basis for \( F \). Suppose

\[
 r_1e_{x_1} + \cdots + r_ne_{x_n} = 0, \tag{2.32}
\]

where \( r_i \in R \) and \( e_{x_i} \in B \). Then \( r_1e_{x_1}(x_1) + \cdots + r_ne_{x_n}(x_i) = 0(x_i) \). It follows that \( r_i = 0(x_i) = 0 \) and thus \( B \) is a linearly independent set.
Suppose further that \( a = (a_x) \in F \). Let \( \{x_1, x_2, \ldots, x_n\} = \text{supp}(a) \). We claim that \( a = a_{x_1}e_{x_1} + \cdots + a_{x_n}e_{x_n} \). There are two cases to consider.

If \( x_i \in \text{supp}(a) \), then \( a_{x_1}e_{x_1}(x_i) + \cdots + a_{x_n}e_{x_n}(x_i) = a_{x_i} = a(x_i) \). If \( x_i \not\in \text{supp}(a) \), then \( a_{x_1}e_{x_1}(x_i) + \cdots + a_{x_n}e_{x_n}(x_i) = 0 = a(x_i) \). The result follows.

(2) \( \Rightarrow \) (3): Suppose \( \{u_x | x \in X\} \) is a basis for \( F \). Define \( \iota : X \to F \) by \( \iota(x) = u_x \). Suppose \( A \) is a left \( R \)-module and \( \iota : X \to A \) is a function. Suppose further that \( \bar{f}, \bar{g} : F \to A \) are left \( R \)-module homomorphisms and \( \bar{f} \iota = f, \bar{g} \iota = f \). For \( y \in F \), \( y = r_1u_{x_1} + \cdots + r_nu_{x_n} \) for some \( r_i \in R, x_i \in X \). So,

\[
\bar{f}(y) = \bar{f}(r_1u_{x_1} + \cdots + r_nu_{x_n}) \\
= \bar{f}(r_1u_{x_1}) + \cdots + \bar{f}(r_nu_{x_n}) \\
= r_1\bar{f}(u_{x_1}) + \cdots + r_n\bar{f}(u_{x_n}) \\
= r_1f(x_1) + \cdots + r_nf(x_n) \\
= r_1\bar{g}(u_{x_1}) + \cdots + r_n\bar{g}(u_{x_n}) \\
= \bar{g}(r_1u_{x_1}) + \cdots + \bar{g}(r_nu_{x_n}) \\
= \bar{g}(y).
\]

(2.33)

It follows that \( \bar{f} \) is unique.

As for existence, define \( \bar{f} : F \to A \) as \( \bar{f}(y) = r_1f(x_1) + \cdots + r_nf(x_n) \) for \( y = r_1u_{x_1} + \cdots + r_nu_{x_n} \in F \). By Theorem 2.5.3, \( \bar{f} \) is well-defined.

(1) Let \( x \in X \). Then \( \bar{f}(\iota(x)) = \bar{f}(u_x) = f(x) \). It follows that \( \bar{f} \iota = f \) for all \( x \in X \).
(2) (a) Let \( r \in R, y = r_1u_{x_1} + \cdots + r_nu_{x_n} \in F \). Then

\[
\tilde{f}(ry) = \tilde{f}(r_1u_{x_1} + \cdots + r_nu_{x_n}) = r_1f(x_1) + \cdots + r_nf(x_n) = r(r_1f(x_1) + \cdots + r_nf(x_n)) = r\tilde{f}(y). \tag{2.34}
\]

(b) Suppose \( y, z \in F \) such that \( y = r_1u_{x_1} + \cdots + r_nu_{x_n}, z = s_1u_{x_1} + \cdots + s_nu_{x_n} \), where \( \{x_i | i \in I\} \) is the union of the indices of the basis elements in which \( y \) and \( z \) are expressed. Then

\[
\tilde{f}(y + z) = (r_1 + s_1)f(x_1) + \cdots + (r_n + s_n)f(x_n) = [r_1f(x_1) + \cdots + r_nf(x_n)] + [s_1f(x_1) + \cdots + s_nf(x_n)] = \tilde{f}(y) + \tilde{f}(z). \tag{2.35}
\]

(3) \( \Rightarrow \) (1) Define \( A = \bigoplus_{x \in X} R \) and define \( f : X \to A \) by \( f(x) = e_x \) for all \( x \in X \) (where \( (e_x) \in A \) such that \( e_x(x) = 1 \) and \( e_x(y) = 0 \) for \( y \neq x \)). By (3), we know that there exists a unique left \( R \)-module homomorphism \( \tilde{f} : F \to A \) such that \( \tilde{f} \iota = f \); that is, \( \tilde{f}(\iota(x)) = e_x \) for all \( x \in X \). We claim that \( \tilde{f} \) is a left \( R \)-module isomorphism.

Let \( a \in A \). Then \( a = a_{x_1}e_{x_1} + \cdots + a_{x_n}e_{x_n} \), where \( \{x_i | 1 \leq i \leq n\} = \text{supp}(a) \). Define \( \bar{g} : A \to F \) by

\[
\bar{g}(a) = a_{x_1}\iota(x_1) + \cdots + a_{x_n}\iota(x_n). \tag{2.36}
\]

Then \( \bar{g} \) is well-defined because \( \text{supp}(a) \) is finite. Additionally, \( 2.36 \) is also holds for any set \( \{y_j\} \) that contains \( \text{supp}(a) \).

We claim that \( \bar{g} \) is a left \( R \)-module homomorphism.

Suppose \( a, b \in A \) and \( r \in R \). Let \( \{x_1, \ldots, x_n\} = \text{supp}(a) \cup \text{supp}(b) \). Then \( a = a_{x_1}e_{x_1} + \cdots + a_{x_n}e_{x_n} \) and \( b = b_{x_1}e_{x_1} + \cdots + b_{x_n}e_{x_n} \) for some \( a_{x_i}, b_{x_j} \in R \), and
\[(a + b) = (a_{x_1} + b_{x_1})e_{x_1} + \cdots + (a_{x_n} + b_{x_n})e_{x_n}.\] So,
\[
g(a + b) = (a_{x_1} + b_{x_1})\iota(x_1) + \cdots + (a_{x_n} + b_{x_n})\iota(x_n)
= (a_{x_1}\iota(x_1) + \cdots + a_{x_n}\iota(x_n)) + (b_{x_1}\iota(x_1) + \cdots + b_{x_n}\iota(x_n))
\]
\[
= g(a) + g(b)
\]
and
\[
g(ra) = ra_{x_1}\iota(x_1) + \cdots + ra_{x_n}\iota(x_n)
= r(a_{x_1}\iota(x_1) + \cdots + a_{x_n}\iota(x_n))
\]
\[
= rg(a).
\]
It remains to show that \(g\) is the inverse of \(\bar{f}\):

Let \(a \in A\) be defined as above. Then
\[
\bar{f}(g(a)) = \bar{f}(a_{x_1}\iota(x_1) + \cdots + a_{x_n}\iota(x_n))
= a_{x_1}\bar{f}(x_1) + \cdots + a_{x_n}\bar{f}(x_n)
= a_{x_1}f(x_1) + \cdots + a_{x_n}f(x_n)
\]
\[
= a_{x_1}e_{x_1} + \cdots + a_{x_n}e_{x_n}
= a,
\]
where (*) follows because \(\bar{f}\) is a left \(R\)-module homomorphism.

Let \(x \in X\). Then \(\iota(x) \in F\) and
\[
g(\bar{f}(\iota(x))) = g(f(x))
= g(e_x)
\]
\[
= \iota(x).
\]

Now, \(g \circ \bar{f} : F \to F\) and \(\text{id}_F : F \to F\) are left \(R\)-module homomorphisms such that \(g \circ \bar{f}(\iota(x)) = \text{id}_F(\iota(x))\). Then by the uniqueness part of property (3), we have that \(g \circ \bar{f} = \text{id}_F\). In other words, by property (3), since \(g \circ \bar{f}\) and \(\text{id}_F\) agree on \(\iota(x)\) for all \(x \in X\), it follows that they agree everywhere.
Thus, \( \tilde{f} \) is a left \( R \)-module isomorphism. \( \Box \)

**Definition 2.5.5.** A module \( F \) over a ring \( R \), that satisfies the equivalent conditions of the theorem above is called a **free \( R \)-module** on the set \( X \).

**Definition 2.5.6.** Let \( B \subseteq \bigoplus_{x \in X} R \) such that \( B = \{ e_x | x \in X \} \), where \( e_x(x) = 1, e_x(y) = 0 \) for \( x \neq y \). Then the elements of the set \( B \) are called the **standard basis vectors** of \( \bigoplus_{x \in X} R \).

### 2.6 Modules over Division Rings

**Theorem 2.6.1.** Let \( D \) be a division ring, and let \( A \) be the left \( D \)-module generated by \( \{x_1, x_2, \ldots, x_n\} \), where \( x_i \in A \). Then a subset of \( \{x_1, x_2, \ldots, x_n\} \) is a basis of \( A \). In particular, every finitely generated left \( D \)-module is free.

**Proof.** We proceed by induction on \( n \). The base case \( n = 0 \) is true because the \( D \)-module \( \{0\} \) is generated by the empty set (that is, \( \{0\} = \langle \emptyset \rangle \)), and since \( \emptyset \) is a (vacuously) linearly independent set by definition, it follows that \( \emptyset \) is a basis for \( \{0\} \).

Suppose that the theorem is true for \( k \) and that \( A = \langle x_1, x_2, \ldots, x_{k+1} \rangle \). If \( \{x_1, x_2, \ldots, x_{k+1}\} \) is a linearly independent set, then we are done. Otherwise, there exists a left \( D \)-linear combination \( a_1x_1 + a_2x_2 + \cdots + a_{k+1}x_{k+1} = 0 \) such that not all \( a_i = 0 \). Without loss of generality, suppose \( a_{k+1} \neq 0 \). Then

\[
x_{k+1} = -a_{k+1}^{-1}a_1x_1 - a_{k+1}^{-1}a_2x_2 - \cdots - a_{k+1}^{-1}a_kx_k
\]

\[
= d_1x_1 + d_2x_2 + \cdots + d_kx_k,
\]

where \( d_i = -a_{k+1}^{-1}a_i \in D \).

We now claim that \( A = \langle x_1, x_2, \ldots, x_{k+1} \rangle = \langle x_1, x_2, \ldots, x_k \rangle \). Suppose \( y \in \langle x_1, x_2, \ldots, x_{k+1} \rangle \). Then \( y = b_1x_1 + \cdots + b_kx_k + b_{k+1}x_{k+1} \), for some \( b_i \in D \). By (2.41),
it follows that

\[ y = b_1 x_1 + \cdots + b_k x_k + b_{k+1} (d_1 x_1 + d_2 x_2 + \cdots + d_k x_k) \]

\[ = (b_1 + b_{k+1} d_1) x_1 + \cdots + (b_k + b_{k+1} d_k) x_k. \]  

(2.42)

Thus, \( y \in \langle x_1, x_2, \ldots, x_k \rangle \). The result follows since \( \langle x_1, x_2, \ldots, x_k \rangle \subseteq \langle x_1, x_2, \ldots, x_{k+1} \rangle \).

By the induction hypothesis, it follows that some subset of \( \{x_1, x_2, \ldots, x_n\} \) is a basis of \( A \). The theorem follows by induction. \( \square \)

**Theorem 2.6.2.** Let \( D \) be a division ring, let \( A \) be a finitely generated left \( D \)-module, let \( \{y_1, \ldots, y_k, y_{k+1}, \ldots, y_m\} \) be a linearly independent subset of \( A \), and suppose \( A = \langle x_1, \ldots, x_j, y_1, \ldots, y_k \rangle \), where \( 0 \leq k < m \). Then for some \( j, 1 \leq j \leq \ell \), it follows that \( A = \langle x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{\ell}, y_1, \ldots, y_k, y_{k+1} \rangle \).

**Proof.** Let \( A = \langle x_1, \ldots, x_\ell, y_1, \ldots, y_k \rangle \). Let \( Y = \{y_1, y_2, \ldots, y_m\} \subseteq A \) be a linearly independent set. Since \( Y \) is a linearly independent set, it follows that \( y_{k+1} \neq 0 \).

Since \( y_{k+1} \in A \), \( y_{k+1} = r_1 x_1 + \cdots + r_\ell x_\ell + s_1 y_1 + \cdots + s_k y_k, r_i, s_t \in D \). We must have \( r_i \neq 0 \) for some \( i \) because otherwise \( y_{k+1} \) would be expressible as a left \( D \)-linear combination of \( y_1, \ldots, y_k \), which would contradict the linear independence of \( Y \). Let \( r_j \) be the first non-zero coefficient. Then,

\[ x_j = r_j^{-1} y_{k+1} - r_j^{-1} r_1 x_1 - r_j^{-1} r_{j-1} x_{j-1} - r_j^{-1} r_{j+1} x_{j+1} - \cdots - r_j^{-1} r_\ell x_\ell - r_j^{-1} s_1 y_1 - \cdots - r_j^{-1} s_k y_k. \]

(2.43)

Let \( A' = \langle x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_\ell, y_1, \ldots, y_k, y_{k+1} \rangle \). We claim that \( A' = A \). First of all, clearly \( A' \subseteq A \). On the other hand, if \( v \in A \), then

\[ v = a_1 x_1 + \cdots + a_j x_j + \cdots + a_\ell x_\ell + b_1 y_1 + \cdots + b_k y_k. \]

(2.44)

By equation 2.43,
\[ v = a_1 x_1 + \cdots + a_j x_j + \cdots + a_\ell x_\ell + b_1 y_1 + \cdots + b_k y_k \]

\[ = a_1 x_1 + \cdots + a_j (r_j^{-1} y_{k+1} - r_j^{-1} r_1 x_1 - r_j^{-1} r_{j-1} x_{j-1} - r_j^{-1} r_{j+1} x_{j+1} - \cdots
   - r_j^{-1} r_\ell x_\ell - r_j^{-1} s_1 y_1 - \cdots - r_j^{-1} s_k y_k) + \cdots + a_\ell x_\ell + b_1 y_1 + \cdots + b_k y_k \]

\[ = (a_1 - a_j r_j^{-1} r_1) + \cdots + (a_{j-1} - a_j r_j^{-1} r_{j-1}) x_{j-1} + (a_{j+1} - a_j r_j^{-1} r_{j+1}) x_{j+1} + \cdots
   + (a_\ell - a_j r_j^{-1} r_\ell) x_\ell + (b_1 - a_j r_j^{-1} s_1) y_1 + \cdots + (b_k - a_j r_j^{-1} s_k) y_k + (a_j r_j^{-1}) y_{k+1} \]

(2.45)

It follows that \( A \subseteq A' \). \( \square \)

**Corollary 2.6.3.** Let \( D \) be a division ring and \( A \) be a finitely generated left \( D \)-module. If \( A = \langle x_1, x_2, \ldots, x_n \rangle \), and \( Y = \{y_1, y_2, \ldots, y_m\} \subseteq A \) is a linearly independent set, then \( m \leq n \).

**Proof.** Suppose \( X = \{x_1, x_2, \ldots, x_n\} \) spans \( A \). Let \( Y = \{y_1, y_2, \ldots, y_m\} \subseteq A \) be a linearly independent set. By Theorem 2.6.2, we know that

\[ \langle x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n, y_k \rangle = A, \]

(2.46)

for some \( j, k, 1 \leq j \leq n, 1 \leq k \leq m \).

By applying Theorem 2.6.2 repeatedly, at the end of the \( p \)th step, we will have a spanning set consisting of \( y_m, y_{m-1}, \ldots, y_{m-p+1} \) together with \( n - p \) of the \( x_i \). If \( m > n \), then at the end of \( n \) steps, we will have that \( \{y_m, \ldots, y_{m-n+1}\} \) spans \( A \). Since \( m - n > 0 \) it follows that \( m - n + 1 \geq 2 \), and thus, \( y_1 \) can be expressed as a left \( D \)-linear combination of \( y_m, \ldots, y_{m-n+1} \). But this contradicts the linear independence of \( Y \). Thus, we must have \( m \leq n \).

Note that the last statement is also true when \( Y = \emptyset \). \( \square \)
Corollary 2.6.4. Let $D$ be a division ring and $A$ be a finitely generated left $D$-module. Then any two bases of $A$ have the same cardinality.

Proof. Suppose $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_m\}$ are two bases of $A$. Then by Theorem 2.6.3 since $X$ spans $A$ and $Y$ is linearly independent, we must have $m \leq n$. On the other hand, since $Y$ spans $A$ and $X$ is linearly independent, we must have $n \leq m$. It follows that $m = n$. \qed

Definition 2.6.5. Let $D$ be a division ring and $F$ a free $D$-module. The number of elements in a basis for $F$ is called the dimension of $F$ over $D$.

Theorem 2.6.6. Let $D$ be a division ring. Let $X$ be a linearly independent subset of a $D$-module $A$ and let $a \in A$ such that $a \notin X$. Then $X \cup \{a\}$ is linearly dependent if and only if $a \in \text{span}(X)$.

Proof. Suppose $X \cup \{a\}$ is linearly dependent. Then there exist $w_1, w_2, \ldots, w_n \in X \cup \{a\}$ such that $b_1w_1 + b_2w_2 + \cdots + b_nw_n = 0$ and not all $b_i$ equal zero. Since $X$ is a linearly independent set, it follows that one of the $w_i$ must be equal to $a$, and its corresponding coefficient $b_i \neq 0$. Without loss of generality, let $w_1 = a$. Then $b_1a + b_2w_2 + \cdots + b_nw_n = 0$ which implies that $a = -b_1^{-1}b_2w_2 - \cdots - b_1^{-1}b_nw_n$. Since $a$ can be expressed as a left $D$-linear combination of $w_2, \ldots, w_n \in X$, it follows that $a \in \text{span}(X)$. Conversely suppose that $a \in \text{span}(X)$. Then there exist $w_1, \ldots, w_n$ and $b_1, \ldots, b_n \in D$ such that $a = b_1w_1 + \cdots + b_nw_n$. It follows that $(-1)a + b_1w_1 + \cdots + b_nw_n = 0$. Thus, $X \cup \{a\}$ is linearly dependent. \qed

Corollary 2.6.7. Let $D$ be a division ring. If $X = \{x_1, x_2, \ldots, x_n\}$ is a linearly independent subset of a finitely generated $D$-module $A$, then $X$ can be expanded to a basis of $A$. 
Proof. If span\((X) = A\), then we are done. Otherwise there exists \(x_{n+1} \in A - \text{span}(X)\). Since \(x_{n+1} \notin \text{span}(X)\), it follows by Theorem 2.6.6 that \(X \cup \{x_{n+1}\}\) is still linearly independent. If \(\text{span}(X \cup \{x_{n+1}\}) = A\), then we are done. Otherwise there exists \(x_{n+2} \in A - \text{span}(X)\). If \(\text{span}(X \cup \{x_{n+2}\}) = A\), then we are done. Since \(A\) is finitely generated, by Corollary 2.6.3, this process can only continue for a finite number of steps. \(\square\)

2.7 Group Algebras

The following definitions are adapted from Serre [Ser77, Sections 1.1 and 6.1].

Definition 2.7.1. Let \(G\) be a group and let \(K\) be a commutative ring. The **group algebra** of \(G\) over \(K\), denoted \(K[G]\), is a ring with elements \(r = \sum_{g \in G} a_g g\) where \(a_g \in K\) and only finitely many of the the \(a_g\) are non-zero. If \(r, s \in K[G]\) such that \(r = \sum_{g \in G} a_g g\) and \(s = \sum_{g \in G} b_g g\), then

\[
  r + s = \sum_{g \in G} (a_g + b_g)g \quad \text{and} \quad (2.47)
\]

\[
  rs = \sum_{g \in G} \sum_{h \in G} a_g b_h (gh). \quad (2.48)
\]

(Note: both of these definitions are well-defined because only finitely many of the \(a_g\) and \(b_h\) are non-zero). We identify \(g \in G\) with \(1_K g \in K[G]\), where \(1_K\) is the identity element of \(K\). The identity element of \(K[G]\) is denoted \(1_{K[G]}\), where \(1_{K[G]} = (1_K)(1_G)\).

Definition 2.7.2. Let \(G\) be a multiplicative group, \(K\) a commutative ring, and \(V_K\) a right \(K\)-module. A **linear representation** of \(G\) in \(V_K\) is a group homomorphism \(\lambda\) from \(G\) into \(GL(V_K)\) given by \(\lambda(g) = \lambda_g\), where

\[
  \lambda_{gh}(x) = \lambda_g \lambda_h(x) = \lambda_g(\lambda_h(x)) \quad (2.49)
\]

for all \(g, h \in G\) and \(x \in V_K\).
**Definition 2.7.3.** Let $G$ be a multiplicative group, $K$ a commutative ring, and $KV$ a left $K$-module. A **linear representation** of $G$ in $KV$ is a group homomorphism $\rho$ from $G$ into $GL(KV)$ given by $\rho(g) = \rho_g$, where

$$(x)\rho_{gh} = (x)\rho_g \rho_h = (x\rho_g)\rho_h$$

(2.50)

for all $g, h \in G$ and $x \in KV$.

**Theorem 2.7.4.** Let $G$ be a multiplicative group, $K$ a commutative ring, and $V_K$ a right $K$-module. Let $\lambda : G \to GL(V_K)$ be a linear representation. For $r = \sum_{g \in G} a_g g \in K[G]$ and $x \in V_K$, define the left action of $K[G]$ on $V_K$ by

$$rx = \left(\sum_{g \in G} a_g g\right)x = \sum_{g \in G} (\lambda_g(x))a_g = \sum_{g \in G} a_g \lambda_g(x).$$

(2.51)

(since $K$ is commutative). Then $V_K$ is a left $K[G]$-module.

Note that, given the definition of the left action of $K[G]$ on $V_K$ in Theorem 2.7.4, for $r = h \in G$ and $x \in V_K$, we have

$$hx = (1_K h)x = [\lambda_h(x)]1_K = \lambda_h(x).$$

(2.52)

**Theorem 2.7.5.** Let $G$ be a multiplicative group, $K$ a commutative ring, and $KV$ a left $K$-module. Let $\rho : G \to GL(KV)$ be a linear representation. For $r = \sum_{g \in G} a_g g \in K[G]$ and $x \in KV$, define the right action of $K[G]$ on $KV$ by

$$xr = x\left(\sum_{g \in G} a_g g\right) = \sum_{g \in G} a_g ((x)\rho_g).$$

(2.53)

Then $V_K$ is a right $K[G]$-module.

Note that, given the definition of the right action of $K[G]$ on $V_K$ in Theorem 2.7.5, for $r = h \in G$ and $x \in V_K$, we have

$$xh = x(1_K h) = 1_K [(x)\rho_h] = (x)\rho_h.$$

(2.54)

We prove Theorem 2.7.4. The proof for Theorem 2.7.5 is similar.
Proof. Let \( r, s \in K[G] \). Then \( r = \sum_{g \in G} a_g g \) and \( s = \sum_{g \in G} b_g g \), where \( a_g, b_g \in K \). Suppose \( x, y \in V \). Then

\[
    r(x + y) = \left( \sum_{g \in G} a_g g \right)(x + y)
    = \sum_{g \in G} a_g \lambda_g(x + y)
    = \sum_{g \in G} a_g \left[ \lambda_g(x) + \lambda_g(y) \right] \quad (*)
    = \sum_{g \in G} \left( a_g \lambda_g(x) + a_g \lambda_g(y) \right) \quad (**) \\
    = \sum_{g \in G} a_g \lambda_g(x) + \sum_{g \in G} a_g \lambda_g(y)
    = rx + ry,
\]

where (*) follows because \( \lambda_g \) is a right \( K \)-module homomorphism and (**) follows because \( \lambda_g(x), \lambda_g(y) \in V_K \) and \( V_K \) has the structure of a left \( K \)-module since \( K \) is commutative. (See Theorem 2.1.4.) Similarly,

\[
    (r + s)x = \left[ \sum_{g \in G} (a_g + b_g) g \right] x
    = \sum_{g \in G} (a_g + b_g) \lambda_g(x)
    = \sum_{g \in G} (a_g \lambda_g(x) + b_g \lambda_g(x)) \quad (*) \\
    = \sum_{g \in G} a_g \lambda_g(x) + \sum_{g \in G} b_g \lambda_g(x)
    = rx + sx,
\]

where (*) follows because \( \lambda_g(x) \in V_K \) and \( V_K \) has the structure of a left \( K \)-module
since $K$ is commutative. Furthermore,

$$ r(sx) = \sum_{g \in G} a_g \left[ \left( \sum_{h \in G} b_h h \right) x \right] $$

$$ = \left( \sum_{g \in G} a_g \right) \left[ \sum_{h \in G} b_h \lambda_h(x) \right] $$

$$ = \sum_{g \in G} a_g \lambda_g \left( \sum_{h \in G} b_h \lambda_h(x) \right) \quad (*) $$

$$ = \sum_{g \in G} \sum_{h \in G} a_g b_h \lambda_g (\lambda_h(x)) \quad (**) $$

$$ = \sum_{g \in G} \sum_{h \in G} a_g b_h \lambda_{gh}(x) $$

$$ = \left[ \sum_{g \in G} \sum_{h \in G} a_g b_h (gh) \right] x $$

$$ = (rs)x, $$

where $(*)$ follows because $\sum_{h \in G} b_h \lambda_h(x)$ is in $V_K$ and $(**)$ follows because $\lambda_g$ is a right $K$-module homomorphism.

Let $1_{K[G]} \in K[G]$. Since $\lambda$ is a group homomorphism, we must have that $\lambda(1_G) = \lambda_{1_G} = \text{id} \in GL(V_K)$. Then $1_{K[G]} x = (1_K)(1_G)x = (1_K)\lambda_{1_G}(x) = 1_K x = x$ because $x \in V_K$ and $V_K$ has the structure of a left $K$-module. \hfill $\square$

**Theorem 2.7.6.** Let $G$ be a group, $K$ a commutative ring, and $V$ a left $K[G]$-module. Then the map $\lambda : G \to GL(V_K)$ given by $\lambda(g) = \lambda_g$, where $\lambda_g(v) = gv$ for all $v \in V$, is a linear representation.

Note that since $V$ is a left $K[G]$-module, it is also a left $K$-module (because $K \subseteq K[G]$). Since $K$ is commutative, it follows by Theorem 2.1.4 that $V$ is also a right $K$-module.
Proof. Suppose $V$ is a left $K[G]$-module. Let $g \in G$.

(1) We claim that $\lambda_g : V \to V$ is a right $K$-module isomorphism (and therefore $\lambda_g \in GL(V_K)$).

(a) Suppose $v, w \in V$ and $\lambda_g(v) = \lambda_g(w)$. Then $gv = gw$. By multiplication on the left by $g^{-1}$, it follows that $v = w$, and therefore $\lambda_g$ is injective.

(b) Suppose $v \in V$. For $g \in G$ it follows that $g^{-1} = (1_K)g^{-1} \in K[G]$. Since $V$ is a left $K[G]$-module, it follows that $g^{-1}v \in V$. Then $\lambda_g(g^{-1}v) = g(g^{-1}v) = (gg^{-1})v = v$. It follows that $\lambda_g$ is surjective.

(c) Suppose $v, w \in V$ and $r \in K$. Then

$$\begin{align*}
\lambda_g(v + w) &= g(v + w) \\
&= gv + gw \\
&= \lambda_g(v) + \lambda_g(w),
\end{align*}$$

where (*) follows by property (1) of the left $K[G]$-module $V$, and

$$\begin{align*}
\lambda_g(rv) &= g(rv) \\
&= 1_Kg[(r1_G)v] \\
&= [(1_Kg)(r1_G)]v \\
&= [1_Kr(g1_G)]v \\
&= (rg)v \\
&= r(gv) \\
&= r\lambda_g(v),
\end{align*}$$

where (*) and (**) follow by property (3) of the left $K[G]$-module $V$ and (**) follows by definition of multiplication in the group algebra $K[G]$. 
(2) We claim that \( \lambda : G \to GL(V) \) is a linear representation. Suppose \( g, h \in G \).

Then \( \lambda(gh) = \lambda_{gh} \). Let \( v \in V \). Then \( \lambda_{gh}(v) = (gh)v = g(hv) \) (by property (3) of the left \( K[G] \)-module \( V \)), and \( g(hv) = \lambda_g(\lambda_h(v)) \). It follows that
\[
\lambda(gh) = \lambda_{gh} = (\lambda_g)(\lambda_h) = \lambda(g)\lambda(h).
\]

Theorem 2.7.7. Let \( G \) be a group, \( K \) a commutative ring, and \( V \) a right \( K[G] \)-module. Then the map \( \rho : G \to GL(KV) \) is a linear representation given by \( \rho(g) = \rho_g \) where \( (v)\rho_g = vg \) for all \( v \in_K V \).

Proof. The proof is similar, mutatis mutandis, to the proof of Theorem 2.7.6. \( \square \)
CHAPTER 3

LINEAR ALGEBRA OVER NON-COMMUTATIVE RINGS

This chapter extends standard linear algebra theory to matrices over an arbitrary ring $R$. The topics that are discussed include matrix multiplication, coordinate matrices, and row/column operations. We use Friedberg, Insel, and Spence [FIS03, Chapters 2 and 3] and Spence, Insel, Friedberg [SIF00, Chapters 1 and 2] as our standard references, as well as Day [Day, Handout] as a reference for the section on coordinate matrices. The final section of this chapter discusses the sufficient conditions under which certain $R$-modules are actually free $R$-modules.

3.1 Conventions

Let $R$ be a ring. Let $T : A \to B$ be a left $R$-module homomorphism and let $x \in A$. In this chapter, the value of $T$ at $x$ will be written as $xT$ instead of $T(x)$.

Furthermore, the elements of the free left $R$-module $R^m = \bigoplus_{i=1}^{m} R$ will be written in the form $\bar{x} = [x_1, \ldots, x_m]$. Let $c \in R$. Then we write the left action of $c$ on $\bar{x}$ as a scalar-vector product $c\bar{x}$, given by $c\bar{x} = [cx_1, \ldots, cx_m]$.

On the other hand, elements of the right $R$-module $R^n = \bigoplus_{i=1}^{n} R$ will be written
as \( \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \), and we write the right action of \( c \) on \( \vec{x} \) as a **vector-scalar product** \( \vec{x}c \), given by \( \vec{x}c = \begin{bmatrix} x_1c \\ \vdots \\ x_nc \end{bmatrix} \).

**Definition 3.1.1.** Let \( R \) be a ring. Then the set of \( m \times n \) matrices with entries in \( R \) will be denoted by \( M_{m \times n}(R) \). We write \( M_{n \times n}(R) \) as \( M_n(R) \). If \( A \in M_{m \times n}(R) \), then the \( ij \)th entry of \( A \) will be denoted as \( a_{ij} \), where \( i \) represents the row and \( j \) represents the column of the matrix. Furthermore, the \( i \)th row of \( A \) will be denoted by \( \vec{a}_i \) (where \( \vec{a}_i \in R^n \) for \( A = M_{m \times n}(R) \)).

### 3.2 Matrix Multiplication

**Definition 3.2.1.** Let \( A \in M_{m \times n}(R) \) and \( \vec{x} \in R^m \). Then the **row vector-matrix product** of \( \vec{x} \) and \( A \) is given by

\[
\vec{x}A = [x_1 \ldots x_m] \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix} = x_1\vec{a}_1 + \cdots + x_m\vec{a}_m.
\]

Let \( A \in M_{m \times n}(R) \) and \( B \in M_{n \times k}(R) \). Then the **matrix product** \( AB \) is given by

\[
AB = \begin{bmatrix} \vec{a}_1B \\ \vdots \\ \vec{a}_mB \end{bmatrix} \in M_{m \times k}(R).
\]

Note that the \( ij \)th entry of the matrix product \( AB \) may also be depicted as \( (AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \).

**Lemma 3.2.2.** Let \( \vec{x}, \vec{y} \in R^m \) and \( A \in M_{m \times n}(R) \). Then \( (\vec{x} + \vec{y})A = \vec{x}A + \vec{y}A \).
Proof. Let $\vec{x} = [x_1 \ldots x_m]$ and $\vec{y} = [y_1 \ldots y_m]$, where $x_i, y_i \in R$. Then

$$(\vec{x} + \vec{y})A = [(x_1 + y_1) \ldots (x_m + y_m)]$$

$$= (x_1 + y_1)\vec{a}_1 + \cdots + (x_m + y_m)\vec{a}_m$$  \hspace{1cm} (3.1)

$$= (x_1\vec{a}_1 + y_1\vec{a}_1) + \cdots + (x_m\vec{a}_m + y_m\vec{a}_m)$$ \hspace{1cm} (*)

$$= (x_1\vec{a}_1 + \cdots + x_m\vec{a}_m) + (y_1\vec{a}_1 + \cdots + y_m\vec{a}_m)$$

$$= \vec{x}A + \vec{y}A,$$

where (*) follows by property (2) of the left $R$-module $R^n$. \hfill \Box

Lemma 3.2.3. Let $\vec{x} \in R^m$ and $A, B \in M_{m \times n}(R)$. Then $\vec{x}(A + B) = \vec{x}A + \vec{x}B$.

Proof. Let $\vec{x} = [x_1 \ldots x_m]$ where $x_i \in R$. Represent the rows of $A$ and $B$ as $\vec{a}_i$ and $\vec{b}_i$. Then $A + B = \begin{bmatrix} (\vec{a}_1 + \vec{b}_1) \\ \vdots \\ (\vec{a}_m + \vec{b}_m) \end{bmatrix}$, and by Definition 3.2.1, we have

$$\vec{x}(A + B) = [x_1 \ldots x_m] \begin{bmatrix} (\vec{a}_1 + \vec{b}_1) \\ \vdots \\ (\vec{a}_m + \vec{b}_m) \end{bmatrix}$$

$$= x_1(\vec{a}_1 + \vec{b}_1) + \cdots + x_m(\vec{a}_m + \vec{b}_m)$$  \hspace{1cm} (3.2)

$$= (x_1\vec{a}_1 + x_1\vec{b}_1) + \cdots + (x_m\vec{a}_m + x_m\vec{b}_m)$$ \hspace{1cm} (*)

$$= (x_1\vec{a}_1 + \cdots + x_m\vec{a}_m) + (x_1\vec{b}_1 + \cdots + x_m\vec{b}_m)$$

$$= \vec{x}A + \vec{x}B,$$

where (*) follows by property (1) of the left $R$-module $R^n$. \hfill \Box
Lemma 3.2.4. Let $\bar{x} \in \mathbb{R}^m$, $A \in M_{m \times n}(\mathbb{R})$, and $c \in \mathbb{R}$. Then $(c\bar{x})A = c(\bar{x}A)$.

Proof. Let $\bar{x} = [x_1 \ldots x_m]$ where $x_i \in \mathbb{R}$. Then

$$(c\bar{x})A = \begin{bmatrix} cx_1 & \ldots & cx_m \end{bmatrix} \begin{bmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_m \end{bmatrix}$$

$$= (cx_1)\bar{a}_1 + \cdots + (cx_m)\bar{a}_m \quad (3.3)$$

$$= c(x_1\bar{a}_1) + \cdots + c(x_m\bar{a}_m) \quad (*)$$

$$= c[x_1\bar{a}_1 + \cdots + x_m\bar{a}_m] \quad (**)$$

$$= c(\bar{x}A),$$

where $(*)$ follows by property $(3)$ of the left $\mathbb{R}$-module $\mathbb{R}^n$ and $(**)$ follows by property (1) of the left $\mathbb{R}$-module $\mathbb{R}^n$. $\square$

Definition 3.2.5. Let $\bar{x} \in \mathbb{R}^m$ and $A \in M_{m \times n}(\mathbb{R})$. Then $\mu_A : \mathbb{R}^m \to \mathbb{R}^n$ is defined by $\bar{x}\mu_A = \bar{x}A$.

Theorem 3.2.6. Let $A \in M_{m \times n}(\mathbb{R})$. Then $\mu_A$ is a left $\mathbb{R}$-module homomorphism.

Proof. The result follows immediately from Lemmas 3.2.2 and 3.2.4. $\square$

Lemma 3.2.7. Let $\bar{v}_1, \ldots, \bar{v}_p \in \mathbb{R}^m$, $A \in M_{m \times n}(\mathbb{R})$, and $c_1, \ldots, c_p \in \mathbb{R}$. Then

$$(c_1\bar{v}_1 + \cdots + c_p\bar{v}_p)A = c_1(\bar{v}_1A) + \cdots + c_p(\bar{v}_pA).$$

Proof. The result follows by induction on $p$ and Lemmas 3.2.2 and 3.2.4. $\square$

Theorem 3.2.8. Let $R$ be a ring and let $\bar{x} \in \mathbb{R}^m$. Let $A \in M_{m \times n}(\mathbb{R})$ and $B \in M_{n \times k}(\mathbb{R})$. Then $\bar{x}(AB) = (\bar{x}A)B$. 
Proof. Let \( \bar{x} = [x_1 \ldots x_m] \). By Definition 3.2.1,

\[
\bar{x}(AB) = [x_1 \ldots x_m] \begin{bmatrix} \bar{a}_1 B \\ \vdots \\ \bar{a}_m B \end{bmatrix}
\]

\[
= x_1(\bar{a}_1 B) + \cdots + x_m(\bar{a}_m B)
\]

(3.4)

\[
= (x_1\bar{a}_1)B + \cdots + (x_m\bar{a}_m)B \quad \text{by Lemma 3.2.4}
\]

\[
= (x_1\bar{a}_1 + \cdots + x_m\bar{a}_m)B \quad \text{by Lemma 3.2.7}
\]

\[
= (\bar{x}A)B. \quad \Box
\]

Corollary 3.2.9. Let \( R \) be a ring and let \( \bar{x} \in R^m \). Let \( A \in M_{m \times n}(R) \) and \( B \in M_{n \times k}(R) \). Then \( \mu_{AB} = \mu_{A\mu_B} \).

Proof. We see that \( \bar{x}\mu_{AB} = \bar{x}(AB) = (\bar{x}A)B = \bar{x}\mu_{A\mu_B} \).

Corollary 3.2.10. Let \( A \in M_{m \times n}(R) \), \( B \in M_{n \times p}(R) \) and \( C \in M_{p \times k}(R) \). Then \( (AB)C = A(BC) \).

Proof. First note that both \( (AB)C \) and \( A(BC) \) are \( m \times k \) matrices. Let \( \bar{u}_j \) be the \( j \)th row of \( AB \). Then by Definition 3.2.1, \( \bar{u}_j = \bar{a}_j B \). Furthermore, \( \bar{u}_j C \) is the \( j \)th row of \( (AB)C \). Now, \( \bar{u}_j C = (\bar{a}_j B)C = \bar{a}_j(BC) \) (by Theorem 3.2.8). But \( \bar{a}_j(BC) \) is the \( j \)th row of \( A(BC) \), so it follows that \( (AB)C = A(BC) \).

Corollary 3.2.11. Let \( R \) be a ring. Then the set of \( n \times n \) matrices \( A \in M_n(R) \) form a ring under the operations of matrix addition and matrix multiplication.

Proof. The set of \( n \times n \) matrices \( A \in M_n(R) \) form an abelian group under the operation of matrix addition. By Corollary 3.2.10, matrix multiplication is associative. We also claim that the right and left distributive laws hold. That is, for \( A, B, C \in M_n(R) \), we have
(1) \((A + B)C = AC + BC\) and

(2) \(A(B + C) = AB + AC\).

For (1), let \(\vec{a}_i, \vec{b}_i\) represent the rows of \(A\) and \(B\), respectively. Then

\[
(A + B)C = \begin{bmatrix}
(a_1 + b_1)C \\
\vdots \\
(a_m + b_m)C
\end{bmatrix} = \begin{bmatrix}
(a_1 + b_1)C \\
\vdots \\
(a_m + b_m)C
\end{bmatrix} \quad \text{by Definition 3.2.1}
\]

\[
= \begin{bmatrix}
\vec{a}_1 C + \vec{b}_1 C \\
\vdots \\
\vec{a}_m C + \vec{b}_1 C
\end{bmatrix} = \begin{bmatrix}
\vec{a}_1 C \\
\vdots \\
\vec{a}_m C
\end{bmatrix} + \begin{bmatrix}
\vec{b}_1 C \\
\vdots \\
\vec{b}_1 C
\end{bmatrix} = AC + BC
\]

The left distributive law, (2), follows using a similar argument, replacing Lemma 3.2.2 with Lemma 3.2.3. \(\square\)

### 3.3 Coordinate Matrices

**Definition 3.3.1.** Let \(\beta = \{u_1, \ldots, u_n\}\) be an ordered basis for a (finitely generated) free left \(R\)-module \(F\). If \(x \in F\), then by Theorem 2.5.3, \(x = c_1 u_1 + \cdots + c_n u_n\) for some unique \(c_i \in R\). We define the coordinate vector of \(x\) relative to \(\beta\) to be \([x]_\beta = [c_1 \ldots c_n]_\beta\).
Lemma 3.3.2. Let $F$ be a free left $R$-module with basis $\beta = \{u_1, \ldots, u_n\}$. Then for $x, y \in F$ and $a \in R$, we have

1. $[x + y]_\beta = [x]_\beta + [y]_\beta$, and
2. $[ax]_\beta = a[x]_\beta$.

Proof. Let $\beta = \{u_1, \ldots, u_n\}$ be a basis for $F$.

(1) Since $x, y \in F$, it follows that

$$x = c_1 u_1 + \cdots + c_n u_n \quad \text{and} \quad y = d_1 u_1 + \cdots + d_n u_n$$

for some $c_i, d_i \in R$. Furthermore, $x + y = (c_1 + d_1) u_1 + \cdots + (c_n + d_n) u_n$.

Then

$$[x + y]_\beta = [(c_1 + d_1) \ldots (c_n + d_n)]$$

$$= [c_1 \ldots c_n] + [d_1 \ldots d_n] \quad (3.6)$$

$$= [x]_\beta + [y]_\beta.$$

(2) Let $x$ be defined as above. Then

$$[ax]_\beta = [a(c_1 u_1 + \cdots + c_n u_n)]_\beta$$

$$= [a(c_1 u_1) + \cdots + a(c_n u_n)]_\beta$$

$$= [(ac_1) u_1 + \cdots + (ac_n) u_n]_\beta$$

$$= [ac_1 \ldots ac_n] \quad (3.7)$$

$$= a[c_1 \ldots c_n]$$

$$= a[x]_\beta. \quad \square$$

Theorem 3.3.3. Let $F$ be a free left $R$-module with basis $\beta = \{u_1, \ldots, u_n\}$. Then the function $\phi : F \to \mathbb{R}^n$ given by $x\phi = [x]_\beta$ is a left $R$-module isomorphism with inverse $\psi : \mathbb{R}^n \to F$ given by $[c_1 \ldots c_n] \psi = c_1 u_1 + \cdots + c_n u_n$. 
Proof. By Lemma 3.3.2, we see that $\phi$ is a left $R$-module homomorphism.

It remains to show that $\phi$ is a bijection. We see that $\text{Ker}(\phi) = \{0\}$, since
\[ x \in \text{Ker}(\phi) \iff x\phi = [0 \ldots 0] \in R^n \]
\[ \iff [x]_\beta = [0 \ldots 0] \quad \text{(3.8)} \]
\[ \iff x = 0u_1 + \cdots + 0u_n \]
\[ \iff x = 0. \]

Furthermore, suppose $\tilde{y} \in R^n$. Then $\tilde{y} = [r_1 \ldots r_n]$ for some $r_i \in R$. Let $x = r_1u_1 + \cdots + r_nu_n \in F$. Then $x\phi = [r_1 \ldots r_n] = \tilde{y}$.

By Theorem 2.2.3, $\psi = \phi^{-1}$ is itself a left $R$-module isomorphism, and the definition of $\psi$ follows from Definition 3.3.1.

\[ \square \]

Definition 3.3.4. Let $T : F_1 \to F_2$ be a left $R$-module homomorphism between the free left $R$-modules $F_1$ and $F_2$, let $\alpha = \{u_1, \ldots, u_n\}$ be an ordered basis for $F_1$, and let $\beta$ be a finite ordered basis for $F_2$. Then the matrix representation of $T$ in $\alpha$ and $\beta$, denoted $[T]_\alpha^\beta$, is given by $[T]_\alpha^\beta = \begin{bmatrix} [u_1T]_\beta \\ \vdots \\ [u_nT]_\beta \end{bmatrix}$. If $\alpha = \beta$, then $[T]_\alpha^\beta$ is denoted $[T]_\alpha$.

Theorem 3.3.5. Let $T : F_1 \to F_2$ be a left $R$-module homomorphism between the free left $R$-modules $F_1$ and $F_2$, let $\alpha = \{u_1, \ldots, u_n\}$ be an ordered basis for $F_1$, and let $\beta$ be a finite ordered basis for $F_2$. If $x \in F_1$, then $[x]_\alpha[T]_\alpha^\beta = [xT]_\beta$.

Proof. Suppose $x \in F_1$. Then $x = c_1u_1 + \cdots + c_nu_n$ for some $c_i \in R$ and $[x]_\alpha = \ldots$. 
Furthermore,

\[[x]_\alpha [T]_\beta = \begin{bmatrix}
[c_1] \\ 
\vdots \\ 
[c_n]
\end{bmatrix}

= c_1[u_1T]_\beta + \cdots + c_n[u_nT]_\beta

(3.9)

= [c_1(u_1T) + \cdots + c_n(u_nT)]_\beta \quad (*)

= [(c_1u_1 + \cdots + c_nu_n)T]_\beta \quad (**) 

= [xT]_\beta,

where (*) follows by Lemma 3.3.2 and (**) follows because T is a left \( R \)-module homomorphism.

**Corollary 3.3.6.** Let \( F \) be a left \( R \)-module with ordered bases \( \beta \) and \( \beta' \). Let \( I_F \) be the left \( R \)-module identity isomorphism on \( F \) and let \( Q = [I_F]_{\beta}^{\beta'} \). Then for any \( x \in F \),

\([x]_{\beta'} = [x]_{\beta} Q \).

**Proof.** Let \( x \in F \). Then

\([x]_{\beta'} = [xI_F]_{\beta'} = [x]_{\beta}[I_F]_{\beta}^{\beta'} = [x]_{\beta}Q. \quad (3.10)\)

Note that if \( F \) is an \( m \)-dimensional free \( R \)-module with basis \( \beta \), and \( \beta' \) is a re-ordering of \( \beta \), then the matrix \( Q = [I_F]_{\beta}^{\beta'} \) is a permutation of the \( m \times m \) identity matrix \( I_m \). If \( A \in M_{n\times m}(R) \), then the product \( AQ \) will permute the columns of \( A \) in the same manner that the columns of \( I_m \) are permuted in \( Q \). Similarly, if \( B \in M_{m\times n}(R) \), then the product \( QB \) will permute the rows of \( B \) is the same manner as \( Q \).
**Theorem 3.3.7.** Let \( T, U : F_1 \to F_2 \) be left \( R \)-module homomorphisms between the free left \( R \)-modules \( F_1 \) and \( F_2 \), let \( \alpha = \{\overline{u}_1, \ldots, \overline{u}_n\} \) be an ordered basis for \( F_1 \), and let \( \beta \) be a finite ordered basis for \( F_2 \). Then \( [T + U]_\alpha^\beta = [T]_\alpha^\beta + [U]_\alpha^\beta \).

**Proof.** By Definition 3.3.4 and Lemma 3.3.2,

\[
[T + U]_\alpha^\beta = \begin{bmatrix}
[u_1(T + U)]_\beta \\
\vdots \\
[u_n(T + U)]_\beta \\
[u_1 T + u_1 U]_\beta \\
\vdots \\
[u_n T + u_n U]_\beta
\end{bmatrix} = \begin{bmatrix}
[u_1 T]_\beta \\
\vdots \\
[u_n T]_\beta \\
[u_1 U]_\beta \\
\vdots \\
[u_n U]_\beta
\end{bmatrix} = [T]_\alpha^\beta + [U]_\alpha^\beta. \tag{3.11}
\]

**Theorem 3.3.8.** Let \( F_1, F_2, \) and \( F_3 \) be free left \( R \)-modules and let \( T : F_1 \to F_2 \) and \( U : F_2 \to F_3 \) be left \( R \)-module homomorphisms. If \( \alpha = \{u_1, \ldots, u_n\} \) is an ordered basis for \( F_1 \), \( \beta \) is an ordered basis for \( F_2 \), and \( \gamma \) is an ordered basis for \( F_3 \), then \( [T]_\alpha^\beta[U]_\beta^\gamma = [TU]_\alpha^\gamma \).

**Proof.** First note that by Theorem 3.3.5, for \( v_i \in F_1 \), we have

\[
[v_i T]_\beta[U]_\beta^\gamma = [v_i TU]_\gamma. \tag{3.12}
\]
Then

\[
[T]_{\alpha}^{\beta} [U]_{\beta}^{\gamma} = \begin{bmatrix}
[u_1 T]_{\beta}^{\gamma} \\
\vdots \\
[u_n T]_{\beta}^{\gamma}
\end{bmatrix}
\]

by definition of matrix multiplication

\[\text{(3.13)}\]

\[
\begin{align*}
= & \begin{bmatrix}
[u_1 T]_{\beta}^{\gamma} [U]_{\beta}^{\gamma} \\
\vdots \\
[u_n T]_{\beta}^{\gamma} [U]_{\beta}^{\gamma}
\end{bmatrix} \\
= & [TU]_{\gamma}^{\gamma}
\end{align*}
\]

by (3.12)

\[
= [TU]_{\alpha}^{\gamma}. \quad \Box
\]

**Theorem 3.3.9.** Let End$_R(R^n)$ be the ring of left $R$-module endomorphisms on $R^n$. Then End$_R(R^n)$ and $M_n(R)$ are isomorphic as rings.

**Proof.** Define $\Phi : \text{End}_R(R^n) \to M_n(R)$ by $\Phi(T) = [T]_{\beta}$ where $\beta = \{e_1, \ldots, e_n\}$ is the standard basis of $R^n$. By Theorems 3.3.7 and 3.3.8, it follows that $\Phi$ is a ring homomorphism. It remains to show that $\Phi$ is a bijection.

By Theorem 2.5.4, for any $\bar{a}_1, \ldots, \bar{a}_n \in R^n$, there exists a unique left $R$-module homomorphism $T$ such that $e_i T = \bar{a}_i$ for all $i$. In other words, if we choose an arbitrary matrix $A \in M_n(R)$ such that $A = \begin{bmatrix}
\bar{a}_1 \\
\vdots \\
\bar{a}_n
\end{bmatrix}$, then there exists a unique left $R$-module homomorphism $T \in \text{End}_R(R^n)$ such that

\[
[T]_{\beta} = \begin{bmatrix}
[e_1 T]_{\beta} \\
\vdots \\
[e_n T]_{\beta}
\end{bmatrix}
= \begin{bmatrix}
\bar{a}_1 \\
\vdots \\
\bar{a}_n
\end{bmatrix}
= A. \quad \text{(3.14)}
\]
This simultaneously proves that $\Phi$ is both surjective and injective. 

### 3.4 Row Operations

**Definition 3.4.1.** Let $A$ be a matrix over a ring $R$ with identity. The following are referred to as *elementary row operations* on $A$:

1. Interchange two rows of $A$.
2. Left multiply a row of $A$ by a unit $k \in R$.
3. For $r \in R$ and $i \neq j$ add $r$ times row $i$ to row $j$. That is, if $A$ is an $n \times m$ matrix, let row $i$ be $[a_{i1} \ a_{i2} \ \ldots \ \ a_{im}]$ and row $j$ be $[a_{j1} \ a_{j2} \ \ldots \ \ a_{jm}]$. Then we replace row $j$ by
   \[
   r[a_{i1} \ a_{i2} \ \ldots \ \ a_{im}] + [a_{j1} \ a_{j2} \ \ldots \ \ a_{jm}] = [ra_{i1} + a_{j1} \ ra_{i2} + a_{j2} \ \ldots \ \ ra_{im} + a_{jm}].
   \]

**Definition 3.4.2.** Let $A$ be an $n \times m$ matrix over a ring $R$ with identity. Let $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n$ represent the $n$ rows of $A$. Then the row space of $A$ (denoted $\text{Row}(A)$) is the set of all left $R$-linear combinations of the rows of $A$. In other words, $\text{Row}(A) = \langle \bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n \rangle = \{x_1 \bar{a}_1 + \cdots + x_n \bar{a}_n | x_i \in R \}$.

**Theorem 3.4.3.** Let $A$ be an $n \times m$ matrix over a ring $R$. Elementary row operations leave $\text{Row}(A)$ unchanged.

**Proof.** Denote the rows of $A$ by $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n$.

1. Let $A'$ be the matrix obtained by interchanging $\bar{a}_i$ and $\bar{a}_j$ of $A$, $1 \leq i, j \leq n$, $(i \neq j)$. Then
   \[
   \text{Row}(A) = \langle \bar{a}_1, \ldots, \bar{a}_i, \ldots, \bar{a}_j, \ldots, \bar{a}_n \rangle = \langle \bar{a}_1, \ldots, \bar{a}_j, \ldots, \bar{a}_i, \ldots, \bar{a}_n \rangle = \text{Row}(A').
   \]
(2) Let $A''$ be the matrix obtained by multiplying $\bar{a}_i$ $(1 \leq i \leq n)$ on the left by a unit $k \in R$. Then the rows of $A''$ are: $\bar{a}_1, \ldots, k\bar{a}_i, \ldots, \bar{a}_n$, and

Row$(A) = \text{Row}(A'')$ because

$$x \in \text{Row}(A) \iff x = c_1\bar{a}_1 + \cdots + c_i\bar{a}_i + \cdots + c_n\bar{a}_n,$$

for some $c_j \in R$

$$\iff x = c_1\bar{a}_1 + \cdots + (c_i k^{-1})(k\bar{a}_i) + \cdots + c_n\bar{a}_n$$

$$\iff x \in \text{Row}(A'').$$

Note that $\{c_j k^{-1} | c_j \in R\} = R$ because $k$ is a unit.

(3) Let $A'''$ be the matrix obtained by adding $p$ times row $i$ to row $j$. Then the rows of $A'''$ are: $\bar{a}_1, \ldots, \bar{a}_i, \ldots, p\bar{a}_i + \bar{a}_j, \ldots, \bar{a}_n$, and Row$(A) = \text{Row}(A''')$ because

$$x \in \text{Row}(A) \iff x = c_1\bar{a}_1 + \cdots + c_i\bar{a}_i + \cdots + c_j\bar{a}_j + \cdots + c_n\bar{a}_n,$$

for some $c_k \in R$

$$\iff x = c_1\bar{a}_1 + \cdots + (c_i - c_j p)\bar{a}_i + \cdots + c_j(p\bar{a}_i + \bar{a}_j) + \cdots + c_n\bar{a}_n$$

$$\iff x \in \text{Row}(A''').$$

Note that given $c_j \in R$, $\{c_j - c_j p | c_j \in R\} = R$.

It follows that elementary row operations leave Row$(A)$ unchanged. \hfill \square

**Definition 3.4.4.** Let $R$ be a ring with identity $1_R$. Then the $n \times n$ identity matrix $A = I_n$ is the matrix with entries

$$a_{ij} = \begin{cases} 1_R & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \quad (3.15)$$

**Definition 3.4.5.** Let $I_n$ be the $n \times n$ identity matrix. Then the **elementary matrix** $E$ is obtained by applying a single elementary row operation to $I_n$. There are three types of elementary matrices:
(1) The first type of elementary matrix, \( E_1(i, j) \), is obtained by interchanging row \( i \) and row \( j \) of the \( n \times n \) identity matrix.

(2) The second type of elementary matrix, \( E_2(i, k) \), is obtained by multiplying row \( i \) of \( I_n \) by a unit \( k \in R \).

(3) Finally, the third type of elementary matrix, \( E_3(i, j, p) \), is obtained by adding \( p \) times row \( i \) of \( I_n \) to row \( j \).

**Lemma 3.4.6.** Let \( A \) be an \( m \times n \) matrix over a ring \( R \). Let \( E \) be an elementary matrix obtained by applying an elementary row operation to the identity matrix \( I_m \). Then the product \( EA \) achieves the same result as directly performing the identical elementary row operation on \( A \). Furthermore, all elementary matrices are invertible.

**Proof.** By definition, \( I_m = \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_k \\ \vdots \\ \vec{e}_m \end{bmatrix} \) where \( \vec{e}_k \) are the standard basis vectors of \( R^m \) and

\[
A = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}
\]

Define \( E_1 = E_1(i, j) = \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_j \\ \vdots \\ \vec{e}_m \end{bmatrix} \). By definition of matrix multiplication and
the standard basis vectors, it follows that

$$E_1 A = \begin{bmatrix} \bar{e}_1 A & \bar{a}_1 \\ \vdots & \vdots \\ \bar{e}_j A & \bar{a}_j \\ \vdots & \vdots \\ \bar{e}_i A & \bar{a}_i \\ \vdots & \vdots \\ \bar{e}_m A & \bar{a}_m \end{bmatrix}$$

Thus, the product $E_1 A$ achieves the same result as interchanging row $i$ and row $j$ of $A$.

Let $k$ be a unit in $R$. Define $E_2 = E_2(i, k) = k\bar{e}_i$. Then

$$E_2 A = \begin{bmatrix} \bar{e}_1 A & \bar{a}_1 \\ \vdots & \vdots \\ k\bar{e}_i A & k\bar{a}_i \\ \vdots & \vdots \\ \bar{e}_m A & \bar{a}_m \end{bmatrix}$$

Thus, the product $E_2 A$ achieves the same result as multiplying row $i$ of $A$ on the left by the unit $k$. 
Finally, let \( p \in R \). Define \( E_3 = E_3(i, j, p) = \begin{bmatrix} \bar{e}_1 \\ \vdots \\ (p\bar{e}_i + \bar{e}_j) \\ \vdots \\ \bar{e}_m \end{bmatrix} \). Then

\[
E_3A = \begin{bmatrix} \bar{e}_1A \\ \vdots \\ (p\bar{e}_i + \bar{e}_j)A \\ \vdots \\ \bar{e}_mA \end{bmatrix} = \begin{bmatrix} \bar{a}_1 \\ \vdots \\ p\bar{a}_i + \bar{a}_j \\ \vdots \\ \bar{a}_m \end{bmatrix}.
\] (3.18)

The product \( E_3A \) attains the same result as adding \( p \) times row \( i \) of \( A \) to row \( j \).

We can use the results above to define the inverses of \( E_1, E_2, \) and \( E_3 \).

Let \( E_1 \) be defined as above, and consider the product \( E_1E_1 \). By equation (3.16), it follows that \( E_1E_1 = I_m \).

Define \( E_2 \) as above, and let \( E'_2 = E_2(i, k^{-1}) = k^{-1}\bar{e}_i \). Then by equation (3.17), it follows that \( E_2E'_2 = E'_2E_2 = I_m \).

Similarly, if we define \( E_3 \) as above and \( E'_3 = E_3(i, j, -p) = (-p\bar{e}_i + \bar{e}_j) \).

Then by equation (3.18), \( E_3E'_3 = E'_3E_3 = I_m \).
Definition 3.4.7. A matrix \( A \) over a ring \( R \) is said to be in **reduced row echelon form** (rref) provided that:

1. Any row containing a non-zero entry precedes any row in which all of the entries are zero;

2. If a row is non-zero, then the first non-zero entry in that row will be the only non-zero entry in its column; and

3. If a row is non-zero, then the first non-zero entry is 1 and occurs in a column strictly to the right of the first non-zero entry in the preceding row.

Definition 3.4.8. Let \( A \) be a matrix over a ring \( R \). Then \( A \) is said to be **\( R \)-row-reducible** provided that it may be put into reduced row echelon form using elementary row operations.

Note that we disregard the issue of whether \( \text{rref}(A) \) is unique and use \( \text{rref}(A) \) to denote any matrix in reduced row echelon form that is \( R \)-row equivalent to \( A \).

Additionally note that by induction and Lemma 3.4.6, if a matrix \( A \) is \( R \)-row-reducible, then there exists an invertible matrix \( P \) such that \( PA = \text{rref}(A) \).

Definition 3.4.9. Let \( A \) be a matrix with entries in a division ring \( D \). The **Gaussian elimination** algorithm is defined as follows:

1. If necessary, interchange rows so that the entry in the first row and leftmost non-zero column of \( A \) is non-zero. Then, if required, multiply the first row by a unit so that the entry in the leftmost non-zero column is \( 1_D \).

2. Add appropriate multiples of the first row to the subsequent rows so that all of the entries below the \( 1_D \) (obtained in step (1)) become zero.
(3) Repeat step (1) for the next row, without using any previous rows.

(4) Using the same procedure as in step (2) zero out all entries below the 1\(_D\).

(5) Repeat steps (3) and (4) on each succeeding row until no non-zero entries remain.

(6) Begin with the last non-zero row and add appropriate multiples of that row to the preceding rows to zero out all entries above its first non-zero entry.

(7) Repeat step (6) on the preceding rows.

**Theorem 3.4.10.** Every matrix with entries in a division ring \(D\) is \(D\)-row-reducible.

**Proof.** We proceed by induction on \(n\), the number of rows in a matrix.

Let \(A\) be a \(1 \times m\) matrix with entries in a division ring \(D\). If \(A = \begin{bmatrix} 0 & 0 & \ldots & 0 \end{bmatrix}\), then we are done. Otherwise, if required, multiply the matrix by a unit \(a \in D\) in order to transform the leftmost entry of \(A\) into 1\(_D\). Then \(A\) will be in reduced row echelon form, and therefore the base case is true.

Suppose the theorem is true for \(k\). Let \(A\) be a \((k+1) \times m\) matrix. Apply steps (1) and (2) of Gaussian elimination to \(A\) to obtain a matrix that looks like

\[
\begin{bmatrix}
0 & \ldots & 0 & 1 & * & \ldots & * \\
0 & \ldots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & A' \\
0 & \ldots & 0 & 0
\end{bmatrix}
\]

where \(A'\) is a matrix with \(k\) rows. We can apply the induction hypothesis to \(A'\) to put \(A'\) in reduced row echelon form and then use steps (6) and (7) on \(A\) to clear any entries in the first row that lie above each leading 1. The resulting matrix will be in reduced row echelon form. \(\square\)
3.5 Column Operations

Please refer to section 3.4 for proofs of the theorems and lemmas below (shown there in terms of row operations).

**Definition 3.5.1.** Let $A$ be a matrix over a ring $R$ with identity. The following are referred to as *elementary column operations* on $A$:

1. Interchange two columns of $A$;
2. Right multiply a column of $A$ by a unit $a \in R$;
3. For $r \in R$ and $i \neq j$ add column $i$ times $r$ to column $j$.

**Definition 3.5.2.** Let $A$ be an $n \times m$ matrix over a ring $R$ with identity. Let $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_m$ represent the $m$ columns of $A$. Then the **column space** of $A$ (denoted $\text{Col}(A)$) is the set of all right $R$-linear combinations of the columns of $A$. In other words, $\text{Col}(A) = \langle \bar{c}_1, \bar{c}_2, \ldots, \bar{c}_m \rangle$ (a right $R$-module).

**Theorem 3.5.3.** Let $A$ be an $n \times m$ matrix over a ring $R$ with identity. Elementary column operations leave $\text{Col}(A)$ unchanged.

*Proof.* See proof of Theorem 3.4.3.

**Definition 3.5.4.** Let $I_n$ be the $n \times n$ identity matrix. Then the **elementary matrix** $E$ is obtained by applying a single elementary column operation to $I_n$.

**Lemma 3.5.5.** Let $A$ be an $m \times n$ matrix over a ring $R$. Let $E$ be an elementary matrix obtained by applying an elementary row operation to the identity matrix $I_n$. Then the product $AE$ achieves the same result as directly performing the identical elementary column operation on $A$. 

\[ \square \]
Definition 3.5.6. A matrix $A$ over a ring $R$ is said to be in \textbf{reduced column echelon form} (reref) provided that:

1. Any column containing a non-zero entry precedes any column in which all of the entries are zero;

2. The first non-zero entry in each column is the only non-zero entry in its row; and

3. The first non-zero entry in each column is 1 and occurs in a row strictly below the first non-zero entry in the preceding column.

Definition 3.5.7. Let $A$ be a matrix over a ring $R$. Then $A$ is said to be \textbf{$R$-column-reducible} provided that it may be put into reduced column echelon form using elementary column operations.

Theorem 3.5.8. If $A$ is $R$-column-reducible, then $\text{reref}(A) = AP$ for some invertible matrix $P$. □

Theorem 3.5.9. Every matrix with entries in a division ring $D$ is $D$-column-reducible. □

3.6 Some Sufficient Conditions for Freeness

The following two theorems specify the conditions under which certain (left) $R$-modules are actually \textit{free} (left) $R$-modules.

Theorem 3.6.1. Let $R$ be a ring, let $M$ be a free left $R$-module with ordered basis $\beta = \{u_1, \ldots, u_m\}$, and let $N = \langle v_1, \ldots, v_n \rangle$ be a submodule of $M$. Suppose that $A = \begin{bmatrix} [v_1]_\beta \\ \vdots \\ [v_n]_\beta \end{bmatrix}$ is an $n \times m$ $R$-row-reducible matrix. Suppose further that $\beta$ is ordered
so that the first $k$ columns of \( \text{rref}(A) \) form the $k \times k$ identity matrix with $n - k$ rows of zeroes underneath. That is,

\[
\text{rref}(A) = \begin{bmatrix}
1 & 0 & \cdots & 0 & r_{1,k+1} & \cdots & r_{1m} \\
0 & 1 & \cdots & 0 & r_{2,k+1} & \cdots & r_{2m} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & r_{k,k+1} & \cdots & r_{km} \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 
\end{bmatrix}
\] (3.19)

Finally, let \( \phi : M \rightarrow \mathbb{R}^m \) be defined as \( \pi \phi = [x]_\beta \) and let \( \psi = \phi^{-1} \). Then

1. \( N \) is a free left \( \mathbb{R} \)-module with basis \( \{ \bar{r}_1 \psi, \ldots, \bar{r}_k \psi \} \), where \( \bar{r}_i \) is the \( i \)th non-zero row of \( \text{rref}(A) \).

2. \( M/N \) is a free left \( \mathbb{R} \)-module with basis \( \{ \bar{e}_{k+1} \psi + N, \ldots, \bar{e}_m \psi + N \} \), where \( \bar{e}_i \) is the \( i \)th standard basis vector of \( \mathbb{R}^m \) (see Definition 2.5.6).

Note that for any \( \mathbb{R} \)-row-reducible matrix \( A \), whose reduced row echelon form does not match the format of (3.19), there exists a permutation matrix \( Q \) such that \( \text{rref}(A)Q \) has the layout of (3.19). In the case of Theorem 3.6.1, where \( A = \begin{bmatrix} [v_1]_\beta \\ \vdots \\ [v_n]_\beta \end{bmatrix} \), the basis \( \beta \) may be reordered to form the matrix (3.19). Specifically, by Corollary 3.3.6 (and the note afterwards), if we let \( \beta' \) represent a reordering of \( \beta \) and \( Q = [I_M]_{\beta'}^\beta \), then

\[
[v_i]_{\beta'} = [v_i]_\beta [I_M]_{\beta'}^\beta = [v_i]_\beta Q, \text{ and the matrix } \begin{bmatrix} [v_1]_{\beta'} \\ \vdots \\ [v_n]_{\beta'} \end{bmatrix} = AQ \text{ will still be a matrix representation of the submodule } N. \text{ Furthermore, } \text{rref}(A)Q = \text{rref}(AQ). \text{ That is, since}
A is $R$-row-reducible, there exists an invertible matrix $P$ such that $PA = \text{rref}(A)$. By the argument above, the right-hand side of (3.19) = $\text{rref}(A)Q = (PA)Q = P(AQ)$. Since (3.19) is in reduced row echelon form, we must have $P(AQ) = \text{rref}(AQ)$.

**Proof.**

(1) Let $A' = \text{rref}(A)$. Let $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_k \in R^m$ be the non-zero rows of $A'$. Suppose $c_1 \bar{r}_1 + c_2 \bar{r}_2 + \cdots + c_k \bar{r}_k = \vec{0}$ (the $1 \times m$ zero row). Then by (3.19) above, it follows that $[c_1 \ldots c_k * \ldots *] = \vec{0}$, so $c_1 = c_2 = \cdots = c_k = 0$. Thus, the non-zero rows of $A'$ are linearly independent.

Furthermore, since $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_k$ are the non-zero rows of $A$ and $\bar{r}_{k+1}, \ldots, \bar{r}_n$ are zero rows, it follows that

$$\langle \bar{r}_1, \ldots, \bar{r}_k \rangle = \langle \bar{r}_1, \ldots, \bar{r}_n \rangle = \text{Row}(A') = \text{Row}(A) = N\phi, \quad (3.20)$$

where $\text{Row}(A') = \text{Row}(A)$ by Theorem 3.4.3.

Thus, the non-zero rows of $A'$ form a basis for $\text{Row}(A)$.

By Theorem 3.3.3, $\phi$ and $\psi$ are $R$-module isomorphisms. Since $R$-module isomorphisms map bases onto bases, the set $\{\bar{r}_1 \psi, \ldots, \bar{r}_k \psi\}$ is a basis for $N$.

(2) We showed in (1) that $\{\bar{r}_1, \ldots, \bar{r}_k\}$ is a basis for $N\phi$. We claim that $\{\bar{r}_1, \ldots, \bar{r}_k, \bar{e}_{k+1}, \ldots, \bar{e}_m\}$ is a basis for $R^m$, where $\bar{e}_i$ are the standard basis vectors in $R^m$.

Suppose

$$c_1 \bar{r}_1 + \cdots + c_k \bar{r}_k + c_{k+1} \bar{e}_{k+1} + \cdots + c_m \bar{e}_m = \vec{0}. \quad (3.21)$$

By assumption, $\bar{r}_i = [0 \ 0 \ldots \ 1 \ldots \ r_{i,k+1} \ldots \ r_{im}]$ where the $i$th coordinate of vector $\bar{r}_i$ is 1. For $k+1 \leq i \leq m$, let

$$b_i = c_1 r_{i1} + \cdots + c_k r_{ki} + c_i. \quad (3.22)$$
Then (3.21) implies that

\[
\begin{bmatrix}
  c_1 & \ldots & c_k & b_{k+1} & \ldots & b_m
\end{bmatrix} = 0.
\] (3.23)

We immediately see that \( c_i = 0 \) for all \( i, 1 \leq i \leq k \). When we plug in the values of \( c_1 = \ldots = c_k = 0 \) into (3.22), and use \( b_i = 0 \), we get \( c_{k+1} = \ldots = c_m = 0 \). It follows that \( c_i = 0 \) for all \( i, 1 \leq i \leq m \).

Suppose \( \vec{y} \in \mathbb{R}^m \). Then \( \vec{y} = [d_1 \ldots d_m] \) for some \( d_i \in \mathbb{R} \).

Let

\[
c_j = \begin{cases} 
  d_j & \text{for } 1 \leq j \leq k \\
  -d_1 r_{1j} - \cdots - d_k r_{kj} + d_j & \text{for } k + 1 \leq j \leq m
\end{cases}
\] (3.21)

where \( r_{ij} \) is the \( j \)th entry in vector \( \vec{r}_i \).

Then by using the fact that

\[
c_1 \vec{r}_1 + \cdots + c_k \vec{r}_k + c_{k+1} \vec{e}_{k+1} + \cdots + c_m \vec{e}_m = \begin{bmatrix} c_1 & \ldots & c_k & b_{k+1} & \ldots & b_m \end{bmatrix} (3.25)
\]

and plugging the \( c_j \) into the right-hand side of (3.25), we get

\[
c_1 \vec{r}_1 + \cdots + c_k \vec{r}_k + c_{k+1} \vec{e}_{k+1} + \cdots + c_m \vec{e}_m = [d_1 \ldots d_m] = \vec{y}. (3.26)
\]

It follows that \( \{\vec{r}_1, \ldots, \vec{r}_k, \vec{e}_{k+1}, \ldots, \vec{e}_m\} \) is a basis for \( \mathbb{R}^m \).

Let \( \vec{N} = \vec{N} \phi \). We claim that \( \alpha = \{\vec{e}_{k+1} + \vec{N}, \ldots, \vec{e}_m + \vec{N}\} \) is a basis for \( \mathbb{R}^m / \vec{N} \).

Suppose

\[
a_{k+1}(\vec{e}_{k+1} + \vec{N}) + \cdots + a_m(\vec{e}_m + \vec{N}) = \vec{N}
\] (3.27)

for some \( a_i \in \mathbb{R} \). Then \( (a_{k+1}\vec{e}_{k+1} + \cdots + a_m\vec{e}_m) + \vec{N} = \vec{N} \). It follows that \( a_{k+1}\vec{e}_{k+1} + \cdots + a_m\vec{e}_m \in \vec{N} \). Since \( \{\vec{r}_1, \ldots, \vec{r}_k\} \) is a basis for \( \vec{N} \), it follows that
\[ a_{k+1} \bar{e}_{k+1} + \cdots + a_m \bar{e}_m = b_1 \bar{r}_1 + \cdots + b_k \bar{r}_k, \] where \( b_j \in R \). Then \((-b_1)\bar{r}_1 + \cdots + (-b_k)\bar{r}_k + a_{k+1} \bar{e}_{k+1} + \cdots + a_m \bar{e}_m = 0\). Since \( \{\bar{r}_1, \ldots, \bar{r}_k, \bar{e}_{k+1}, \ldots, \bar{e}_m\} \) is a basis for \( R^m \), it follows that \( a_i = b_j = 0 \) for all \( i, j \). Therefore, \( \alpha \) is linearly independent.

Suppose that \( x = \bar{m} + \bar{N} \) for some \( \bar{m} \in R^m \). Since \( \bar{m} \in R^m \), \( \bar{m} = a_1 \bar{r}_1 + \cdots + a_k \bar{r}_k + a_{k+1} \bar{e}_{k+1} + \cdots + a_m \bar{e}_m \) for some \( a_j \in R \). So,

\[
x = (a_1 \bar{r}_1 + \cdots + a_k \bar{r}_k + a_{k+1} \bar{e}_{k+1} + \cdots + a_m \bar{e}_m) + \bar{N}
\]

\[
= (a_{k+1} \bar{e}_{k+1} + \cdots + a_m \bar{e}_m) + \bar{N}
\]

\[
= a_{k+1} (\bar{e}_{k+1} + \bar{N}) + \cdots + a_m (\bar{e}_m + \bar{N})
\]

\[
\in \langle \bar{e}_{k+1} + \bar{N}, \ldots, \bar{e}_m + \bar{N} \rangle,
\]

where \( (*) \) follows because \( a_1 \bar{r}_1 + \cdots + a_k \bar{r}_k \in \bar{N} \).

It follows that \( \langle \bar{e}_{k+1} + \bar{N}, \ldots, \bar{e}_m + \bar{N} \rangle = R^m / \bar{N} \). Again, since \( \psi \) is a left \( R \)-module isomorphism and \( \bar{N} \psi = N \), it follows that \( \{\bar{e}_{k+1} \psi + N, \ldots, \bar{e}_m \psi + N\} \) is a basis for \( M/N \).

**Theorem 3.6.2.** Let \( A \) be an \( m \times n \) matrix over a ring \( R \) such that \( A \) is \( R \)-column-reducible. Let \( A' = \text{rref}(A) \). Then \( \text{Ker}(\mu_{A'}) = \text{Ker}(\mu_A) \).

**Proof.** Suppose \( A' = \text{rref}(A) \). Then by Theorem 3.5.8 there exists an invertible \( n \times n \)
matrix $P$ such that $AP = A'$. Then

$$x \in \text{Ker}(\mu_{A'}) \iff x_{\mu_{A'}} = 0$$
$$\iff xA' = 0$$
$$\iff x(AP) = 0$$
$$\iff (xA)P = 0$$
$$\iff (xA)PP^{-1} = 0P^{-1}$$
$$\iff xA = 0$$
$$\iff x \in \text{Ker}(\mu_A).$$

**Theorem 3.6.3.** Let $M$ and $N$ be finitely generated free left $R$-modules with ordered bases $\alpha = \{u_1, \ldots, u_m\}$ and $\beta = \{w_1, \ldots, w_n\}$, respectively, let $T : M \to N$ be a left $R$-module homomorphism, and let $A = [T]^\beta_\alpha$. Assume that $A$ is $R$-column-reducible and that

$$\text{rcf}(A) = \begin{bmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\
a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,k} & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
a_{m,1} & a_{m,2} & \cdots & a_{m,k} & 0 & \cdots & 0
\end{bmatrix}. \quad (3.30)$$

Let

$$\vec{v}_i = \begin{bmatrix}
-a_{i1} & -a_{i2} & \cdots & -a_{ik} & 0 & \cdots & 1 & \cdots & 0
\end{bmatrix}, \quad (3.31)$$

(where the $i$th coordinate of $\vec{v}_i$ is 1), let $\phi : M \to R^m$ be defined as $x\phi = [x]_\alpha$, and let $\psi = \phi^{-1}$.

Then $\text{Ker}(T)$ is a free left $R$-module with basis $\{\vec{v}_{k+1}\psi, \ldots, \vec{v}_m\psi\}$. 

Note that for any $R$-column-reducible matrix $A$ whose reduced column echelon form does not match the format of (3.30), there exists a permutation matrix $Q$ such that $Q \text{rref}(A)$ has the layout of (3.30). In the case of Theorem 3.6.2, where $A = [T]_{\alpha}^\beta$, the basis $\alpha$ may be reordered to form the matrix (3.30). Specifically, by Corollary 3.3.6 (and the note afterwards), if we let $\alpha'$ represent a reordering of $\alpha$ and $Q = [I_M]_{\alpha'}^\beta$, then $QA = [I_M]_{\alpha'}^\beta[T]_{\alpha}^\beta = [I_M T]_{\alpha'}^\beta$ and $Q \text{rref}(A) = \text{rref}(QA)$. That is, since $A$ is $R$-column-reducible, there exists an invertible matrix $P$ such that $AP = \text{rref}(A)$. By the argument above, the right-hand side of (3.30) = $Q \text{rref}(A) = Q(AP) = (QA)P$. Since (3.30) is in reduced column echelon form, we must have $(QA)P = \text{rref}(QA)$.

Proof. Let $A$ be an $m \times n$ matrix. Let $A' = \text{rref}(A)$ be as above.

Suppose $xA' = 0$ where $x = [x_1 \ x_2 \ \ldots \ x_m]^T \in \mathbb{R}^m$. Then $xA' = 0$ implies:

\begin{align*}
 x_1 + x_{k+1}a_{k+1,1} + \cdots + x_m a_{m1} &= 0 \\
 x_2 + x_{k+1}a_{k+1,2} + \cdots + x_m a_{m2} &= 0 \\
 &\vdots \\
 x_k + x_{k+1}a_{k+1,k} + \cdots + x_m a_{mk} &= 0.
\end{align*}

Therefore,

\begin{align*}
 x_1 &= x_{k+1}(-a_{k+1,1}) + \cdots + x_m(-a_{m1}) \\
 x_2 &= x_{k+1}(-a_{k+1,2}) + \cdots + x_m(-a_{m2}) \\
 &\vdots \\
 x_k &= x_{k+1}(-a_{k+1,k}) + \cdots + x_m(-a_{mk}).
\end{align*}
so,

\[
\begin{bmatrix}
  x_1 & x_2 & \ldots & x_m
\end{bmatrix}
= x_{k+1}
\begin{bmatrix}
  -a_{k+1,1} & -a_{k+1,2} & \ldots & -a_{k+1,k} & 1 & 0 & \ldots & 0
\end{bmatrix}
+ x_{k+2}
\begin{bmatrix}
  -a_{k+2,1} & -a_{k+2,2} & \ldots & -a_{k+2,k} & 0 & 1 & \ldots & 0
\end{bmatrix}
\]

\vdots

+ x_m
\begin{bmatrix}
  -a_{m1} & -a_{m2} & \ldots & -a_{mk} & 0 & \ldots & 0 & 1
\end{bmatrix}.
\]

(3.34)

Let

\[
\vec{v}_{k+1} =
\begin{bmatrix}
  -a_{k+1,1} & -a_{k+1,2} & \ldots & -a_{k+1,k} & 1 & 0 & \ldots & 0
\end{bmatrix}
\]

\vdots

\vec{v}_m =
\begin{bmatrix}
  -a_{m1} & -a_{m2} & \ldots & -a_{mk} & 0 & \ldots & 0 & 1
\end{bmatrix}.
\]

so that (3.34) may be written as:

\[
\begin{bmatrix}
  x_1 & x_2 & \ldots & x_m
\end{bmatrix} = x_{k+1}\vec{v}_{k+1} + \cdots + x_m\vec{v}_m.
\]

We claim that the set \(\{\vec{v}_{k+1}, \ldots, \vec{v}_m\}\) is a basis for \(\text{Ker}(\mu_A)\).

First of all, (3.32) – (3.35) show that \(\text{Ker}(\mu_{A'}) \subseteq \text{span}\{\vec{v}_{k+1}, \ldots, \vec{v}_m\}\). Additionally, note that \(\vec{v}_i \in \text{Ker}(\mu_{A'})\) for \(k+1 \leq i \leq m\) because \(\vec{v}_i A' = \vec{0}\) for all \(i\). So, \(\text{span}\{\vec{v}_{k+1}, \ldots, \vec{v}_m\} \subseteq \text{Ker}(\mu_{A'})\) (because \(\text{Ker}(\mu_{A'})\) is a submodule of the left \(R\)-module \(M\)).

Furthermore, suppose \(d_{k+1}\vec{v}_{k+1} + \cdots + d_m\vec{v}_m = \vec{0}\). Then by definition of \(\vec{v}_{k+1}, \ldots, \vec{v}_m\), it follows that \([* * \ldots * d_{k+1} d_{k+2} \ldots d_m] = \vec{0}\) which implies that \(d_{k+1} = \cdots = d_m = 0\).

It follows that \(\{\vec{v}_{k+1}, \ldots, \vec{v}_m\}\) is a basis for \(\text{Ker}(\mu_{A'})\). By Theorem 3.6.2, we have that \(\{\vec{v}_{k+1}, \ldots, \vec{v}_m\}\) is a basis for \(\text{Ker}(\mu_A) \subseteq R^m\).

By Theorem 3.3.3, \(\psi : R^m \rightarrow M\) is an \(R\)-module isomorphism. Since \(R\)-module isomorphisms map bases onto bases, the set \(\{\vec{v}_{k+1}\psi, \ldots, \vec{v}_m\psi\}\) is a basis for \(\text{Ker}(T)\). \(\square\)
4.1 Review of Analysis/ Topology

4.1.1 Basic Topology of Metric Spaces

The following definitions and theorems reference Rudin [Rud76, Ch. II].

Definition 4.1.1. A metric space is a set $S$ and a function $d : S \times S \to \mathbb{R}$ such that for every $p, q, r \in S$,

1. $d(p, q) > 0$ if $p \neq q$, $d(p, p) = 0$;
2. $d(p, q) = d(q, p)$; and
3. $d(p, q) \leq d(p, r) + d(r, q)$.

Such a function $d$ is called a metric on $S$, and the metric space together with its metric is denoted as $(S, d)$.

Examples:

1. The set $\mathbb{R}$ of real numbers is a metric space with the absolute value metric

$$d(x, y) = |x - y|.$$
(2) For \( x, y \in \mathbb{C} \), where \( x = a + bi \) and \( y = c + di \), define the absolute value metric by 
\[
 d(x, y) = |x - y| = |(a - c) + (b - d)i| = \sqrt{(a - c)^2 + (b - d)^2}.
\]
Then \( \mathbb{C} \) is a metric space.

(3) Every subset of a metric space \( S \) is itself a metric space with the same metric as \( S \).

**Definition 4.1.2.** Let \( \prod_{i=1}^{k} V_i \) be a direct product of metric spaces, each with metric \( d_{V_i} \), respectively. Let \( \vec{x}, \vec{y} \in \prod_{i=1}^{k} V_i \). Then the **product metric** \( d = d_{\prod V_i} \) is given by 
\[
 d(\vec{x}, \vec{y}) = \sqrt{d_{V_1}(x_1, y_1)^2 + \cdots + d_{V_k}(x_k, y_k)^2},
\]
where \( x_i \) and \( y_i \) represent the \( i \)th coordinate of \( \vec{x} \) and \( \vec{y} \), respectively.

**Lemma 4.1.3.** Let \( a, b, c, d \geq 0 \). Then 
\[
 \sqrt{a^2 + c^2} + \sqrt{b^2 + d^2} \geq \sqrt{(a + b)^2 + (c + d)^2}.
\]

**Proof.** Let \( a, b, c, d \in \mathbb{R} \) such that \( a, b, c, d \geq 0 \). We have that 
\[
 (ad - cb)^2 = (ad)^2 - 2abcd + (cb)^2 \geq 0.
\]
That is, 
\[
 (ad)^2 + (cb)^2 \geq 2abcd \tag{4.1}
\]
or 
\[
 (ab)^2 + (ad)^2 + (cb)^2 + (cd)^2 \geq (ab)^2 + 2abcd + (cd)^2. \tag{4.2}
\]
Then 
\[
 (a^2 + c^2)(b^2 + d^2) \geq (ab + cd)^2 \tag{4.3}
\]
which implies that 
\[
 2\sqrt{(a^2 + c^2)(b^2 + d^2)} \geq 2(ab + cd). \tag{4.4}
\]
We can add \( a^2 + c^2 \) and \( b^2 + d^2 \) to both sides to get 
\[
 (a^2 + c^2) + 2\sqrt{(a^2 + c^2)(b^2 + d^2)} + (b^2 + d^2) \geq (a^2 + c^2) + 2(ab + cd) + (b^2 + d^2) \tag{4.5}
\]
which implies that

$$(\sqrt{a^2 + c^2} + \sqrt{b^2 + d^2})^2 \geq (a + b)^2 + (c + d)^2.$$  \hfill (4.6)

Thus,

$$\sqrt{a^2 + c^2} + \sqrt{b^2 + d^2} \geq \sqrt{(a + b)^2 + (c + d)^2}.$$  \hfill (4.7)

**Theorem 4.1.4.** The product metric as given in Definition 4.1.2 is a metric.

Note that since $\prod_{i=1}^{k} V_i = V_1 \times (V_2 \times \cdots \times V_k)$, it is sufficient to prove Theorem 4.1.4 for the case $k = 2$.

**Proof.** Let $V_1 \times V_2$ be a direct product of metric spaces, each with metric $d_{V_1}$ and $d_{V_2}$, respectively. Let $(x_1, x_2), (y_1, y_2) \in V_1 \times V_2$. Let $d = d_{V_1 \times V_2}$ be given by

$$d((x_1, x_2), (y_1, y_2)) = \sqrt{d_{V_1}(x_1, y_1)^2 + d_{V_2}(x_2, y_2)^2}.$$

Since $d_{V_i}(x_i, y_i) \geq 0$ for $i = 1, 2$, by definition of metric space, it follows that $d((x_1, x_2), (y_1, y_2)) \geq 0$. Additionally,

$$d((x_1, x_2), (y_1, y_2)) = 0 \iff \sqrt{d_{V_1}(x_1, y_1)^2 + d_{V_2}(x_2, y_2)^2} = 0$$

$$\iff d_{V_1}(x_1, y_1) = 0 \quad \text{and} \quad d_{V_2}(x_2, y_2) = 0$$

$$\iff x_1 = y_1 \quad \text{and} \quad x_2 = y_2$$

$$\iff (x_1, x_2) = (y_1, y_2).$$

Furthermore, $d_{V_i}(x_i, y_i) = d_{V_i}(y_i, x_i)$ for $i = 1, 2$. Thus,

$$d((x_1, x_2), (y_1, y_2)) = \sqrt{d_{V_1}(x_1, y_1)^2 + d_{V_2}(x_2, y_2)^2}$$

$$= \sqrt{d_{V_1}(y_1, x_1)^2 + d_{V_2}(y_2, x_2)^2}$$

$$= d((y_1, y_2), (x_1, x_2)).$$  \hfill (4.8)
Finally, let \((v_1, v_2) \in V_1 \times V_2\). Then \(d_{V_i}(x_i, y_i) \leq d_{V_i}(x_i, v_i) + d_{V_i}(v_i, y_i)\) for \(i = 1, 2\), and \(d_{V_i}(x_i, y_i)^2 \leq (d_{V_i}(x_i, v_i) + d_{V_i}(v_i, y_i))^2\). It follows that

\[
\begin{align*}
  d((x_1, x_2), (y_1, y_2)) &= \sqrt{d_{V_1}(x_1, y_1)^2 + d_{V_2}(x_2, y_2)^2} \\
  &\leq \sqrt{(d_{V_1}(x_1, v_1) + d_{V_1}(v_1, y_1))^2 + (d_{V_2}(x_2, v_2) + d_{V_2}(v_2, y_2))^2} \\
  &\leq \sqrt{d_{V_1}(x_1, v_1)^2 + d_{V_2}(x_2, v_2)^2} + \sqrt{d_{V_1}(v_1, y_1)^2 + d_{V_2}(v_2, y_2)^2} \quad (*) \\
  &= d((x_1, x_2), (v_1, v_2)) + d((v_1, v_2), (y_1, y_2)),
\end{align*}
\]

where (*) follows by Lemma 4.1.3. \(\square\)

Thus, the direct product of metric spaces is itself a metric space.

**Definition 4.1.5.** Let \(X\) and \(Y\) be metric spaces, and let \(f : (X, d_X) \rightarrow (Y, d_Y)\) be a function. Then \(f\) is an isometry if for all \(x_1, x_2 \in X\), we have \(d_X(x_1, x_2) = d_Y(f(x_1), f(x_2))\). If \(f\) is also bijective, then the metric spaces \(X\) and \(Y\) are said to be isometric.

**Lemma 4.1.6.** Let \(f : X \rightarrow Y\) be an isometry of metric spaces, each with metric \(d_X, d_Y\), respectively. If \(f\) is bijective, then \(f^{-1} : Y \rightarrow X\) is also an isometry.

**Proof.** Suppose \(y_1, y_2 \in Y\). Then since \(f\) is bijective, there exists \(x_1, x_2 \in X\) such that \(f(x_1) = y_1\) and \(f(x_2) = y_2\). Furthermore,

\[
\begin{align*}
  d_Y(y_1, y_2) &= d_Y(f(x_1), f(x_2)) \\
  &= d_X(x_1, x_2) \quad (*) \\
  &= d_X(f^{-1}(y_1), f^{-1}(y_2)) \quad (**),
\end{align*}
\]

where (*) follows since \(f\) is an isometry, and (**) follows by the fact that \(f^{-1}(y_1) = x_1\) and \(f^{-1}(y_2) = x_2\). \(\square\)

**Theorem 4.1.7.** The metric spaces \(\mathbb{C}\) and \(\mathbb{R}^2\) are isometric.
Proof. Let $d_{\mathbb{R}}$ and $d_{\mathbb{C}}$ be the absolute value metric defined above, and let $d_{\mathbb{R}^2}$ be the product metric. Define $f : \mathbb{C} \mapsto \mathbb{R}^2$ by $f(x + yi) = (x, y)$. Let $z_1, z_2 \in \mathbb{C}$ where $z_1 = x_1 + y_1i$ and $z_2 = x_2 + y_2i$. Then

$$d_{\mathbb{C}}(z_1, z_2) = |z_1 - z_2|$$

$$=|(x_1 + y_1i) - (x_2 + y_2i)|$$

$$=|(x_1 - x_2) + (y_1 - y_2)i|$$

$$=\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

$$=d_{\mathbb{R}^2}((x_1, y_1), (x_2, y_2)). \quad \square$$

For Definitions 4.1.8–4.1.13 and Theorem 4.1.14, let $(S, d)$ be a metric space and $E \subseteq S$.

**Definition 4.1.8.** Let $\epsilon > 0$. The $\epsilon$-neighborhood of a point $p \in S$, denoted $N_\epsilon(p)$, is the set of all points $q \in S$ such that $d(p, q) < \epsilon$.

**Definition 4.1.9.** A point $p$ is said to be a limit point of $E$ if every neighborhood of $p$ contains a point $q \in E$ such that $p \neq q$. Note that a limit point may or may not be in $E$. If $p \in E$ is not a limit point of $E$, then $p$ is called an isolated point of $E$.

**Definition 4.1.10.** A set $E$ is closed if every limit point of $E$ is a point of $E$. In other words, $E$ is closed if it contains all of its limit points.

**Definition 4.1.11.** Let $E'$ denote the set of all limit points of $E$ in $S$. The closure of $E$, denoted $\overline{E}$, is the set $E \cup E'$.

**Definition 4.1.12.** A point $p$ is an interior point of a set $E$ if there exists a neighborhood $N_\epsilon(p)$ of $p$ such that $N_\epsilon(p) \subseteq E$.

**Definition 4.1.13.** A set $E$ is open if every point of $E$ is an interior point of $E$. 
Theorem 4.1.14. If $p$ is a limit point of $E$, then every neighborhood of $p$ contains infinitely many elements of $E$.

The following proof refers to Sprecher [Spr87, Theorem 17.2].

Proof. Let $p$ be a limit point of $E$ and let $d$ be a metric on $E$. Suppose there exists $\epsilon > 0$ such that $N_\epsilon(p) \cap E$ is a finite set. Then the set $\{d(p, q) : q \in N_\epsilon(p) \cap E, q \neq p\}$ has a minimum. In other words, there exists a point $q' \in N_\epsilon(p) \cap E$ such that $q' \neq p$ and $d(p, q') \leq d(p, q)$ for all $q \in E$. Let $\delta < d(p, q')$. Then $N_\delta(p)$ will not contain any point of $E$ that is different from $p$. This contradicts the fact that $p$ is a limit point. Thus, every neighborhood of $p$ must contain infinitely many points of $E$. \hfill \Box

Theorem 4.1.15. The point $p \in E$ is an isolated point of $E$ if and only if there exists a neighborhood, $N_\epsilon(p)$, such that $N_\epsilon(p) \cap (E \{p\}) = \emptyset$.

Proof. Suppose $p \in E$ is an isolated point of $E$. Then by negation of the definition of limit point, there exists a neighborhood, $N_\epsilon(p)$, such that $N_\epsilon(p) \cap (E \{p\}) = \emptyset$. On the other had, suppose $p \in E$ and there exists a neighborhood, $N_\epsilon(p)$, such that $N_\epsilon(p) \cap (E \{p\}) = \emptyset$. Then by Theorem 4.1.14, $p$ is not a limit point of $E$ and as such, since $p \in E$, $p$ must be an isolated point of $E$. \hfill \Box

4.1.2 Sequences, Limits, and Convergence

Definition 4.1.16. Let $A$ be a set. A sequence in $A$ is a function $f : \mathbb{N} \to A$ given by $f(n) = x_n$ for all $n \in \mathbb{N}$. The sequence $f$ is denoted by $\{x_n\}$.

Definition 4.1.17. A sequence $\{x_n\}$ in a metric space $S$ is said to converge to $x \in S$, if for every $\epsilon > 0$, there exists $N \in \mathbb{R}$ such that for $n \geq N$, $d(x_n, x) < \epsilon$. We write $x_n \to x$ or $\lim_{n \to \infty} x_n = x$. 


Theorem 4.1.18. Let $S$ be a metric space and $E \subseteq S$. Then the following are equivalent:

(1) $S \setminus E$ is open.

(2) Every convergent sequence in $E$ converges to a point in $E$.

(3) $E$ is closed.

Proof. (1) $\Rightarrow$ (2) Let $\{x_n\}$ be a sequence in $E$ such that $x_n \to x \in S$. Suppose that $x \notin E$. Then $x \in S \setminus E$. Since $S \setminus E$ is open, it follows that there exists $\epsilon > 0$ such that $N_\epsilon(x) \subseteq S \setminus E$. Now since $x_n \to x$, it follows that there exists $N$ such that for $n \geq N$, $|x_n - x| < \epsilon$. Thus, $x_N \in N_\epsilon(x)$ which implies that $x_N \in S \setminus E$. This contradicts the assumption that $x_n \in E$ (for all $n$). Thus, we must have that $x \in E$.

(2) $\Rightarrow$ (3) Suppose $x$ is a limit point of $E$. Then every $\epsilon$-neighborhood of $x$ contains an element of $E$ that is different from $x$. We can construct a sequence, $\{x_n\}$ in $E$, that converges to $x$ in the following manner. Let $x_1$ be an element of $N_\epsilon(x)$ such that $x_1 \neq x$. Choose $x_2 \in N_{1/2}(x)$ such that $x_2 \neq x$. Continue in this manner, letting $x_n \in N_{1/n}(x)$ and $x_n \neq x$. By the Archimedean Property, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $N > 1/\epsilon$ and thus, $\epsilon > 1/N$. Thus, for $n \geq N$, we have $x_n \in N_{1/N}(x)$ which implies that $|x_n - x| < 1/N < \epsilon$. So, $x_n \to x$. Since every convergent sequence in $E$ converges to a point in $E$, it follows that $x \in E$.

(3) $\Rightarrow$ (1) Suppose $E$ is closed. Suppose further that $x \in S \setminus E$. Then $x$ is not a limit point of $E$ (because $E$ contains all of its limit points). So, there exists an $\epsilon$-neighborhood $N_\epsilon(x)$ such that $N_\epsilon(x) \cap E = \emptyset$. Thus, $N_\epsilon(x) \subseteq S \setminus E$, and it follows that $S \setminus E$ is open. \[\square\]

Theorem 4.1.19. Let $S$ be a metric space and let $E \subseteq S$. Then the closure of $E$, $\overline{E}$, is closed.
The following proof refers to Rudin [Rud76, Theorem 2.27].

Proof. Let $E'$ denote the set of all limit points of $E$. Then by definition, $\overline{E} = E \cup E'$. Suppose $p \in S/\overline{E}$. Then $p \notin E$ and $p$ is not a limit point of $E$. Thus, there exists $\epsilon > 0$ and a neighborhood $N_\epsilon(p)$ that does not intersect $E$ and therefore does not intersect $E'$. (If $N_\epsilon(p)$ intersected $E'$, then it would have to contain elements of $E$ as every neighborhood of every point of $E'$ contains infinitely many elements of $E$.) Thus, $N_\epsilon(p) \subseteq S/\overline{E}$. Since $S/\overline{E}$ is open, by Theorem 4.1.18, it follows that $\overline{E}$ is closed. □

Theorem 4.1.20. Let \( \{x_n\} \) and \( \{y_n\} \) be convergent sequences in \( \mathbb{R} \), and let \( \lim_{n \to \infty} x_n = x \) and \( \lim_{n \to \infty} y_n = y \). If \( x_n \leq y_n \) for all \( n \), it follows that \( x \leq y \).

Proof. See Bartle and Sherbet [BS00, Theorem 3.2.5]. □

Theorem 4.1.21. Let \( \{x_n\} \) be a sequence in a metric space \( S \). Then \( \{x_n\} \) converges to \( x \in S \) if and only if \( \lim_{n \to \infty} d(x_n, x) = 0 \).

Proof. We have

\[
x_n \to x \iff \text{for every } \epsilon > 0, \text{ there exists } N \text{ such that for } n \geq N, d(x_n, x) < \epsilon
\]

\[
\iff \text{for every } \epsilon > 0, \text{ there exists } N \text{ such that for } n \geq N, |d(x_n, x)| < \epsilon
\]

\[
\iff \lim_{n \to \infty} d(x_n, x) = 0. \quad \Box
\]

Definition 4.1.22. Let \( \{x_n\} \) be a sequence in \( \mathbb{R} \). Then the limit inferior of \( \{x_n\} \), denoted \( \liminf_{n \to \infty} x_n \), is defined to be \( \sup\{\inf\{x_k : k \geq n\} : n > 0\} \). Similarly, the limit superior of \( \{x_n\} \), denoted \( \limsup_{n \to \infty} x_n \), is defined to be \( \inf\{\sup\{x_k : k \geq n\} : n > 0\} \).

Note that \( \liminf_{n \to \infty} x_n \) and \( \limsup_{n \to \infty} x_n \) always exist in \( \mathbb{R} \cup \{\pm \infty\} \), and the relationship \( \liminf_{n \to \infty} x_n \leq \limsup_{n \to \infty} x_n \) always holds.
Theorem 4.1.23 (Monotone Convergence Theorem). A monotone sequence in \( \mathbb{R} \) is convergent if and only if it is bounded. Furthermore, a bounded, increasing sequence in \( \mathbb{R} \) converges to its supremum, while a bounded, decreasing sequence in \( \mathbb{R} \) converges to its infimum.

*Proof.* See Bartle and Sherbet [BS00, Theorem 3.3.2]. \( \square \)

Definition 4.1.24. A sequence \( \{x_n\} \) in a metric space \( S \) is said to be a **Cauchy sequence** if for every \( \epsilon > 0 \), there exists \( N \in \mathbb{R} \) such that for \( n, m \geq N \), \( d(x_n, x_m) < \epsilon \).

Lemma 4.1.25. Let \( f : X \to Y \) be a bijective isometry of metric spaces, each with metric \( d_X, d_Y \), respectively. If \( \{y_n\} \) is a Cauchy sequence in \( Y \), then \( \{f^{-1}(y_n)\} \) is a Cauchy sequence in \( X \), and if \( x_n \to x \), then \( f(x_n) \to f(x) \).

*Proof.* First of all, since \( f \) is bijective, by Lemma 4.1.6 we have that \( f^{-1} : Y \to X \) is an isometry. Let \( \epsilon > 0 \). Then since \( \{y_n\} \) is a Cauchy sequence in \( Y \), there exists \( N \) such that for \( n, m \geq N \),

\[
    d_Y(y_n, y_m) = d_X(f^{-1}(y_n), f^{-1}(y_m)) < \epsilon. \tag{4.12}
\]

It follows that \( \{f^{-1}(y_n)\} \) is a Cauchy sequence.

Suppose that \( \{x_n\} \) is a sequence in \( X \) such that \( x_n \to x \). Let \( \epsilon > 0 \). Then there exists \( N \) such that for \( n \geq N \), we have \( d_X(x_n, x) = d_Y(f(x_n), f(x)) < \epsilon \). It follows that \( f(x_n) \to f(x) \). \( \square \)

Theorem 4.1.26. Let \( \{x_n\} \) be a convergent sequence in a metric space \( S \). Then \( \{x_n\} \) is a Cauchy sequence.

*Proof.* Suppose \( \{x_n\} \) converges to \( x \in S \). Then for \( \epsilon > 0 \), there exists \( N \) such that \( d(x_n, x) < \epsilon/2 \) for \( n \geq N \). For \( n, m \geq N \) we have

\[
    d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon/2 + \epsilon/2 = \epsilon. \tag{4.13}
\]

\( \square \)
Lemma 4.1.27. Suppose \( \{x_n\} \) is a Cauchy sequence in a metric space \( S \) with a convergent subsequence \( \{x_{n_k}\} \) such that \( \lim_{k \to \infty} x_{n_k} = x \). Then \( \lim_{n \to \infty} x_n = x \).

Proof. Let \( \epsilon > 0 \). Then there exists \( N_1 \) such that \( d(x_n, x_m) < \epsilon/2 \) for \( n, m \geq N_1 \), and there exists \( N_2 \) such that \( d(x_{n_k}, x) < \epsilon/2 \) for \( k \geq N_2 \). Let \( N = \max\{N_1, N_2\} \). Then for \( n, k \geq N \), since \( n \leq k \), it follows that

\[
d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \epsilon/2 + \epsilon/2 = \epsilon.
\] (4.14)

Definition 4.1.28. A metric space in which every Cauchy sequence converges is called a complete metric space.

Theorem 4.1.29. Let \( S \) be a complete metric space, and let \( E \subseteq S \). Then \( E \) is a closed subspace of \( S \) if and only if \( E \) is a complete metric space.

Proof. First of all, suppose that \( E \) is a closed subspace of \( S \). By the example after Definition 4.1.1, since \( E \subseteq S \), it follows that \( E \) is a metric space itself with the same metric as \( S \). Suppose \( \{x_n\} \) is a Cauchy sequence in \( E \). Then since \( \{x_n\} \) is in \( S \), and \( S \) is complete, we must have that \( \{x_n\} \) converges to some \( x \in S \). By Theorem 4.1.18 (2), since \( \{x_n\} \) is a convergent sequence in \( E \), it must converge to \( x \in E \). It follows that \( E \) is complete.

On the other hand, suppose that \( E \) is a complete metric space. Let \( \{x_n\} \) be in \( E \) such that \( x_n \to x \) for some \( x \in S \). Then by Theorem 4.1.26, since \( E \) is a metric space, we must have that \( \{x_n\} \) is a Cauchy sequence. Since \( E \) is complete, we must have that \( x \in E \). Thus, \( E \) is closed. \( \square \)

Theorem 4.1.30. The reals, \( \mathbb{R} \), are a complete metric space.

Proof. See Rudin [Rud76, Theorem 3.11]. \( \square \)
Theorem 4.1.31. Let $X$ and $Y$ be complete metric spaces. Then, the direct product $X \times Y$, is also a complete metric space.

Proof. Let $d$ be the product metric on $X \times Y$ and let $(x_n, y_n)$ be a Cauchy sequence in $X \times Y$. Then for every $\epsilon > 0$, there exists $N$ such that for $m, n \geq N$, $d((x_n, y_n), (x_m, y_m)) < \epsilon$. It follows that $d_X(x_n, x_m)^2 + d_Y(y_n, y_m)^2 < \epsilon^2$ which implies that $d_X(x_n, x_m), d_Y(y_n, y_m) < \epsilon$. It follows that $\{x_n\}$ is a Cauchy sequence in $X$ and $\{y_n\}$ is a Cauchy sequence in $Y$. Since $X$ and $Y$ are complete metric spaces, we must have that $x_n \to x$ and $y_n \to y$ for some $x \in X$ and $y \in Y$.

We claim that $(x_n, y_n)$ converges to $(x, y)$. Let $\epsilon > 0$. Then there exists $N_1$ such that for $n \geq N_1$, $d_X(x_n, x) < \epsilon/2$. Similarly, there exists $N_2$ such that for $n \geq N_2$, $d_Y(y_n, y) < \epsilon/2$. Let $N = \max\{N_1, N_2\}$. Then for $n \geq N$, we have

\[
d((x_n, y_n), (x, y)) \leq d((x_n, y_n), (x_n, y)) + d((x_n, y), (x, y))
\]

\[
= \sqrt{d_X(x_n, x_n)^2 + d_Y(y_n, y)^2} + \sqrt{d_X(x_n, x)^2 + d_Y(y, y)^2}
\]

\[
= d_Y(y_n, y) + d_X(x_n, x)
\]

\[
< \epsilon/2 + \epsilon/2 = \epsilon. \quad \Box
\]

Corollary 4.1.32. For any $k \in \mathbb{N}$, $\mathbb{R}^k$ is a complete metric space. \hfill \Box

Corollary 4.1.33. For any $k \in \mathbb{N}$, $\mathbb{C}^k$ is a complete metric space. \hfill \Box

Proof. The result follows by Theorems 4.1.7, 4.1.31, and Corollary 4.1.32. \hfill \Box

Corollary 4.1.34. [Cauchy Criterion] A sequence in $\mathbb{R}^k$ (or $\mathbb{C}^k$) converges if and only if it is a Cauchy sequence.

4.1.3 Series

Definition 4.1.35. Let $\{a_i\}$ be a sequence in $\mathbb{C}$ and let $s_k = \sum_{i=1}^{k} a_i$ denote the $k$th partial sum of the sequence. Then the series $\sum_{i=1}^{\infty} a_i$ is said to be convergent if
\[ \lim_{k \to \infty} s_k \text{ exists in } \mathbb{C}. \] A series \( \sum_{i=1}^{\infty} a_i \) is **absolutely convergent** if \( \sum_{i=1}^{\infty} |a_i| \) is convergent.

**Lemma 4.1.36.** Let \( \sum_{i=1}^{\infty} a_i, \sum_{i=1}^{\infty} b_i \) be convergent, complex-valued series. Then

1. \( \sum_{i=1}^{\infty} ca_i \) converges and \( \sum_{i=1}^{\infty} ca_i = c \sum_{i=1}^{\infty} a_i \),

2. \( \sum_{i=1}^{\infty} (a_i + b_i) \) converges and \( \sum_{i=1}^{\infty} (a_i + b_i) = \sum_{i=1}^{\infty} a_i + \sum_{i=1}^{\infty} b_i \).

**Proof.** See Rudin [Rud76, Theorem 3.47]. \( \square \)

**Lemma 4.1.37.** Let \( \sum_{i=1}^{\infty} a_i, \sum_{i=1}^{\infty} b_i \) be convergent, real-valued series. If \( a_i \leq b_i \) for all \( i \), then \( \sum_{i=1}^{\infty} a_i \leq \sum_{i=1}^{\infty} b_i \).

**Proof.** Let \( s_k = \sum_{i=1}^{k} a_i \) and let \( t_k = \sum_{i=1}^{k} b_i \) for all \( i \). Let \( \sum_{i=1}^{\infty} a_i = \lim_{k \to \infty} s_k = S \) and \( \sum_{i=1}^{\infty} b_i = \lim_{k \to \infty} t_k = T \). Then since \( a_i \leq b_i \) for all \( i \), it follows that \( s_k \leq t_k \) for all \( k \). By Theorem 4.1.20, it follows that \( S \leq T \). \( \square \)

**Definition 4.1.38.** Let \( k_i \) be a bijection from \( \mathbb{N} \) to \( \mathbb{N} \) given by \( i \mapsto k_i \). If we let \( a_i' = a_{k_i} \), then \( \sum_{i=1}^{\infty} a_i' \) is a called a **rearrangement** of \( \sum_{i=1}^{\infty} a_i \).

**Theorem 4.1.39.** Let \( \sum_{i=1}^{\infty} a_i \) be an absolutely convergent series. Then every rearrangement of \( \sum_{i=1}^{\infty} a_i \) converges to the same sum.

**Proof.** See Rudin [Rud76, Theorem 3.55]. \( \square \)

### 4.1.4 Continuity

The following definition refers to Rudin [Rud76, Definition 4.5].
Definition 4.1.40. Suppose that $X$ and $Y$ are metric spaces with metrics $d_X$ and $d_Y$, respectively, and that $f : X \to Y$ is a function. Let $c \in X$. Then $f$ is said to be **continuous at $c$** if for every $\epsilon > 0$, there exists $\delta > 0$ such that if $x \in X$ satisfies $d_X(x, c) < \delta$, then $d_Y(f(x), f(c)) < \epsilon$. If $f$ is continuous at every point of $X$, then $f$ is said to be **continuous on $X$**. The function $f$ is said to be **uniformly continuous** if for every $\epsilon > 0$, there exists $\delta > 0$ such that if $x, y \in X$ satisfies $d_X(x, y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$.

**Theorem 4.1.41.** If $f : X \to Y$ is uniformly continuous, then $f$ is continuous on $X$.

*Proof.* See Sprecher [Spr87, Chapter 6, section 27]. \qed

**Theorem 4.1.42.** Let $X$ and $Y$ be metric spaces and let $f : X \to Y$ be a function. Then $f$ is continuous at a point $c \in X$ if and only if for every sequence $\{x_n\}$ in $X$ that converges to $c$, the sequence $f(x_n)$ converges to $f(c)$, i.e.,

$$
\lim_{n \to \infty} f(x_n) = f \left( \lim_{n \to \infty} x_n \right),
$$

for all convergent sequences $\{x_n\}$ in $X$.

*Proof.* Suppose $f$ is continuous at the point $c \in X$. Let $\epsilon > 0$. Then there exists $\delta > 0$ such that for $x \in X$, $d(f(x), f(c)) < \epsilon$ whenever $d(x, c) < \delta$.

Now since $\lim_{n \to \infty} x_n = c$, we know there exists $N$ such that $d(x_n, c) < \delta$ for $n \geq N$. So, for $n \geq N$, we have $d(x_n, c) < \delta$ which implies that $d(f(x_n), f(c)) < \epsilon$.

On the other hand, suppose that $f$ is not continuous at $c$. Then there exists $\epsilon > 0$ such that for every $\delta > 0$, there exists $x \in X$ such that $d_X(x, c) < \delta$, and $d_Y(f(x), f(c)) \geq \epsilon$.

For $n \in \mathbb{N}$, fix $\delta(n) = 1/n$. Then, by the discontinuity of $f$ at $c$, for every $n \in \mathbb{N}$, there exists $x_n \in X$ such that $d_X(x_n, c) < 1/n$ with $d_Y(f(x_n), f(c)) \geq \epsilon$. Thus, $\{x_n\}$
is a sequence in $X$ such that $x_n$ converges to $c$ (see Theorem 3.1.19), but $f(x_n)$ does not converge to $f(c)$. The result follows.

4.2 Topological Vector Spaces

The following definition is adapted from Horn and Johnson [HJ85, Definition 5.1.3].

**Definition 4.2.1.** Let $V$ be a vector space over $\mathbb{C}$. Let $a, b \in V$. Then an inner product on $V$ is a function from $V \times V$ to $\mathbb{C}$, given by $(a, b) \mapsto \langle a, b \rangle$, such that for all $a, b, c \in V$ and $k \in \mathbb{C}$, we have:

1. $\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle$;
2. $\langle ka, b \rangle = k \langle a, b \rangle$;
3. $\langle a, b \rangle = \overline{\langle b, a \rangle}$; and
4. $\langle a, a \rangle \geq 0$, where $\langle a, a \rangle = 0$ if and only if $a = 0$.

Note that $\langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$, $\langle a, kb \rangle = \overline{k} \langle a, b \rangle$, and $\langle a, b \rangle + \overline{\langle a, b \rangle} = 2 \Re \langle a, b \rangle$, where $\Re \langle a, b \rangle$ represents the real part of the inner product $\langle a, b \rangle$. Furthermore, $\langle a, a \rangle \in \mathbb{R}$ because of property (3). In other words, (3) implies that $\langle a, a \rangle = \overline{\langle a, a \rangle}$, and since $\langle a, a \rangle$ is equal to its conjugate, it must be a real number. Finally, it can be shown via induction that $\left\langle \sum_{i=1}^{n} a_i, c \right\rangle = \sum_{i=1}^{n} \langle a_i, c \rangle$ and $\left\langle a, \sum_{i=1}^{n} b_i \right\rangle = \sum_{i=1}^{n} \langle a, b_i \rangle$.

**Definition 4.2.2.** A vector space $V$ over $\mathbb{C}$ that is assigned a specific inner product is called an inner product space.

Note that a subspace of an inner product space is itself an inner product space.

**Lemma 4.2.3.** Let $V$ be an inner product space. Then for $a, b \in V$, we have

$$\langle a + b, a + b \rangle = \langle a, a \rangle + 2 \Re \langle a, b \rangle + \langle b, b \rangle. \quad (4.17)$$
Proof. By Definition 4.2.1, we have

\[
(a + b, a + b) = (a, a + b) + (b, a + b) \\
= (a, a) + (a, b) + (b, a) + (b, b) \\
= (a, a) + (a, b) + (a, b) + (b, b) \\
= (a, a) + (a, b) + (b, b) + (b, b)
\]

(4.18)

\[
= (a, a) + 2R (a, b) + (b, b) .
\]

\[\Box\]

Definition 4.2.4. Let \( V \) and \( W \) be inner product spaces, and let \( f : V \rightarrow W \) be a linear function. Then \( f \) is an **isometry** if \( (a, b) = (f(a), f(b)) \), for all \( a, b \in V \).

Lemma 4.2.5. Let \( V \) be an inner product space. Suppose \( a, b \in V \) such that \( a \) and \( b \) are linearly dependent. Then \( |(a, b)|^2 = (a, a) (b, b) \).

Proof. Suppose that \( a, b \in V \) are linearly dependent. Then, without loss of generality, \( a = cb \) for some \( c \in \mathbb{C} \). Furthermore,

\[
| (a, b) |^2 = | (cb, b) |^2 \\
= |c|^2 | (b, b) |^2 \\
= (cb, cb) (b, b) \\
= (a, a) (b, b) .
\]

(4.19)

\[\Box\]

Theorem 4.2.6 (Cauchy-Schwarz Inequality). Let \( V \) be an inner product space. Then for \( a, b \in V \),

\[
| (a, b) |^2 \leq (a, a) (b, b) ,
\]

(4.20)

with equality holding if and only if \( \{a, b\} \) is linearly dependent.

Note that parts of the following proof are taken from Horn and Johnson [HJ85, Theorem 5.1.4].

Proof. We show by cases that the inequality in (4.20) is sharp when \( a \) and \( b \) are independent, and equality holds when \( a \) and \( b \) are dependent.
If \( \langle a, a \rangle = 0 \) or \( \langle b, b \rangle = 0 \), then by Definition 4.2.1, either \( a = 0 \) or \( b = 0 \), which implies that \{a, b\} is a linearly dependent set. Furthermore, if \( a = 0 \) or \( b = 0 \), then

\[
| \langle a, b \rangle |^2 = 0 = \langle a, a \rangle \langle b, b \rangle.
\]

If \( \langle a, b \rangle = 0 \) with \( a, b \neq 0 \), then

\[
0 = | \langle a, b \rangle |^2 < \langle a, a \rangle \langle b, b \rangle.
\] (4.21)

We claim that \( a \) and \( b \) are linearly independent. Suppose \( a \) and \( b \) are dependent. Then by Lemma 4.2.5, we must have that

\[
| \langle a, b \rangle |^2 = \langle a, a \rangle \langle b, b \rangle.
\]

This contradicts (4.21). Thus, we must have that \( a \) and \( b \) are independent.

If neither \( a, b \), nor \( | \langle a, b \rangle | \) are zero, let \( t \in \mathbb{R} \). For \( a, b \in V \), consider \( p(t) = \langle a + \langle a, b \rangle tb, a + \langle a, b \rangle tb \rangle \). Then by Lemma 4.2.3 and Definition 4.2.1 we have:

\[
0 \leq \langle a + \langle a, b \rangle tb, a + \langle a, b \rangle tb \rangle
\]

\[
= \langle a, a \rangle + 2t \Re \langle a, \langle a, b \rangle tb \rangle + \langle \langle a, b \rangle tb, \langle a, b \rangle tb \rangle
\]

\[
= \langle a, a \rangle + 2t \Re \langle a, b \rangle \langle a, b \rangle + t^2 | \langle a, b \rangle |^2 \langle b, b \rangle
\]

\[
= \langle a, a \rangle + 2t | \langle a, b \rangle |^2 + t^2 | \langle a, b \rangle |^2 \langle b, b \rangle.
\] (4.22)

Now since \( p(t) \) is greater than or equal to zero for all \( t \in \mathbb{R} \), it follows that \( p(t) \) can have at most one real root.

If \( p(t) \) has no real roots, then by the quadratic equation, we get

\[
4| \langle a, b \rangle |^4 - 4| \langle a, b \rangle |^2 \langle a, a \rangle \langle b, b \rangle < 0
\] (4.23)

which implies that

\[
| \langle a, b \rangle |^4 < | \langle a, b \rangle |^2 \langle a, a \rangle \langle b, b \rangle
\] (4.24)
or

\[
| \langle a, b \rangle |^2 < \langle a, a \rangle \langle b, b \rangle.
\] (4.25)

By Lemma 4.2.5, we have that \( a \) and \( b \) are linearly independent.
Finally, if $p(t)$ has one real root, then by the quadratic equation, we get

$$4|\langle a, b \rangle|^4 - 4|\langle a, b \rangle|^2 \langle a, a \rangle \langle b, b \rangle = 0 \quad (4.26)$$

and since $\langle a, b \rangle$ is non-zero, we get

$$|\langle a, b \rangle|^2 = \langle a, a \rangle \langle b, b \rangle \quad (4.27)$$

Furthermore, there exists $t \in \mathbb{R}$ such that $p(t) = \langle a + \langle a, b \rangle tb, a + \langle a, b \rangle tb \rangle = 0$. That is, there exists $t \in \mathbb{R}$ such that $a = -\langle a, b \rangle tb$. Thus, $a$ and $b$ are linearly dependent. \qed

**Definition 4.2.7.** Let $V$ be an inner product space. Let $a \in V$. Then the $\ell_2$ norm of $a$, denoted $\|a\|_2$ is defined to be $\|a\|_2 = \sqrt{\langle a, a \rangle}$.

**Lemma 4.2.8 (Triangle Inequality).** Let $V$ be an inner product space. Let $a, b \in V$. Then $\|a + b\|_2 \leq \|a\|_2 + \|b\|_2$.

**Proof.** We have

$$\|a + b\|_2^2 = \langle a + b, a + b \rangle = \|a\|_2^2 + 2\Re \langle a, b \rangle + \|b\|_2^2 \quad (*)$$

$$\leq \|a\|_2^2 + 2|\langle a, b \rangle | + \|b\|_2^2 \quad (4.28)$$

$$\leq \|a\|_2^2 + 2\|a\|_2 \|b\|_2 + \|b\|_2^2 \quad (**)$$

$$= (\|a\|_2 + \|b\|_2)^2,$$

where $(*)$ follows by Lemma 4.2.3 and $(**)$ follows by the Cauchy-Schwarz Inequality. Thus, $\|a + b\|_2 \leq \|a\|_2 + \|b\|_2$. \qed

The following definition is adapted from Royden [Roy88, Chapter 6].

**Definition 4.2.9.** A vector space $V$ over $\mathbb{C}$ is called a **normed linear space** if for every $v \in V$, there exists a non-negative real number $\|v\|$ such that
(1) \( \|cv\| = |c|\|v\| \) for all \( c \in \mathbb{C} \);

(2) \( \|v\| = 0 \) if and only if \( v = 0 \); and

(3) \( \|v + w\| \leq \|v\| + \|w\| \), where \( w \in V \).

**Theorem 4.2.10.** Let \( V \) be an inner product space. Then \( V \) is a normed linear space with norm \( \|v\|_2 \).

*Proof.* We see that

\[
(1) \quad \|cv\|_2 = \sqrt{\langle cv, cv \rangle} = \sqrt{|c|^2 \langle v, v \rangle} = |c| \sqrt{\langle v, v \rangle} = |c| \|v\|_2.
\]

(2) Furthermore,

\[
\|v\|_2 = 0 \iff \sqrt{\langle v, v \rangle} = 0
\]

\[
\iff \langle v, v \rangle = 0
\]

\[
\iff v = 0 \quad \text{(\( (*) \))}
\]

where \((*)\) follows by Property (4) in Definition 4.2.1.

(3) Let \( v, w \in V \). Then \( \|v + w\|_2 \leq \|v\|_2 + \|w\|_2 \) by Lemma 4.2.8. \(\Box\)

**Theorem 4.2.11.** Let \( V \) be an inner product space. Then for \( v, w \in V \), \( d(v, w) = \|v - w\|_2 \) is a metric on \( V \) and therefore normed linear spaces are metric spaces.

*Proof.* Let \( v, w \in V \). Then \( d(v, w) = \|v - w\|_2 = \sqrt{\langle v - w, v - w \rangle} \geq 0 \) by Definition 4.2.1. Furthermore,

\[
d(v, w) = 0 \iff \|v - w\|_2 = 0
\]

\[
\iff v - w = 0 \quad \text{by Definition 4.2.9}
\]

\[
\iff v = w.
\]
Additionally,
\[ d(v, w) = \|v - w\|_2 = \sqrt{\langle v - w, v - w\rangle} = \sqrt{\langle (-1)(w - v), (-1)(w - v)\rangle} \]
\[ = \sqrt{(-1)^2 \langle w - v, w - v\rangle} \]
\[ = \sqrt{\langle w - v, w - v\rangle} \]
\[ = \|w - v\|_2 = d(w, v). \] (4.29)

Finally, let \( x \in V \). Then
\[ d(v, w) = \|v - w\|_2 = \|(v - x) + (x - w)\|_2 \]
\[ \leq \|v - x\|_2 + \|x - w\|_2 = d(v, x) + d(x, w). \] (4.30)

The following definition is taken from Royden [Roy88, Chapter 6].

**Definition 4.2.12.** A normed linear space that is complete as a metric space is called a **Banach space**.

The following definition is adapted from Royden [Roy88, Chapter 10].

**Definition 4.2.13.** A vector space \( V \) is called a **Hilbert space** if it is a complete inner product space (and therefore it is a Banach space).

**Theorem 4.2.14.** Let \( V \) and \( W \) be Hilbert spaces. Then the direct sum of Hilbert spaces \( V \oplus W \) is itself a Hilbert space with inner product
\[ \langle (v_1, w_1), (v_2, w_2) \rangle = \langle v_1, v_2 \rangle_V + \langle w_1, w_2 \rangle_W. \] (4.31)

**Proof.** Let \( V \) and \( W \) be Hilbert spaces. Then by Definition 4.2.13, \( V \) and \( W \) are complete inner product spaces. By Theorems 4.2.10 and 4.2.11, \( V \) and \( W \) are complete metric spaces. By Theorem 4.1.31, it follows that \( V \oplus W \) is a complete metric space. It remains to show that \( V \oplus W \) is an inner product space.
First of all, let \((v_1, w_1), (v_2, w_2), \) and \((v_3, w_3) \in V \oplus W\). Then

\[
\langle (v_1 + v_3, w_1 + w_3), (v_2, w_2) \rangle = \langle v_1 + v_3, v_2 \rangle_V + \langle w_1 + w_3, w_2 \rangle_W
\]

\[
= \langle v_1, v_2 \rangle_V + \langle v_3, v_2 \rangle_V + \langle w_1, w_2 \rangle_W + \langle w_3, w_2 \rangle_W
\]

\[
= \langle v_1, v_2 \rangle_V + \langle w_1, w_2 \rangle_W + \langle v_3, v_2 \rangle_V + \langle w_3, w_2 \rangle_W
\]

\[
= \langle (v_1, w_1), (v_2, w_2) \rangle + \langle (v_3, w_3), (v_2, w_2) \rangle .
\]

Additionally, let \(k \in \mathbb{C}\). Then

\[
\langle k(v_1, w_1), (v_2, w_2) \rangle = \langle kv_1, v_2 \rangle_V + \langle kw_1, w_2 \rangle_W
\]

\[
= k \langle v_1, v_2 \rangle_V + k \langle w_1, w_2 \rangle_W
\]

\[
= k(\langle v_1, v_2 \rangle_V + \langle w_1, w_2 \rangle_W)
\]

\[
= k \langle (v_1, w_1), (v_2, w_2) \rangle .
\]

Furthermore,

\[
\langle (v_1, w_1), (v_2, w_2) \rangle = \langle v_2, v_1 \rangle_V + \langle w_2, w_1 \rangle_W
\]

\[
= \langle v_2, v_1 \rangle_V + \langle w_2, w_1 \rangle_W
\]

\[
= \langle v_2, v_1 \rangle_V + \langle w_2, w_1 \rangle_W
\]

\[
= \langle (v_2, w_2), (v_1, w_1) \rangle .
\]

Finally, since \(\langle v_1, v_1 \rangle_V \geq 0\) and \(\langle w_1, w_1 \rangle_W \geq 0\), it follows that \(\langle (v_1, w_1), (v_1, w_1) \rangle = \langle v_1, v_1 \rangle_V + \langle w_1, w_1 \rangle_W \geq 0\), and

\[
\langle (v_1, w_1), (v_1, w_1) \rangle = 0 \iff \langle v_1, v_1 \rangle_V + \langle w_1, w_1 \rangle_W = 0
\]

\[
\iff \langle v_1, v_1 \rangle_V = 0 \quad \text{and} \quad \langle w_1, w_1 \rangle_W = 0
\]

\[
\iff v_1, w_1 = 0
\]

\[
\iff (v_1, w_1) = (0, 0).
\]

The result follows. \(\square\)
Theorem 4.2.15. Let $V$ be an inner product space. Then $f : V \times V \to \mathbb{C}$ given by $(a, b) \mapsto \langle a, b \rangle$ is continuous on $V \times V$, where the metric $d_{V \times V}$ on $V \times V$ is given by

$$d_{V \times V}((x_1, y_1), (x_2, y_2)) = \sqrt{\|x_1 - x_2\|_2^2 + \|y_1 - y_2\|_2^2}.$$

Proof. Let $\epsilon > 0$. Let $(a, b) \in V \times V$, and let $\delta = \min \left\{ \frac{\epsilon}{3\|b\|_2^2 + 1}, \frac{\epsilon}{3\|a\|_2^2 + 1}, \frac{\epsilon}{3} \right\}$.

Consider $a + x_\epsilon$ and $b + y_\epsilon$ such that $d((a + x_\epsilon, b + y_\epsilon), (a, b)) < \delta$. Then

$$\sqrt{d(a + x_\epsilon, a)^2 + d(b + y_\epsilon, b)^2} < \delta$$

which implies that

$$\sqrt{\|x_\epsilon\|_2^2 + \|y_\epsilon\|_2^2} < \delta$$

or

$$\|x_\epsilon\|_2^2 + \|y_\epsilon\|_2^2 < \delta^2$$

which implies that $\|x_\epsilon\|_2, \|y_\epsilon\|_2 < \delta$.

By the Cauchy-Schwarz Inequality (Theorem 4.2.6), it follows that

$$|\langle x_\epsilon, b \rangle| \leq \sqrt{\langle x_\epsilon, x_\epsilon \rangle} \sqrt{\langle b, b \rangle} = \|x_\epsilon\|_2 \|b\|_2.$$

Using a similar argument, we have $|\langle a, y_\epsilon \rangle| \leq \|a\|_2 \|y_\epsilon\|_2$ and $|\langle x_\epsilon, y_\epsilon \rangle| \leq \|x_\epsilon\|_2 \|y_\epsilon\|_2$. Then

$$|\langle a + x_\epsilon, b + y_\epsilon \rangle - \langle a, b \rangle| = |\langle a, b \rangle + \langle x_\epsilon, b \rangle + \langle a, y_\epsilon \rangle + \langle x_\epsilon, y_\epsilon \rangle - \langle a, b \rangle|$$

$$= |\langle x_\epsilon, b \rangle + \langle a, y_\epsilon \rangle + \langle x_\epsilon, y_\epsilon \rangle|$$

$$\leq |\langle x_\epsilon, b \rangle| + |\langle a, y_\epsilon \rangle| + |\langle x_\epsilon, y_\epsilon \rangle|$$

$$\leq \|x_\epsilon\|_2 \|b\|_2 + \|a\|_2 \|y_\epsilon\|_2 + \|x_\epsilon\|_2 \|y_\epsilon\|_2$$

$$\leq \frac{\epsilon \|b\|_2^2}{3\|b\|_2^2 + 1} + \frac{\|a\|_2 \epsilon}{3\|a\|_2^2 + 1} + \left( \frac{\epsilon}{3} \right)^2 \left( \frac{\epsilon}{3} \right)$$

$$\leq \epsilon/3 + \epsilon/3 + \epsilon/9$$

$$< \epsilon. \quad \square$$
Lemma 4.2.16 (Parallelogram Law). Let \( V \) be an inner product space, and let \( x, y \in V \). Then
\[
\|x + y\|_2^2 + \|x - y\|_2^2 = 2 \|x\|_2^2 + 2 \|y\|_2^2.
\]

Proof. Let \( x, y \in V \). Then
\[
\|x + y\|_2^2 + \|x - y\|_2^2 = (x + y, x + y) + (x - y, x - y)
\]
\[
= (x, x) + 2 \Re \langle x, y \rangle + (y, y)
\]
\[
+ (x, x) - 2 \Re \langle x, y \rangle + (y, y)
\]
\[
= 2 \langle x, x \rangle + 2 \langle y, y \rangle
\]
\[
= 2 \|x\|_2^2 + 2 \|y\|_2^2. \quad \square
\]

4.3 Convergence Theorems

Definition 4.3.1. A numbering of a countably infinite set \( G \) is a bijection from \( \mathbb{N} \) to \( G \), given by \( i \mapsto g_i \) and denoted \( \{g_i\} \).

Definition 4.3.2. Let \( G \) be a countable set. Then the vector space \( \mathbb{C}^G \) has elements \( a : G \to \mathbb{C} \), where the coordinates of \( a \) are denoted \( a(g) \).

Note that the notation in this chapter differs slightly from the notation in chapter 2, where the element \( a \) is denoted as \( (a_g) \), and the coordinates \( a(g) = a_g \).

Theorem 4.3.3. Let \( G \) be a countably infinite set. Let \( b \in \mathbb{C}^G \). If, for some numbering \( \{g_i\} \) of \( G \), \( \sum_{i=1}^\infty |b(g_i)| < \infty \), then \( \sum_{i=1}^\infty b(g_i) \) converges to a finite value and is independent of the numbering of \( G \).

Proof. See Theorem 4.1.39. \( \square \)

Definition 4.3.4. Let \( G \) be a countably infinite set. If for some (and therefore, any) numbering \( \{g_i\} \) of \( G \), we have \( \sum_{i=1}^\infty |b(g_i)| < \infty \), then \( \sum_{g \in G} b(g) = \sum_{i=1}^\infty b(g_i) \).
Corollary 4.3.5. Let $G$ be a countable set. If $b(g) \in \mathbb{R}^+ \cup \{0\} \cup \{\infty\}$ for all $g \in G$, then $\sum_{g \in G} b(g)$ is well-defined in $\mathbb{R}^+ \cup \{0\} \cup \{\infty\}$. \qed

Lemma 4.3.6. Suppose $\{b_n\}$ and $\{c_n\}$ are two sequences such that $b_n \leq c_n$ for all $n$, and $\lim_{n \to \infty} b_n = B$. Then $B \leq \liminf_{n \to \infty} c_n$.

Proof. Let $d_n = \inf\{c_k | k \geq n\}$. Then $\liminf_{n \to \infty} c_n = \sup\{d_n\}$. Since $d_n \leq \liminf_{n \to \infty} c_n$ for all $n$, it is enough to show that for every $\varepsilon > 0$, there exists $N$ such that $B - \varepsilon \leq d_N$ (because if $B - \varepsilon \leq \liminf_{n \to \infty} c_n$ for all $\varepsilon > 0$, then $B \leq \liminf_{n \to \infty} c_n$).

Let $\varepsilon > 0$. Since $\lim_{n \to \infty} b_n = B$, there exists $N$ such that $|b_n - B| < \varepsilon$ for $n \geq N$. Then $B - \varepsilon < b_k \leq c_k$ for all $k \geq N$. So $B - \varepsilon$ is a lower bound of the sequence $\{c_k | k \geq N\}$. Thus, $B - \varepsilon \leq d_N = \inf\{c_k | k \geq N\}$. The result follows. \qed

Theorem 4.3.7 (Fatou's Lemma). Let $G$ be a countable set with numbering $\{g_i\}$, and let $\{a_n\}$ be a non-negative sequence in $\mathbb{R}^G$. Suppose $\lim_{n \to \infty} a_n(g_i) = a(g_i)$ for all $i$. Then $\sum_{i=1}^{\infty} a(g_i) \leq \liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i)$.

Proof. For fixed $k \in \mathbb{N}$, we have $\sum_{i=1}^{k} a(g_i) = \sum_{i=1}^{k} \lim_{n \to \infty} a_n(g_i)$. Since $\lim_{n \to \infty} a_n(g_i) = a(g_i)$ for all $i$ by assumption, and we are dealing with finite sums, it follows that

$$\sum_{i=1}^{k} a(g_i) = \sum_{i=1}^{k} a_n(g_i) = \lim_{n \to \infty} \sum_{i=1}^{k} a_n(g_i).$$ (4.42)

Now, $\sum_{i=1}^{k} a_n(g_i) \leq \sum_{i=1}^{\infty} a_n(g_i)$ for fixed $n$. So, (4.42) and Lemma 4.3.6 imply that $\sum_{i=1}^{k} a(g_i) \leq \liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i)$.

Define $b_k = \sum_{i=1}^{k} a(g_i)$. Then $\{b_k\}$ is an increasing sequence bounded by $\liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i)$. By the Monotone Convergence Theorem (Theorem 4.1.23), it follows that
\( \{b_k\} \) converges to its supremum. In other words, \( \sup \{b_k\} = \sum_{i=1}^{\infty} a(g_i) \). Then, since \( \liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) \) is an upper bound of \( \{b_k\} \), it follows that \( \sum_{i=1}^{\infty} a(g_i) \leq \liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) \). \( \square \)

**Theorem 4.3.8 (Lebesgue Convergence Theorem).** Let \( G \) be a countable set. Suppose \( b(g) \in \mathbb{R}^+ \) for all \( g \in G \) and \( \sum_{g \in G} b(g) < \infty \). Let \( \{a_n\} \) be a non-negative sequence in \( \mathbb{R}^G \) such that \( a_n(g) \leq b(g) \) for all \( g \in G \) and \( \lim_{n \to \infty} a_n(g) = a(g) \) for all \( g \in G \). Then \( \sum_{g \in G} a(g) = \lim_{n \to \infty} \sum_{g \in G} a_n(g) \). In particular, the sums on both sides converge.

Note that parts of the following proof are adapted from Royden [Roy88, Chapter 4, proof of Theorem 16].

**Proof.** Choose a numbering \( \{g_i\} \) for \( G \). First note that since \( a_n(g_i) \leq b(g_i) \) for all \( i \), it follows that \( \sum_{i=1}^{\infty} a_n(g_i) \leq \sum_{i=1}^{\infty} b(g_i) < \infty \). Furthermore, since \( b(g_i) \) is an upper bound for the sequence \( a_n(g_i) \), it follows by Theorem 4.1.20 that \( \lim_{n \to \infty} a_n(g_i) = a(g_i) \leq b(g_i) \) for all \( i \) and thus, \( \sum_{i=1}^{\infty} a(g_i) \leq \sum_{i=1}^{\infty} b(g_i) < \infty \).

Now since \( a_n(g_i) \leq b(g_i) \) for all \( i \), it follows that \( b(g_i) - a_n(g_i) \geq 0 \). By the limit laws, it follows that \( \lim_{n \to \infty} (b(g_i) - a_n(g_i)) = b(g_i) - a(g_i) \).

Thus, we can apply Fatou's Lemma to the sequence \( b(g_i) - a_n(g_i) \) and can conclude that \( \sum_{i=1}^{\infty} (b(g_i) - a(g_i)) \leq \liminf_{n \to \infty} \left( \sum_{i=1}^{\infty} (b(g_i) - a_n(g_i)) \right) \). Since \( \sum_{i=1}^{\infty} b(g_i), \sum_{i=1}^{\infty} a(g_i), \sum_{i=1}^{\infty} a_n(g_i) < \infty \), by Lemma 4.1.36, we get

\[
\sum_{i=1}^{\infty} b(g_i) - \sum_{i=1}^{\infty} a(g_i) \leq \liminf_{n \to \infty} \left( \sum_{i=1}^{\infty} b(g_i) - \sum_{i=1}^{\infty} a_n(g_i) \right) \tag{4.43}
\]

or

\[
\sum_{i=1}^{\infty} b(g_i) - \sum_{i=1}^{\infty} a(g_i) \leq \sum_{i=1}^{\infty} b(g_i) - \limsup_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) \tag{4.44}
\]
Since $\sum_{i=1}^{\infty} b(g_i) < \infty$, it follows that

$$\limsup_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) \leq \sum_{i=1}^{\infty} a(g_i). \quad (4.45)$$

We can apply a similar argument to the sequence $b(g_i) + a_n(g_i) \geq 0$ to get

$$\sum_{i=1}^{\infty} a(g_i) \leq \liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i). \quad (4.46)$$

Combining (4.45) and (4.46), we get

$$\limsup_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) \leq \sum_{i=1}^{\infty} a(g_i) \leq \liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i). \quad (4.47)$$

Since the relationship $\liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) \leq \limsup_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i)$ is always true, we can conclude that

$$\liminf_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) = \limsup_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) = \lim_{n \to \infty} \sum_{i=1}^{\infty} a_n(g_i) = \sum_{i=1}^{\infty} a(g_i). \quad (4.48)$$

\[\square\]

4.4 $\ell^2(G)$

**Definition 4.4.1.** Let $G$ be a countable set and let $g \in G$. Then $\ell^2(G)$ is the set of all elements $a \in \mathbb{C}^G$ such that $\sum_{g \in G} |a(g)|^2 < \infty$ (i.e., $\sum_{g \in G} |a(g)|^2$ converges; see Theorem 4.3.3).

**Lemma 4.4.2.** Let $G$ be a countable set. Let $a, b \in \ell^2(G)$. Then $\sum_{g \in G} |a(g)b(g)|$ converges.

**Proof.** We know that $(|a(g)| - |b(g)|)^2 \geq 0$ for all $a(g), b(g) \in \mathbb{C}$, which implies that

$$|a(g)|^2 - 2|a(g)||b(g)| + |b(g)|^2 \geq 0, \quad (4.49)$$

or

$$|a(g)|^2 + |b(g)|^2 \geq 2|a(g)||b(g)| = 2|a(g)b(g)|. \quad (4.50)$$
Choose a numbering \( \{g_i\} \) for \( G \). Then

\[
\sum_{i=1}^{n} |a(g_i)b(g_i)| \leq \frac{1}{2}(\sum_{i=1}^{n} |a(g_i)|^2 + |b(g_i)|^2)
\]

\[
\leq \sum_{i=1}^{\infty} |a(g_i)|^2 + |b(g_i)|^2
\]

\[
= \sum_{i=1}^{\infty} |a(g_i)|^2 + \sum_{i=1}^{\infty} |b(g_i)|^2 \tag{4.51}
\]

\[
< \infty \quad \text{because } a(g) \in \ell^2(G) \quad \text{and} \quad < \infty \quad \text{because } b(g) \in \ell^2(G)
\]

\[
< \infty.
\]

where (*) follows by Lemma 4.1.36.

So, the sequence of partial sums \( u_n = \sum_{i=1}^{n} |a(g_i)b(g_i)| \) is an increasing, bounded sequence in \( \mathbb{R} \). By Theorem 4.1.23, it follows that \( \lim_{n \to \infty} \sum_{i=1}^{n} |a(g_i)b(g_i)| \) exists. In other words, the series \( \sum_{i=1}^{\infty} |a(g_i)b(g_i)| \) converges.

\[\square\]

**Lemma 4.4.3.** Let \( G \) be a countable set. Then \( \ell^2(G) \) is a subspace of the vector space \( \mathbb{C}^G \).

**Proof.** If \( a(g) = 0 \) for all \( g \in G \), then \( \sum_{g \in G} |a(g)|^2 = 0 \) which implies that \( 0 \in \ell^2(G) \).

Furthermore, let \( a, b \in \ell^2(G) \) and let \( k \in \mathbb{C} \). Then

\[
\sum_{g \in G} |a(g) + b(g)|^2 = \sum_{g \in G} |a(g)|^2 + 2a(g)b(g) + b(g)^2 |
\]

\[
\leq \sum_{g \in G} |a(g)|^2 + 2|a(g)b(g)| + |b(g)|^2
\]

\[
= \sum_{g \in G} |a(g)|^2 + \sum_{g \in G} 2|a(g)b(g)| + \sum_{g \in G} |b(g)|^2 \tag{4.52}
\]

\[
< \infty, \quad (**)
\]
where (*) follows by Lemma 4.1.36 and (**) follows by Lemma 4.4.2 and the fact that \(a, b \in \ell^2(G)\). Additionally,

\[
\sum_{g \in G} |k a(g)|^2 = \sum_{g \in G} |k|^2 |a(g)|^2 \\
= |k|^2 \sum_{g \in G} |a(g)|^2 \quad \text{by Lemma 4.1.36} \quad (4.53)
\]

\(< \infty.

It follows that \(\ell^2(G)\) is a subspace of the vector space \(\mathbb{C}^G\).

**Theorem 4.4.4.** Let \(a, b \in \ell^2(G)\). Then \(\langle a, b \rangle = \sum_{g \in G} a(g)\overline{b(g)}\) converges and is an inner product on \(\ell^2(G)\).

**Proof.** By Lemma 4.4.2, we know that \(\sum_{g \in G} |a(g)\overline{b(g)}|\) converges. By Theorem 4.3.3, it follows that \(\sum_{g \in G} a(g)\overline{b(g)}\) converges.

Suppose \(a \in \ell^2(G)\). Then \(\langle a, a \rangle = \sum_{g \in G} a(g)\overline{a(g)} = \sum_{g \in G} |a(g)|^2 \geq 0\). Furthermore,

\[
\langle a, a \rangle = 0 \iff \sum_{g \in G} a(g)\overline{a(g)} = 0 \\
\iff \sum_{g \in G} |a(g)|^2 = 0 \\
\iff a(g) = 0 \quad \text{for all } g \in G.
\]

Additionally,

\[
\langle a + b, c \rangle = \sum_{g \in G} (a(g) + b(g))\overline{c(g)} \\
= \sum_{g \in G} (a(g)\overline{c(g)} + b(g)\overline{c(g)}) \quad (4.54) \\
= \sum_{g \in G} a(g)\overline{c(g)} + \sum_{g \in G} b(g)\overline{c(g)} \quad (*) \\
= \langle a, c \rangle + \langle b, c \rangle,
\]
where (*) follows by Lemmas 4.1.36 and 4.4.2. Also,
\[
\langle ka, b \rangle = \sum_{g \in G} (ka(g)) b(g) = k \sum_{g \in G} (a(g)) b(g) \quad (*)
\]
\[
= k \langle a, b \rangle,
\]
where (*) follows by Lemmas 4.1.36 and 4.4.2.

Finally,
\[
\langle b, a \rangle = \sum_{g \in G} b(g) \overline{a(g)} = \sum_{g \in G} \overline{b(g)} \overline{a(g)} = \sum_{g \in G} a(g) \overline{b(g)} = \langle a, b \rangle.
\]

**Corollary 4.4.5.** Let \( G \) be a countable set. Then \( \ell^2(G) \) is an inner product space and therefore a normed linear space with norm \( \|a\|_2 \) for \( a \in \ell^2(G) \).

**Proof.** The result follows immediately from Definition 4.2.2, Theorem 4.2.10, and Theorem 4.4.4. □

**Theorem 4.4.6.** Let \( G \) be a countable set. Suppose \( \{a_n\} \) is a Cauchy sequence in \( \ell^2(G) \). Then there exists \( a \in \ell^2(G) \) such that \( \lim_{n \to \infty} \|a_n - a\|_2 = 0 \).

**Proof.** Since \( \{a_n\} \) is a Cauchy sequence in \( \ell^2(G) \), it follows that for every \( \epsilon > 0 \), there exists \( N \) such that for \( n, m \geq N \), \( \|a_n - a_m\|_2 < \epsilon \). By the definition of \( \ell^2 \) norm, it follows that
\[
\sqrt{\sum_{g \in G} |a_n(g) - a_m(g)|^2} < \epsilon
\]
for \( n, m \geq N \), or
\[
\sum_{g \in G} |a_n(g) - a_m(g)|^2 < \epsilon^2
\]
which implies that for fixed \( g \in G \),
\[
|a_n(g) - a_m(g)|^2 < \epsilon^2
\]
for $n, m \geq N$, or

$$|a_n(g) - a_m(g)| < \epsilon,$$  \hspace{1cm} (4.60)

for $n, m \geq N$.

Thus, for fixed $g \in G$, $\{a_n(g)\}$ is a Cauchy sequence in $\mathbb{C}$. By Theorem 4.1.30, it follows that there exists $a(g) \in \mathbb{C}$ such that $\lim_{n \to \infty} a_n(g) = a(g)$. This shows that for all $g \in G$, each sequence of coordinates $\{a_n(g)\}$ converges to some $a(g) \in \mathbb{C}$. Thus, the original sequence $\{a_n\}$ converges to some $a \in \mathbb{C}^G$, where the coordinates of $a$ are given by $a(g) = \lim_{n \to \infty} a_n(g)$.

It remains to show that $a \in \ell^2(G)$ and $\lim_{n \to \infty} \|a_n(g) - a\|_2 = 0$.

Let $\epsilon = 1/2$. Then since $\{a_n\}$ is a Cauchy sequence, there exists $n_1$ such that $\|a_s - a_m\|_2 < 1/2$ for $s, m \geq n_1$. Continuing in this manner, for $k \geq 2$, if we let $\epsilon = 1/2^k$, there exists $n_k > n_{k-1}$ such that $\|a_s - a_m\|_2 < 1/2^k$ for $s, m \geq n_k$. So, $\{a_{n_k}\}$ is a subsequence of $\{a_n\}$.

For $g \in G$, let $b_k(g) = |a_{n_1}(g)| + \sum_{i=2}^{k} |a_{n_i}(g) - a_{n_{i-1}}(g)|$, and let

$$b(g) = |a_{n_1}(g)| + \sum_{i=2}^{\infty} |a_{n_i}(g) - a_{n_{i-1}}(g)|.$$

Now since $|a_{n_i}(g) - a_{n_{i-1}}(g)|^2 \leq \sum_{g \in G} |a_{n_i}(g) - a_{n_{i-1}}(g)|^2$, it follows that $|a_{n_i}(g) - a_{n_{i-1}}(g)| \leq \|a_{n_i} - a_{n_{i-1}}\|_2$. Thus,

$$b(g) \leq |a_{n_1}(g)| + \sum_{i=2}^{\infty} \|a_{n_i} - a_{n_{i-1}}\|_2$$

$$\leq |a_{n_1}(g)| + 1.$$  \hspace{1cm} (4.61)

This shows that $b(g)$ is finite for each $g \in G$ and thus, $\lim_{k \to \infty} b_k(g) = b(g)$. 
Also,

\[ \|b_k\|_2 = \left\| a_{n_1} + \sum_{i=2}^{k} (a_{n_i} - a_{n_{i-1}}) \right\|_2 \leq \|a_{n_1}\|_2 + \sum_{i=2}^{k} \|a_{n_i} - a_{n_{i-1}}\|_2 \]

\[ \leq \|a_{n_1}\|_2 + 1 < \infty. \] (4.62)

This implies that \( \sum_{g \in G} |b_k(g)|^2 \leq (\|a_{n_1}\|_2 + 1)^2 < \infty. \)

Now since \( \lim_{k \to \infty} b_k(g) = b(g) \), it follows that \( \lim_{k \to \infty} |b_k(g)|^2 = |b(g)|^2 \). Applying Fatou’s Lemma (Theorem 4.3.7) to the sequence \( \{b_k^2\} \), we get

\[ \|b\|_2^2 = \sum_{g \in G} |b(g)|^2 \leq \liminf_{k \to \infty} \sum_{g \in G} |b_k(g)|^2 \leq (\|a_{n_1}\|_2 + 1)^2 < \infty. \] (4.63)

Furthermore,

\[ |a_{n_k}(g)| = |a_{n_1}(g) + a_{n_2}(g) - a_{n_1}(g) + a_{n_3}(g) - a_{n_2}(g) + \cdots + a_{n_k}(g) - a_{n_{k-1}}(g)| \]

\[ \leq |a_{n_1}(g)| + |a_{n_2}(g) - a_{n_1}(g)| + |a_{n_3}(g) - a_{n_2}(g)| + \cdots + |a_{n_k}(g) - a_{n_{k-1}}(g)| \]

\[ \leq b(g). \] (4.64)

So, \( |a_{n_k}(g)| \leq b(g) \). By Theorem 4.1.20, it follows that

\[ \lim_{k \to \infty} |a_{n_k}(g)| \leq \lim_{k \to \infty} b(g). \] (4.65)

which implies that \( |a(g)| \leq b(g) \).

By (4.63) and (4.65), it follows that

\[ \|a\|_2^2 = \sum_{g \in G} |a(g)|^2 \leq \sum_{g \in G} |b(g)|^2 < \infty. \] (4.66)

Thus, \( a \in \ell^2(G) \).
We now claim that \( \lim_{i \to \infty} \| a_n - a \|^2 = 0 \). We just showed that \( |a_n(g)| \leq b(g) \) for all \( g \). It follows that \( |a_n(g) - a(g)| \leq |a_n(g)| + |a(g)| \leq 2b(g) \) and therefore

\[
|a_n(g) - a(g)|^2 \leq 4|b(g)|^2. \tag{4.67}
\]

Since \( \lim_{i \to \infty} a_n(g) = a(g) \) by assumption, it follows that

\[
\lim_{i \to \infty} |a_n(g) - a(g)|^2 = 0. \tag{4.68}
\]

Since \( \sum_{g \in G} 4|b(g)|^2 < \infty \), we can apply the Lebesgue Convergence Theorem (Theorem 4.3.8) to the sequence \( |a_n(g) - a(g)|^2 \) to get

\[
\lim_{i \to \infty} \left( \sum_{g \in G} |a_n(g) - a(g)|^2 \right) = \sum_{g \in G} 0 = 0. \tag{4.69}
\]

Now,

\[
\lim_{i \to \infty} \left( \sum_{g \in G} |a_n(g) - a(g)|^2 \right) = 0 \iff \lim_{i \to \infty} \| a_n - a \|^2 = 0
\]

\[
\iff \lim_{i \to \infty} d(a_n, a) = 0
\]

\[
\iff \lim_{i \to \infty} a_n = a,
\]

by Theorem 4.1.21. Since \( \{a_n\} \) is a Cauchy sequence with a convergent subsequence, it follows by Theorem 4.1.27, that \( \lim_{n \to \infty} \| a_n - a \|^2 = 0 \).

**Corollary 4.4.7.** Let \( G \) be a countable set. Then \( \ell^2(G) \) is a Hilbert space.

**Proof.** Lemma 4.4.3 and Corollary 4.4.5 shows that \( \ell^2(G) \) is an inner product space. Theorems 4.4.6 and 4.1.21 show that \( \ell^2(G) \) is complete. \( \square \)

### 4.5 Complete Orthonormal Sets

**Definition 4.5.1.** The Kronecker delta, \( \delta_{ij} \), is defined such that \( \delta_{ij} = 1 \), if \( i = j \) and \( \delta_{ij} = 0 \) if \( i \neq j \).
**Definition 4.5.2.** Let $H$ be a Hilbert space and let $G$ be a countable set. The set \( \{e_g : g \in G\} \subseteq H \) is a complete orthonormal set if

1. For \( i, j \in G \), \( \langle e_i, e_j \rangle = \delta_{ij} \) (See Definition 4.5.1).
2. For \( a \in H \), if \( \langle a, e_i \rangle = 0 \) for all \( i \in G \), then \( a = 0 \).

**Theorem 4.5.3.** Let \( e_g(i) = 1 \) if \( i = g \) and 0 otherwise. Then the set \( \{e_g : g \in G\} \) is a complete orthonormal set in \( \ell^2(G) \).

**Proof.** In \( \ell^2(G) \), \( \langle e_i, e_j \rangle = \sum_{g \in G} e_i(g) \overline{e_j(g)} \). Additionally, \( e_i(g) = 1 \) if and only if \( i = g \), otherwise, \( e_i(g) = 0 \).

If \( i \neq j \), then for all \( g \in G \), either \( e_i(g) \) or \( e_j(g) \) is 0. As a result,

\[
\langle e_i, e_j \rangle = \sum_{g \in G} e_i(g) \overline{e_j(g)} = \sum_{g \in G} 0 = 0 = \delta_{ij}.
\]

If \( i = j \), then

\[
\langle e_i, e_j \rangle = \langle e_i, e_i \rangle = \sum_{g \in G} e_i(g) \overline{e_i(g)} = e_i(i) \overline{e_i(i)} = 1 = \delta_{ij}.
\]

Furthermore, suppose that \( a \in \ell^2(G) \) and that \( \langle a, e_i \rangle = 0 \) for all \( i \in \{e_g : g \in G\} \). Then

\[
\langle a, e_i \rangle = \sum_{g \in G} a(g) \overline{e_i(g)} = \langle a(g) \rangle(1) = a(g) = 0.
\]

Since \( \langle a, e_i \rangle = 0 \) for all \( i \), it follows that \( a(g) = 0 \) for all \( g \in G \), and thus, \( a = 0 \). It follows that \( \{e_g : g \in G\} \) is a complete orthonormal set in \( \ell^2(G) \). \( \square \)

**Theorem 4.5.4.** Let \( H \) be a Hilbert space and \( \{e_n : n \in \mathbb{N}\} \) a complete orthonormal set. Then for \( a \in H \), \( \lim_{n \to \infty} \left\| a - \sum_{i=1}^{n} \langle a, e_i \rangle e_i \right\|_2 = 0 \) and \( \|a\|_2^2 = \sum_{i=1}^{\infty} |\langle a, e_i \rangle| \). In other words, we say that the Fourier series \( \sum_{i=1}^{\infty} \langle a, e_i \rangle e_i \) converges to \( a \) in the \( \ell^2 \) norm.
Proof. By definition,

\[ 0 \leq \left\| a - \sum_{i=1}^{n} \langle a, e_i \rangle e_i \right\|_2^2 = \langle a - \sum_{i=1}^{n} \langle a, e_i \rangle e_i, a - \sum_{j=1}^{n} \langle a, e_j \rangle e_j \rangle \]

\[ = \langle a, a \rangle - 2 \Re \left\langle a, \sum_{i=1}^{n} \langle a, e_i \rangle e_i \right\rangle + \left\langle \sum_{i=1}^{n} \langle a, e_i \rangle e_i, \sum_{j=1}^{n} \langle a, e_j \rangle e_j \right\rangle \]

\[ = \langle a, a \rangle - \sum_{i=1}^{n} 2 \Re \left( \langle a, e_i \rangle \langle a, e_i \rangle \right) + \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \langle a, e_i \rangle, \langle a, e_j \rangle \right) \]

(4.71)

\[ = \langle a, a \rangle - 2 \sum_{i=1}^{n} |\langle a, e_i \rangle|^2 + \sum_{i=1}^{n} |\langle a, e_i \rangle|^2 \]

\[ = \langle a, a \rangle - \sum_{i=1}^{n} |\langle a, e_i \rangle|^2. \]

Thus, we have \( \sum_{i=1}^{n} |\langle a, e_i \rangle|^2 \leq \|a\|_2 < \infty. \) If we let \( s_k = \sum_{i=1}^{k} |\langle a, e_i \rangle|^2, \) then by the Monotone Convergence Theorem (4.1.23), it follows that \( \lim_{k \to \infty} s_k \) exists and thus \( \sum_{i=1}^{\infty} |\langle a, e_i \rangle|^2 \) is a convergent series. Since \( s_k \in \mathbb{R} \) for all \( k \) and \( \{s_k\} \) is convergent, it follows by the Cauchy Criterion (Corollary 4.1.34) that \( \{s_k\} \) is a Cauchy sequence.

Thus, for \( \varepsilon > 0, \) there exists \( N \) such that for \( m \geq n \geq N, \)

\[ |s_n - s_m| = \left| \sum_{i=n+1}^{m} |\langle a, e_i \rangle|^2 \right| < \varepsilon^2. \]  

(4.72)

Let \( z_n = \sum_{i=1}^{n} \langle a, e_i \rangle e_i. \) If \( m \geq n, \) then \( z_m - z_n = \sum_{i=n+1}^{m} \langle a, e_i \rangle e_i, \) and thus,

\[ \|z_m - z_n\|_2^2 = \left( \sum_{i=n+1}^{m} \langle a, e_i \rangle e_i, \sum_{j=n+1}^{m} \langle a, e_j \rangle e_j \right) = \sum_{i=n+1}^{m} |\langle a, e_i \rangle|^2 < \varepsilon^2 \] by (4.72).

It follows that for \( m, n \geq N, \) \( \|z_m - z_n\|_2 < \varepsilon. \) Thus, \( \{z_n\} \) is a Cauchy sequence.

Since \( z_n \in H \) for all \( n, \) it follows by the completeness of \( H \) that there exists \( b \in H \)
such that \( \lim_{n \to \infty} z_n = \sum_{i=1}^{\infty} \langle a, e_i \rangle e_i = b. \)

It remains to show that \( a = b. \) Let \( c = a - b. \) Then

\[
\langle c, e_j \rangle = \langle a - b, e_j \rangle
\]

\[
= \langle a - \lim_{n \to \infty} z_n, e_j \rangle
\]

\[
= \langle a - \lim_{n \to \infty} \sum_{i=1}^{n} \langle a, e_i \rangle e_i, e_j \rangle
\]

\[
= \langle a, e_j \rangle - \left( \lim_{n \to \infty} \sum_{i=1}^{n} \langle a, e_i \rangle e_i, e_j \right) \tag{4.73}
\]

\[
= \langle a, e_j \rangle - \lim_{n \to \infty} \sum_{i=1}^{n} \langle a, e_i \rangle \langle e_i, e_j \rangle \quad (*)
\]

\[
= \langle a, e_j \rangle - \langle a, e_j \rangle = 0,
\]

where (*) follows by Theorems 4.1.42 and 4.2.15. It follows that \( \langle c, e_j \rangle = 0 \) for all \( j. \)

Since \( \{e_n : n \in \mathbb{N}\} \) is a complete orthonormal set, it follows by Definition 4.5.2 that \( c = 0. \) Thus, \( a = b. \)

Finally, we have

\[
\|a\|_2^2 = \left\langle \lim_{n \to \infty} \sum_{i=1}^{n} \langle a, e_i \rangle e_i, \lim_{n \to \infty} \sum_{j=1}^{n} \langle a, e_j \rangle e_j \right\rangle
\]

\[
= \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \langle a, e_i \rangle \langle e_i, e_j \rangle \quad (*)
\]

\[
= \lim_{n \to \infty} \sum_{i=1}^{n} | \langle a, e_i \rangle |^2
\]

\[
= \sum_{i=1}^{\infty} | \langle a, e_i \rangle |^2.
\]

where (*) follows by Theorems 4.1.42 and 4.2.15.

\[\square\]

4.6 Projections

**Definition 4.6.1.** Let \( H \) be a Hilbert space and let \( W \) be a subset of \( H. \) Then

\( W^\perp = \{v \in H : \langle v, w \rangle = 0 \text{ for all } w \in W\}. \)
Definition 4.6.2. Let $H$ be a Hilbert space and let $W$ be a subspace of $H$. Then $W$ is a closed subspace of $H$ if $W$ is closed as a subset of the metric space $H$.

Theorem 4.6.3. Let $W$ be a closed subspace of a Hilbert space $H$. Then $W$ is itself a Hilbert space.

Proof. Since $W$ is a subspace of an inner product space, $W$ is itself an inner product space. (See note after Definition 4.2.2.) Let $\{x_n\}$ be a Cauchy sequence in $W \subseteq H$. Then by the completeness of $H$ it follows that $\{x_n\}$ converges to some value $x \in H$. Since $W$ is closed, we must have that $x \in W$. Thus, $W$ is complete. It follows that $W$ is a Hilbert space. $\square$

Theorem 4.6.4. Let $H$ be a Hilbert space and let $W \subseteq H$. Then $W^\perp$ is a closed subspace of $H$ (and hence a Hilbert space itself).

Proof. We have $0 \in W^\perp$ because $\langle 0, w \rangle = 0$ for all $w \in W$. Additionally, suppose that $x, y \in W^\perp$. Then $\langle x, w \rangle = \langle y, w \rangle = 0$ for all $w \in W$. It follows that

$$\langle x + y, w \rangle = \langle x, w \rangle + \langle y, w \rangle = 0$$

for all $w \in W$. Thus, $x + y \in W^\perp$. Finally, if $x \in W^\perp$, then $\langle cx, w \rangle = c \langle x, w \rangle = c0 = 0$ for all $w \in W$. So, $cx \in W^\perp$.

Suppose that $\{x_n\}$ is a convergent sequence in $W^\perp$. Then $x_n \to x$ for some $x \in H$. Furthermore, for all $w \in W$,

$$\langle x, w \rangle = \langle \lim_{n \to \infty} x_n, w \rangle$$

$$= \lim_{n \to \infty} \langle x_n, w \rangle \quad (\ast)$$

$$= \lim_{n \to \infty} 0 = 0, \quad (\ast\ast)$$

where (\ast) follows by the continuity of inner product (Theorem 4.2.15) and (\ast\ast) follows by the fact that $x_n \in W^\perp$ for all $n$. Since $\langle x, w \rangle = 0$, it follows that $x \in W^\perp$ and hence $W^\perp$ is closed. $\square$
Lemma 4.6.5. Let $W$ be a closed subspace of a Hilbert space $H$. Then $W \cap W^\perp = \{0\}$.

Proof. Since $W$ and $W^\perp$ are both subspaces of $H$, we have that $\{0\} \subseteq W \cap W^\perp$.

Suppose $x \in W \cap W^\perp$. Then $\langle x, w \rangle = 0$ for all $w \in W$. In particular, $\langle x, x \rangle = 0$ which implies that $x = 0$. The result follows. \qed

Note: we use Reed and Simon [RS80, Ch. II] as a reference for the following theorems and proofs.

Lemma 4.6.6. Let $W$ be a closed subspace of a Hilbert space $H$. For $v \in H$, there exists a unique vector $w \in W$ that minimizes $\|v - w\|_2$.

Proof. If $v \in W$, then $\|v - v\|_2 = 0$ and thus $v$ is the unique vector that minimizes the norm.

Otherwise, suppose $v \notin W$. Let $d = \inf\{\|v - w\|_2 : w \in W\}$. Then for every $n \in \mathbb{N}$, there exists $x_n \in W$ such that $d \leq \|v - x_n\|_2 \leq d + 1/n$. Thus, there exists a sequence $\{x_n\}$ in $W$ such that $\lim_{n \to \infty} \|v - x_n\|_2 = d$. Furthermore,

$$
\|x_n - x_m\|_2^2 = \|(x_n - v) - (x_m - v)\|_2^2 \\
= 2 \|(x_n - v)\|_2^2 + 2 \|(x_m - v)\|_2^2 - \|2v + x_n + x_m\|_2^2 \\
= 2 \|(x_n - v)\|_2^2 + 2 \|(x_m - v)\|_2^2 - 4 \|v - (1/2)(x_n + x_m)\|_2^2 \\
\leq 2 \|(x_n - v)\|_2^2 + 2 \|(x_m - v)\|_2^2 - 4d^2
$$

where (*) follows from the parallelogram law and (**) follows from the fact that $\|v - (1/2)(x_n + x_m)\|_2 \geq d$.

Now since $\lim_{n \to \infty} \|v - x_n\|_2 = d$, it follows that for every $\delta > 0$, there exists $N$ such that for $n \geq N$, $\|x_n - v\|_2 - d < \delta$, or in other words, $\|x_n - v\|_2 < d + \delta$. Then for $m \geq N$, we have $\|x_m - v\|_2 < d + \delta$. 
Given $\epsilon > 0$, let $\delta < \min \left\{ 1, \frac{\epsilon}{8d + 4} \right\}$. Then for $m, n \geq N$, we get

$$
\|x_n - x_m\|_2^2 \leq 2\|(x_n - v)\|_2^2 + 2\|(x_m - v)\|_2^2 - 4d^2
< 2(d + \delta)^2 + 2(d + \delta)^2 - 4d^2
= 8d\delta + 4\delta^2
\leq \delta(8d + 4)
< \epsilon.
$$

Thus, $\{x_n\}$ is a Cauchy sequence. By the completeness of $W$, we must have that $x_n \to w$ for some $w \in W$. Thus, there exists $w \in W$ such that $\|v - w\|_2 = d$.

It remains to show that the vector $w$ that minimizes $\|v - x\|_2$ is unique. Suppose there exists two distinct vectors, $w$ and $w'$, such that $\|v - w\|_2 = \|v - w'\|_2 = d$. We first show that $(v - w)$ and $(v - w')$ are linearly independent. Suppose without loss of generality that $v - w = c(v - w')$ for some $c \in \mathbb{C}$. Then

$$
v - w = cv - cw'
$$

which implies that

$$
v - cv = w - cw'
$$

and

$$
v(1 - c) = w - cw'.
$$

If $c = 1$, then $w = w'$. If $c \neq 1$, then $v = \frac{w - cw'}{1 - c}$. Then since $v$ is a linear combination of elements of $W$, we must have that $v \in W$. But this contradicts our assumption that $v \notin W$. It follows that $(v - w)$ and $(v - w')$ must be linearly independent.
We see that
\[
\left\| v - \left( \frac{w + w'}{2} \right) \right\|^2_2 = \langle (v - w/2) - w'/2, (v - w/2) - w'/2 \rangle
\]
\[
= \langle (v/2 - w/2) + (v/2 - w'/2), (v/2 - w/2) + (v/2 - w'/2) \rangle
\]
\[
= (1/4) \| v - w \|^2_2 + 2 \Re \langle (v/2 - w/2, v/2 - w'/2) \rangle + (1/4) \| v - w' \|^2_2
\]
\[
= (1/2) d^2 + (1/2) \Re \langle v - w, v - w' \rangle
\]
\[
\leq (1/2) d^2 + (1/2) | \langle v - w, v - w' \rangle |
\]
\[
< (1/2) d^2 + (1/2) \| v - w \|_2 \| v - w' \|_2
\]
\[
= d^2,
\]
(4.82)
where (*) follows by the Cauchy-Schwarz Inequality (Theorem 4.2.6). Thus, we have that
\[
\left\| v - \left( \frac{w + w'}{2} \right) \right\|^2 < d.
\]
(4.83)
But this is a contradiction with the minimality of \( d \), so we must have that \( w = w' \). \( \square \)

**Lemma 4.6.7.** Let \( H \) be a Hilbert space. Let \( W \) be a closed subspace of \( H \). Then every element \( x \in H \) can be uniquely written in the form \( x = z + w \) where \( z \in W \) and \( w \in W^\perp \).

Let \( x \in H \). Then by Lemma 4.6.6, we know that there is a unique element \( z \in W \) that minimizes \( \| v - w \| \). Let \( w = x - z \). Then clearly, \( x = z + w \). It remains to show that \( w \in W^\perp \).

Let \( t \in \mathbb{R} \) and \( d = \| x - z \|_2 = \| w \|_2 \). Then since \( d = \inf \{ \| v - w \|_2 : w \in W \} \), it follows that \( d \leq \| x - r \|_2 \) for any \( r \in W \). Now, if \( y \in W \), then \( z + ty \in W \). Thus,
\[
\begin{align*}
    d^2 & \leq \|x - (z + ty)\|_2^2 \\
    & = \|w - ty\|_2^2 \quad (4.84) \\
    & = d^2 - 2t\Re\langle w, y \rangle + t^2 \|y\|_2^2.
\end{align*}
\]

It follows that \(-2t\Re\langle w, y \rangle + t^2 \|y\|_2^2 \geq 0\) for all \(t \in \mathbb{R}\) and thus we must have that \(\Re\langle w, y \rangle = 0\).

If we choose \(t_i \in \mathbb{C}\), then we still have that \(z + tiy \in W\) for \(y \in W\). Thus,

\[
\begin{align*}
    d^2 & \leq \|x - (z + tiy)\|_2^2 \\
    & = \|w - tiy\|_2^2 \quad (4.85) \\
    & = d^2 - 2t\Im\langle w, y \rangle + t^2 \|y\|_2^2.
\end{align*}
\]

It follows that \(-2t\Im\langle w, y \rangle + t^2 \|y\|_2^2 \geq 0\) for all \(t \in \mathbb{R}\) and thus we must have that \(\Im\langle w, y \rangle = 0\).

It follows that \(\langle w, y \rangle = \Re\langle w, y \rangle + i\Im\langle w, y \rangle = 0\). Thus, \(w \in W^\perp\).

Finally, suppose that \(x \in H\) such that \(x = z + w = z' + w'\) for \(z, z' \in W\) and \(w, w' \in W^\perp\). Then \(z - z' = w' - w\) which implies that \((z - z'), (w' - w) \in W \cap W^\perp\).

But, by Lemma 4.6.5 \(W \cap W^\perp = \{0\}\). Thus, we must have that \(z - z' = w' - w = 0\) which implies that \(z = z'\) and \(w' = w\).

**Theorem 4.6.8.** Let \(W\) be a closed subspace of a Hilbert space \(H\). Then \(H\) is isometrically isomorphic to \(W \oplus W^\perp\) via the function \(f : W \oplus W^\perp \to H\) given by \(f(z, w) = z + w\).

**Proof.** First of all, Lemma 4.6.7 shows that \(f\) is a bijection.

Let \(z \in W\), \(w \in W^\perp\), and \(c \in \mathbb{C}\). Then

\[
f[c(z, w)] = f(cz, cw) = cz + cw = c(z + w) = cf(z, w). \quad (4.86)
\]
Furthermore, let \((z, w), (z', w') \in W \oplus W^\perp\). Then
\[
f((z, w) + (z', w')) = f(z + z', w + w') = (z + z') + (w + w')
\]
\[
= (z + w) + (z' + w') = f(z, w) + f(z', w').
\]
(4.87)

It follows that \(f\) is an isomorphism.

Finally,
\[
\langle (z, w), (z', w') \rangle = \langle z, z' \rangle + \langle w, w' \rangle
\]
\[
= \langle z + w, z' + w' \rangle \quad (\ast)
\]
\[
= \langle f(z, w), f(z', w') \rangle,
\]
where \((\ast)\) follows by the fact that \(\langle z, w' \rangle = \langle w, z' \rangle = 0\), since \(w, w' \in W\) and \(z, z' \in W^\perp\). It follows that \(f\) is an isometry. \(\Box\)

**Definition 4.6.9.** Let \(W\) be a closed subspace of the Hilbert space \(H\). The orthogonal projection onto \(W\) is the operator \(\text{pr}_W : H \to H\) such that \(\text{pr}_W\) is the identity on \(W\) and the zero map on \(W^\perp\).
CHAPTER 5
GROUP VON NEUMANN ALGEBRAS

This final chapter defines the notions of bounded operators, group von Neumann algebra (denoted $\mathcal{N}(G)$), Hilbert $\mathcal{N}(G)$-modules, and von Neumann dimension. Section 5.4 demonstrates a particular circumstance under which von Neumann dimension may take on a fractional value. Section 5.5 relates the notion of von Neumann dimension to "$\mathcal{U}(G)$"-dimension, and uses the results of Chapter 3, section 3.6, to demonstrate the conditions under which certain $\mathcal{U}(G)$-modules have integral dimension.

5.1 Bounded Linear Operators

The following definitions refer to section 2.2 and Royden [Roy88, Chapter 10].

**Definition 5.1.1.** Let $V$ be a Hilbert space. We say that $A \in \text{End}(V_C)$ is **bounded** if there exists $M \in \mathbb{R}$ such that for all $x \in V$, we have $\|Ax\|_2 \leq M \|x\|_2$. The least such $M$ is called the **norm** of $A$ and is denoted $\|A\|$. Specifically, $\|A\| = \sup_{x \in V, x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$.

Definition 5.1.1 holds, *mutatis mutandis*, for $A \in \text{End}(cV)$.

**Definition 5.1.2.** Let $V$ be a Hilbert space. Then the set

$$\mathcal{B}(V_C) = \{A \in \text{End}(V_C) : \|A\| < \infty\}.$$ (5.1)
Similarly,

$$\mathcal{B}(cV) = \{ A \in \text{End}(cV) : \| A \| < \infty \}.$$  \hfill (5.2)

**Lemma 5.1.3.** A $\mathbb{C}$-linear combination of elements of $\mathcal{B}(V_c)$ is itself an element of $\mathcal{B}(V_c)$.

*Proof.* We first show that a $\mathbb{C}$-linear combination of elements of $\mathcal{B}(V_c)$ is in $\text{End}(V_c)$. Let $A_1, \ldots, A_k \in \mathcal{B}(V_c)$ and $c_1, \ldots, c_k \in \mathbb{C}$. We know by Theorem 2.2.6 that $c_i A_i \in \text{End}(V_c)$ for all $i$. By Theorem 2.2.5, we know that $\text{End}(V_c)$ is a ring, which implies that $c_1 A_1 + \cdots + c_k A_k$ is in $\text{End}(V_c)$. It remains to show that $c_1 A_1 + \cdots + c_k A_k$ is bounded. Denote the norm of $c_1 A_1 + \cdots + c_k A_k$ as $\left\| \sum_{i=1}^k c_i A_i \right\|$. Then

$$\left\| \sum_{i=1}^k c_i A_i \right\| = \sup_{x \in V, x \neq 0} \frac{\left\| \left( \sum_{i=1}^k c_i A_i \right)x \right\|_2}{\| x \|_2}$$

$$\leq \sup_{x \in V, x \neq 0} \sum_{i=1}^k |c_i| \frac{\| A_i x \|_2}{\| x \|_2} \quad (\ast)$$

$$\leq \sum_{i=1}^k |c_i| \sup_{x \in V, x \neq 0} \frac{\| A_i x \|_2}{\| x \|_2}$$

$$< \infty \quad (**),$$

where $(\ast)$ follows by the Triangle Inequality (Lemma 4.2.8) and Definition 4.2.9, and $(\ast\ast)$ follows because each $A_i$ is bounded by assumption. \qed

**Lemma 5.1.4.** Let $A$ and $B \in \mathcal{B}(V_c)$. Then the composition of $A$ and $B$ is also in $\mathcal{B}(V_c)$. 
Proof. By Theorem 2.2.5, we know that \( AB \in \text{End}(V_c) \). Furthermore,
\[
\|AB\| = \sup_{x \in V, x \neq 0} \frac{\|ABx\|}{\|x\|_2} \\
\leq \sup_{x \in V, x \neq 0} \frac{\|A\|\|Bx\|}{\|x\|_2} \\
= \|A\| \sup_{x \in V, x \neq 0} \frac{\|Bx\|}{\|x\|_2} < \infty. \quad (5.4)
\]

**Theorem 5.1.5.** Let \( V \) be a Hilbert space. Then \( \mathcal{B}(V_c) \) is a subring of the ring \( \text{End}(V_c) \).

**Proof.** Let \( A, B \in \mathcal{B}(V) \). By Lemma 5.1.3, we know that \( A + B \) and \( A - B \in \mathcal{B}(V_c) \). Furthermore, by Lemma 5.1.4, we have that \( AB \) is also in \( \mathcal{B}(V_c) \). The result follows. \( \square \)

**Theorem 5.1.6.** Let \( V \) be a Hilbert space, and let \( A \in \text{End}(V_c) \). Then the following are equivalent:

1. \( A \) is continuous on \( V \).
2. \( A \) is continuous at one point of \( V \).
3. \( A \) is bounded.

The following proof is taken from Royden [Roy88, Chapter 10, Proposition 2].

**Proof.** (1) \( \Rightarrow \) (2): The result follows immediately since \( A \) is continuous at every point of \( V \).

(2) \( \Rightarrow \) (3): Suppose \( A \) is continuous at \( x_0 \in V \). Then for \( \epsilon = 1 \), there exists \( \delta > 0 \) such that \( \|Ax - Ax_0\|_2 < 1 \) for \( x \in V \) such that \( \|x - x_0\|_2 < \delta \). Let \( \alpha = \delta / 2 \). For any \( z \in V \) such that \( z \neq 0 \), set \( w = \frac{\alpha z}{\|z\|_2} \). Then \( \|w\|_2 = \left( \frac{\alpha}{\|z\|_2} \right) \|z\|_2 = \alpha \).

Additionally, since \( A \) is linear,
\[
\frac{\alpha}{\|z\|_2} Az = Aw = A(w + x_0) - A(x_0) \quad (5.5)
\]
and
\[
\frac{\alpha}{\| z \|_2} \| A z \|_2 = \| A(w + x_0) - A(x_0) \|_2 < 1, \tag{5.6}
\]
because \( \| (w + x_0) - x_0 \|_2 = \| w \|_2 = \alpha < \delta \). Thus, \( \| A z \|_2 \leq \alpha^{-1} \| z \|_2 \) for all \( z \in V \).
Hence, \( A \) is bounded.

(3) \( \Rightarrow (1) \): Suppose \( A \) is bounded. Then \( A \) has norm \( \| A \| < \infty \). Let \( \epsilon > 0 \).
Choose \( \delta = \frac{\epsilon}{\| A \| + 1} \). Then for \( x_1, x_2 \in V \) such that \( \| x_1 - x_2 \|_2 < \delta \), we have
\[
\| A x_1 - A x_2 \|_2 = \| A(x_1 - x_2) \|_2 \leq \| A \| \| x_1 - x_2 \|_2 < \epsilon. \tag{5.7}
\]
Thus, by Definition 4.1.40, we have that \( A \) is uniformly continuous and therefore by Theorem 4.1.41, \( A \) is continuous on \( V \). \( \square \)

5.2 \( \ell^2(G) \) is a \( \mathbb{C}[G] \)-bimodule

**Theorem 5.2.1.** Let \( G \) be a countable multiplicative group. Let \( a = \sum_{g \in G} a_g g \in \ell^2(G) \).
Define \( \lambda : G \to GL(\ell^2(G)) \) by \( \lambda(h) = \lambda_h \) for all \( h \in G \), where
\[
\lambda_h(a) = \sum_{g \in G} a_g (hg). \tag{5.8}
\]
Define \( \rho : G \to GL(\mathbb{C} \ell^2(G)) \) by \( \rho(h) = \rho_h \) for all \( h \in G \), where
\[
(a)\rho_h = \sum_{g \in G} a_g (gh). \tag{5.9}
\]
Then \( \| \lambda_h(a) \|_2 = \| a \|_2 = \| (a)\rho_h \|_2 \) and \( \lambda \) and \( \rho \) are representations of \( G \) in \( \ell^2(G) \).

**Proof.** Let \( h \in G \), and \( a = \sum_{g \in G} a_g g \in \ell^2(G) \). We first show that
\[
\| \lambda_h(a) \|_2 = \| a \|_2 = \| (a)\rho_h \|_2. \tag{5.10}
\]
We have that
\[
\lambda_h(a) = \sum_{g \in G} a_g (hg) = \sum_{k \in G} a_{h^{-1}k} k, \tag{5.11}
\]
where $k = hg$. We can see by (5.11) that $\lambda_h$ has the effect of re-arranging the coefficients of $a = \sum_{g \in G} a_g g$. Since $a \in \ell^2(G)$, we know that

$$\|a\|_2^2 = \sum_{g \in G} |a_g|^2 < \infty. \quad (5.12)$$

By Theorem 4.3.3, we know that any arrangement (or numbering) of the terms of $a$ will converge to the same value. Thus, $\|\lambda_h(a)\|_2 = \|a\|_2$. A similar argument shows that $\|(a)\rho_h\|_2 = \|a\|_2$. Thus, we have that $\lambda_h(a) \in \ell^2(G)_C$ and $(a)\rho_h \in C\ell^2(G)$.

Next, we show that $\lambda$ is a group homomorphism. Suppose $h, k \in G$. Then $\lambda(hk) = \lambda_{hk}$. Let $a \in \ell^2(G)_C$. Then

$$\lambda_{hk}(a) = \sum_{g \in G} a_g [(hk)g]$$

$$= \sum_{g \in G} a_g [h(kg)]$$

$$= \lambda_h \left( \sum_{g \in G} a_g (kg) \right)$$

$$= \lambda_h \left[ \lambda_k \left( \sum_{g \in G} a_g g \right) \right]$$

$$= \lambda_h(\lambda_k(a))$$

$$= (\lambda_h \lambda_k)(a).$$

It follows that $\lambda(hk) = \lambda_{hk} = \lambda_h \lambda_k = \lambda(h)\lambda(k)$. A similar argument can be used to show that $\rho(hk) = \rho(h)\rho(k)$.

Next, we show that $\lambda_h$ is a right $C$-module isomorphism. Let $a = \sum_{g \in G} a_g g$ and $b = \sum_{g \in G} b_g g$ be in $\ell^2(G)$. 

Then
\[ \lambda_h(a + b) = \lambda_h \left( \sum_{g \in G} a_g g + \sum_{g \in G} b_g g \right) \]
\[ = \lambda_h \left( \sum_{g \in G} (a_g + b_g) g \right) \]
\[ = \sum_{g \in G} (a_g + b_g) (h g) \]
\[ = \sum_{g \in G} (a_g (h g) + b_g (h g)) \]
\[ = \sum_{g \in G} a_g (h g) + \sum_{g \in G} b_g (h g) \]
\[ = \lambda_h(a) + \lambda_h(b). \]  \hfill (5.14)

Furthermore, let \( c \in \mathbb{C} \). Then
\[ \lambda_h(ac) = \lambda_h \left( \left( \sum_{g \in G} a_g g \right) c \right) \]
\[ = \lambda_h \left( \sum_{g \in G} (a_g c) g \right) \]
\[ = \sum_{g \in G} a_g c (h g) \]
\[ = \left( \sum_{g \in G} a_g (h g) \right) c \]
\[ = \lambda_h(a)c. \]  \hfill (5.15)

Since \( h \in G \), and \( G \) is a group, we know that \( h^{-1} \in G \). Consider \( \lambda_{h^{-1}} \). Then by (5.13), we have that
\[ \lambda_h \lambda_{h^{-1}} = \lambda_{h^{-1}} = \lambda_{1_G} = \lambda_{h^{-1}h} = \lambda_{h^{-1}} \lambda_h, \]  \hfill (5.16)

where \( \lambda(1_G) = \lambda_{1_G} \) is the identity map on \( \ell^2(G)_{\mathbb{C}} \) (by (5.8)). Thus, we have that \( \lambda_h \) is bijective.

It follows that \( \lambda_h \in GL(\ell^2(G)_{\mathbb{C}}) \). A similar argument can be used to show that \( \rho_h \in GL(\ell^2(G)) \).
Thus, $\lambda$ and $\rho$ are representations of $G$ in $\ell^2(G)$. \qed

Note that if we combine the results of Theorem 5.2.1 with Theorems 2.7.4 and 2.7.5, we have that

$$\|ha\|_2 = \|\lambda_h(a)\|_2 = \|a\|_2 = \|(a)\rho_h\|_2 = \|ah\|_2.$$  

(5.17)

**Theorem 5.2.2.** Let $G$ be a countable group, $\lambda$ and $\rho$ be as in Theorem 5.2.1, $\sum_{h \in G} r_h h$, $\sum_{k \in G} s_k k \in \mathbb{C}[G]$ and $\sum_{g \in G} a_g g \in \ell^2(G)$. Then $\ell^2(G)$ is a $\mathbb{C}[G]$-bimodule with left action given by

$$\left(\sum_{h \in G} r_h h\right) \left(\sum_{g \in G} a_g g\right) = \sum_{h \in G} \sum_{g \in G} r_h \lambda_h(a_g g) = \sum_{h \in G} \sum_{g \in G} r_h a_g(hg).$$  

(5.18)

and the right action of $\mathbb{C}[G]$ on $\ell^2(G)$ given by

$$\left(\sum_{g \in G} a_g g\right) \left(\sum_{k \in G} s_k k\right) = \sum_{k \in G} \sum_{g \in G} (a_g g)s_k \rho_k = \sum_{k \in G} \sum_{g \in G} a_g s_k(gk).$$  

(5.19)

Note that the summations in the left and right actions, (5.18) and (5.19), are well-defined because the sums $\sum_{h \in G} r_h h$, $\sum_{k \in G} s_k k \in \mathbb{C}[G]$ each have only finitely many non-zero terms. Even though $\sum_{g \in G} a_g g \in \ell^2(G)$ may have infinitely many non-zero terms, the left and right actions are well-defined because $\sum_{g \in G}$ is the innermost summation.

*Proof.* Since $\mathbb{C}$ is a commutative ring, $\ell^2(G)$ is a left [resp. right] $\mathbb{C}$-module (see Lemma 4.4.3), and $\lambda$ and $\rho$ are representations of $G$ in $\ell^2(G)$, by Theorems 2.7.4 and 2.7.5, it follows that $\ell^2(G)$ is a left and right $\mathbb{C}[G]$-module with left and right actions given by (5.18) and (5.19).
Let $r = \sum_{h \in G} r_h h$ and $s = \sum_{k \in G} s_k k$ be in $\mathbb{C}[G]$ and let $a = \sum_{g \in G} a_g g \in \ell^2(G)$. Then

$$r(as) = \left( \sum_{h \in G} r_h h \right) \left[ \left( \sum_{g \in G} a_g g \right) \left( \sum_{k \in G} s_k k \right) \right]$$
$$= \left( \sum_{h \in G} r_h h \right) \left[ \sum_{k \in G} \sum_{g \in G} a_g s_k (gk) \right]$$
$$= \sum_{h \in G} \sum_{k \in G} \sum_{g \in G} r_h (a_g s_k) [h(gk)]$$
$$= \sum_{k \in G} \sum_{g \in G} \sum_{h \in G} (r_h a_g) s_k [(hg)k]$$
$$= \sum_{k \in G} \sum_{h \in G} \sum_{g \in G} (r_h a_g) s_k [(hg)k]$$

(5.20)

where (*) follows by switching the finite summations $\sum$ and $\sum$.

It follows that $\ell^2(G)$ is a $\mathbb{C}[G]$-bimodule.

5.3 Group von Neumann Algebras

The following definition is adapted from Luck [Lüc02, Definition 1.1].

**Definition 5.3.1.** Let $G$ be a countable group. Then the **Group von Neumann algebra** of $G$, denoted $\mathcal{N}(G)$, is the subring

$$\mathcal{N}(G) = \text{End}(\ell^2(G)_{\mathbb{C}[G]}) \cap \mathcal{B}(\ell^2(G)_\mathbb{C})$$

(5.21)

$$= \mathcal{B}(\ell^2(G)_{\mathbb{C}[G]}),$$

of the ring $\text{End}(\ell^2(G)_\mathbb{C})$. 

Theorem 5.3.2. Let $G$ be a countable group and let $\lambda$ be as in Theorem 5.2.1. Then there exists an injective ring homomorphism $\iota : \mathbb{C}[G] \rightarrow \mathcal{N}(G)$ given by $\iota(r) = \lambda_r$, for $r = \sum_{h \in G} r_h h \in \mathbb{C}[G]$, where $\lambda_r(a) = ra$ for $a \in \ell^2(G)$.

Proof. Let $h \in G$. We showed in Theorem 5.2.1 that $\lambda_h \in GL(\ell^2(G)_C) \subseteq \text{End}(\ell^2(G)_C)$ and that $\|\lambda_h(a)\|_2 = 1 \|a\|_2$. Thus, $\lambda_h \in \mathcal{B}(\ell^2(G)_C)$. By Lemma 5.1.3 we know that a $\mathbb{C}$-linear combination of elements of $\mathcal{B}(\ell^2(G)_C)$ is in $\mathcal{B}(\ell^2(G)_C)$ as well. This shows that any for $r = \sum_{h \in G} r_h h \in \mathbb{C}[G]$, we have that $\lambda_r = \sum_{h \in G} r_h \lambda_h$ is also in $\mathcal{B}(\ell^2(G)_C)$.

On the other hand, Theorem 5.2.2 shows us that $\lambda_r \in \text{End}(\ell^2(G)_{\mathbb{C}[G]})$. Specifically, let $r, s \in \mathbb{C}[G]$ and $a \in \ell^2(G)$. Then

$$\lambda_r(as) = r(as) = (ra)s = ([\lambda_r(a)]s). \quad (5.22)$$

We also have that

$$\lambda_r(a + b) = r(a + b) = ra + rb = \lambda_r(a) + \lambda_r(b), \quad (5.23)$$

since $\ell^2(G)$ is also a left $\mathbb{C}[G]$-module.

Thus, $\iota$ is well-defined because $\lambda_r \in \text{End}(\ell^2(G)_{\mathbb{C}[G]}) \cap \mathcal{B}(\ell^2(G)_C) = \mathcal{N}(G)$.

Now, suppose $r, s \in \mathbb{C}[G]$ and $a \in \ell^2(G)$. Then $\iota(r + s) = \lambda_{r+s}$, and

$$\lambda_{r+s}(a) = (r + s)a = ra + sa \quad (*)$$

$$= \lambda_r(a) + \lambda_s(a)$$

$$= (\lambda_r + \lambda_s)(a), \quad (5.24)$$
where (*) follows since $\ell^2(G)$ is a left $C[G]$-module. Furthermore, $\iota(rs) = \lambda_{rs}$, and

$$
\lambda_{rs}(a) = (rs)a = r(sa) \quad (*)
= \lambda_r(\lambda_s(a)) = (\lambda_r\lambda_s)(a),
$$

where (*) follows since $\ell^2(G)$ is a left $C[G]$-module. Thus, we have that

$$
\iota(r + s) = \lambda_{(r+s)} = \lambda_r + \lambda_s = \iota(r) + \iota(s). \quad (5.26)
$$

and

$$
\iota(rs) = \lambda_{rs} = \lambda_r\lambda_s = \iota(r)\iota(s). \quad (5.27)
$$

It remains to show that $\ker(\iota) = \{0\}$. By definition, $\{0\} \subseteq \ker(\iota)$. Suppose $r \in \ker(\iota)$. Then $\iota(r) = \lambda_r = \lambda_0$. So, $\lambda_r(x) = 0$ for all $x \in \ell^2(G)$. Specifically, $r1_G = 0$, which implies that $r = 0$. The result follows.

\textbf{Theorem 5.3.3.} Let $G$ be a countable group, and let $1_G$ be the identity element of $G$. For $S, T \in \mathcal{N}(G)$, if $S(1_G) = T(1_G)$, then $S = T$.

\textbf{Proof.} For $x \in C[G]$, $S(x) = S(1_G)x = T(1_G)x = T(x)$.

By Theorem 4.5.4, for $x \in \ell^2(G)$, $x = \lim_{n \to \infty} x_n$ for some $x_n \in C[G]$. So,

$$
S(x) = S\left(\lim_{n \to \infty} x_n\right) = \lim_{n \to \infty} S(x_n) \quad (*)
= \lim_{n \to \infty} S(1_G)x_n = \lim_{n \to \infty} T(1_G)x_n = \lim_{n \to \infty} T(x_n)
= T\left(\lim_{n \to \infty} x_n\right) \quad (**) = T(x),
$$
where (*) and (**) follow by the continuity of $S$ and $T$ as given by Theorem 5.1.6. □

**Corollary 5.3.4.** Let $T \in \mathcal{N}(G)$. Then if $T(1_G) \in \mathbb{C}[G]$, then there exists $a \in \mathbb{C}[G]$ such that $T(x) = ax$.

**Proof.** Suppose $T(1_G) = a \in \mathbb{C}[G]$. We know that the operator $S(x) = ax$ is such that $S \in \mathcal{N}(G)$. Then $S(1_G) = a$, and by Theorem 5.3.3, we have that $S = T$. □

The following definition is adapted from Lück [Luc02, Definition 1.5].

**Definition 5.3.5.** Let $G$ be a countable group. A (finitely generated) **left Hilbert $\mathcal{N}(G)$-module** is a Hilbert space $V$ such that

1. $V$ is a left $\mathbb{C}[G]$-module;

2. $\|ga\|_2 = \|a\|_2$ for all $g \in G$ and $a \in V$; and

3. there exists $n$ and an injective, isometric left $\mathbb{C}[G]$-module homomorphism

$$f : V \to \bigoplus_{i=1}^n \ell^2(G).$$

The following definition is taken from Lück [Luc02, Section 1.1.3].

**Definition 5.3.6.** Let $H$ be a Hilbert space. A bounded operator $t : H \to H$ is called **positive** if $(t(v), v) \in \mathbb{R}^+ \cup \{0\}$ for all $v \in H$.

Example: The id operator on $\ell^2(G)$ is bounded because $\|\text{id}(v)\|_2 = \|v\|_2 = 1 \|v\|_2$ for all $v \in \ell^2(G)$. The id operator is positive because $(\text{id}(v), v) = (v, v) \geq 0$ for all $v \in \ell^2(G)$.

Example: Let $W$ be a closed subspace of the Hilbert space $H$. Let $\text{pr}_W : H \to H$ be as in Definition 4.6.9. Let $z \in H$. Then by Lemma 4.6.7, $z = x + y$ for $x \in W$
and \( y \in W^\perp \). So,

\[
\|z\|_2^2 = \|x + y\|_2^2 = \langle x + y, x + y \rangle \\
= \langle x, x \rangle + 2\Re \langle x, y \rangle + \langle y, y \rangle \\
= \|x\|_2^2 + \|y\|_2^2,
\]

(5.29)

where (*) follows because \( \langle x, y \rangle = 0 \). Thus, we have that \( \|x\|_2, \|y\|_2 \leq \|z\|_2 \). We know that \( \text{pr}_W \) is bounded because by (5.29),

\[
\|\text{pr}_W(z)\|_2 = \|x\|_2 \leq 1 \|z\|_2.
\]

(5.30)

Furthermore,

\[
\langle \text{pr}_W(z), z \rangle = \langle \text{pr}_W(x + y), (x + y) \rangle \\
= \langle x, x + y \rangle \\
= \langle x, x \rangle + \langle x, y \rangle \\
= \langle x, x \rangle \geq 0,
\]

(5.31)

where (*) follows because \( \langle x, y \rangle = 0 \). It follows that \( \text{pr}_W \) is a positive operator.

The following definition is adapted from Luck [Lüc02, Definitions 1.2 and 1.8].

**Definition 5.3.7.** Let \( s : V \to V \) be a positive endomorphism of a Hilbert left \( \mathcal{N}(G) \)-module \( V \). Choose \( n \) and \( f : V \to \bigoplus_{i=1}^n \ell^2(G) \) as in Definition 5.3.5. Let \( \tilde{s} : \bigoplus_{i=1}^n \ell^2(G) \to \bigoplus_{i=1}^n \ell^2(G) \) be the positive operator given by

\[
\tilde{s} : \bigoplus_{i=1}^n \ell^2(G) \to \bigoplus_{i=1}^n \ell^2(G),
\]

(5.32)

where \( \text{pr}_{f(V)} \) is the projection onto the closed subspace \( f(V) \) (see Lemma 5.3.14 and Definition 4.6.9).

So, \( \tilde{s} = f \circ s \circ f^{-1} \circ \text{pr}_{f(V)} \), and the von Neumann trace of \( s : V \to V \) is defined to be

\[
\text{tr}_{\mathcal{N}(G)}(s) = \sum_{i=1}^n \langle \tilde{s}(E_i), E_i \rangle,
\]

(5.33)
where \( E_t = (a_{i1}, \ldots, a_{in}) \) such that
\[
a_{ij} = \begin{cases} 
1_G & \text{if } i = j \\
0 & \text{otherwise}
\end{cases}
\] (5.34)

**Theorem 5.3.8.** Let \( V, \mathcal{N}(G) \), and \( s \) be as in Definition 5.3.7. Then the von Neumann trace of \( s \) is independent of the choice of \( n \) and \( f \) (as specified in Definition 5.3.5).

**Proof.** See Luck [Lüc02, Chapter 1, after Definition 1.8]. \( \Box \)

The following definition is taken from Luck [Lüc02, Definition 1.10].

**Definition 5.3.9.** The von Neumann dimension of a left Hilbert \( \mathcal{N}(G) \)-module \( V \) is
\[
\dim_{\mathcal{N}(G)} V = \text{tr}(id : V \to V).
\] (5.35)

**Lemma 5.3.10.** Let \( V \) be a left Hilbert \( \mathcal{N}(G) \)-module with \( n \) and \( f \) as in Definition 5.3.5. Let \( id : V \to V \) be the identity map on \( V \). Then \( \overline{id} = \text{pr}_{f(V)} \), and thus
\[
\dim_{\mathcal{N}(G)} V = \text{tr}_{\mathcal{N}(G)}(id) = \sum_{i=1}^{n} \langle \overline{id}(E_i), E_i \rangle = \sum_{i=1}^{n} \langle \text{pr}_{f(V)}(E_i), E_i \rangle.
\] (5.36)

**Proof.** By Definition 5.3.7, we have that
\[
\overline{id} : \bigoplus_{i=1}^{n} \ell^2(G) \xrightarrow{\text{pr}_{f(V)}} f(V) \xrightarrow{f^{-1}} V \xrightarrow{id} V \xrightarrow{f} \bigoplus_{i=1}^{n} \ell^2(G),
\] (5.37)

or \( \overline{id} = f \circ id \circ f^{-1} \circ \text{pr}_{f(V)} = \text{pr}_{f(V)} \). Equation (5.36) follows immediately from Definitions 5.3.7 and 5.3.9. \( \Box \)

**Theorem 5.3.11.** Let \( G \) be a countable group. Then \( \bigoplus_{i=1}^{n} \ell^2(G) \) is a left Hilbert \( \mathcal{N}(G) \)-module and \( \dim_{\mathcal{N}(G)} \bigoplus_{i=1}^{n} \ell^2(G) = n \).
Proof. First of all, \( \ell^2(G) \) satisfies the conditions of Definition 5.3.5. Specifically, Theorem 5.2.2 shows that \( \ell^2(G) \) is a left \( \mathbb{C}[G] \)-module and thus by Theorem 2.4.3, it follows that \( \bigoplus_{i=1}^{n} \ell^2(G) \) is a left \( \mathbb{C}[G] \)-module. Furthermore, let \( \bar{a} = (a_1, a_2, \ldots, a_n) \in \bigoplus_{i=1}^{n} \ell^2(G) \) and let \( g \in G \). Then

\[
\|g\bar{a}\|_2^2 = \|(ga_1, ga_2, \ldots, ga_n)\|_2^2 \\
= \|ga_1\|_2^2 + \|ga_2\|_2^2 + \cdots + \|ga_n\|_2^2 \\
= \|a_1\|_2^2 + \|a_2\|_2^2 + \cdots + \|a_n\|_2^2 \quad (\star)
\]

where (\star) follows by (5.17) after Theorem 5.2.1.

Finally, \( f = \text{id} : \bigoplus_{i=1}^{n} \ell^2(G) \rightarrow \bigoplus_{i=1}^{n} \ell^2(G) \) is an isometric, injective left \( \mathbb{C}[G] \)-module homomorphism. It follows that \( \bigoplus_{i=1}^{n} \ell^2(G) \) is a left Hilbert \( \mathcal{N}(G) \)-module.

Let \( V = \bigoplus_{i=1}^{n} \ell^2(G) \). By Lemma 5.3.10, and the fact that

\[
f(V) = \text{id} \left( \bigoplus_{i=1}^{n} \ell^2(G) \right) = \bigoplus_{i=1}^{n} \ell^2(G) = V.
\]

we have

\[
\dim_{\mathcal{N}(G)}(\ell^2(G))^n = \sum_{i=1}^{n} \langle \text{pr}_V(E_i), E_i \rangle \\
= \sum_{i=1}^{n} \langle E_i, E_i \rangle \quad (\star)
\]

\[
= \sum_{i=1}^{n} 1 \quad (\star \star)
\]

\[
= n
\]

where (\star) follows from the fact that the projection of \( \bigoplus_{i=1}^{n} \ell^2(G) \) onto \( \bigoplus_{i=1}^{n} \ell^2(G) \) is just
the identity map, and (**) follows from the fact that

\[ \langle E_i, E_i \rangle_{\rho^2(G)} = \langle 1_G, 1_G \rangle_{\rho^2(G)} = (1)(1) = 1. \]  

(5.41)

**Theorem 5.3.12.** Let \( V \) and \( W \) be left Hilbert \( \mathcal{N}(G) \)-modules. Then \( V \oplus W \) is a left Hilbert \( \mathcal{N}(G) \)-module and \( \dim_{\mathcal{N}(G)}(V \oplus W) = \dim_{\mathcal{N}(G)} V + \dim_{\mathcal{N}(G)} W. \)

Before proving the theorem, we prove the following lemmas first.

**Lemma 5.3.13.** For \( i = 1, 2 \), let \( W_i \) be a closed subspace of the Hilbert space \( H_i \). Then \( W_1 \oplus W_2 \) is a closed subspace of \( H_1 \oplus H_2 \) and \( (W_1 \oplus W_2)^\perp = W_1^\perp \oplus W_2^\perp. \)

**Proof.** By Theorem 2.4.3, we have that \( W_1 \oplus W_2 \) is a subspace of \( H_1 \oplus H_2 \).

Let \( (v_n, w_n) \in W_1 \oplus W_2 \) be such that \( (v_n, w_n) \to (v, w) \). Let \( \epsilon > 0 \). Then there exists \( N \) such that for \( n \geq N \), we have \( d((v_n, w_n), (v, w)) < \epsilon \) which implies that \( \|v_n - v\|^2 + \|w_n - w\|^2 < \epsilon^2 \), which further implies that \( \|v_n - v\| < \epsilon \). Thus, \( v_n \to v \), and a similar argument shows that \( w_n \to w \). Since \( W_1 \) and \( W_2 \) are closed, it follows that \( v \in W_1 \) and \( w \in W_2 \). Thus, \( (v, w) \in W_1 \oplus W_2 \) and hence \( W_1 \oplus W_2 \) is closed.

It remains to show that \( (W_1 \oplus W_2)^\perp = W_1^\perp \oplus W_2^\perp \), where

\[ (W_1 \oplus W_2)^\perp = \{(x, y) \in H_1 \oplus H_2 : \langle (x, y), (v, w) \rangle = 0 \text{ for all } (v, w) \in W_1 \oplus W_2 \}. \]  

(5.42)

Suppose \( (x, y) \in W_1^\perp \oplus W_2^\perp \). Then \( \langle x, v \rangle = 0 \) and \( \langle y, w \rangle = 0 \) for all \( v \in W_1 \) and \( w \in W_2 \). It follows that \( \langle x, v \rangle + \langle y, w \rangle = 0 \) and thus, \( \langle (x, y), (v, w) \rangle = 0 \). It follows that \( (x, y) \in (W_1 \oplus W_2)^\perp \).

On the other hand, suppose that \( (x, y) \in (W_1 \oplus W_2)^\perp \). Then \( \langle (x, y), (v, w) \rangle = 0 \) for all \( (v, w) \in W_1 \oplus W_2 \). In particular, \( (v, 0) \in W_1 \oplus W_2 \). So, we have that

\[ \langle (x, y), (v, 0) \rangle = 0, \]  

(5.43)
which implies that
\[ \langle x, v \rangle + \langle y, 0 \rangle = 0, \] (5.44)
or
\[ \langle x, v \rangle = 0. \] (5.45)
Thus, \( x \in W_1^\perp \). A similar argument can be used to show that \( y \in W_2^\perp \).

**Lemma 5.3.14.** Let \( V \) be a left Hilbert \( \mathcal{N}(G) \)-module. Let \( f : V \rightarrow \bigoplus_{i=1}^n \ell^2(G) \) be an isometric, injective left \( C[G] \)-module homomorphism. Then \( f(V) \) is a closed subspace of \( \bigoplus_{i=1}^n \ell^2(G) \).

**Proof.** Let \( \{y_n\} \in f(V) \) such that \( y_n \rightarrow y \), and \( y \in \bigoplus_{i=1}^n \ell^2(G) \). Let \( x_n \in V \) such that \( f(x_n) = y_n \). Since \( f \) is an isometry, it follows by Theorem 4.1.25 that \( \{x_n\} \in V \) is a Cauchy sequence. Since \( V \) is a Hilbert space, it follows that \( x_n \rightarrow x \) for some \( x \in V \), which implies, by Theorem 4.1.25, that \( f(x_n) = y_n \rightarrow f(x) \). Since limits are unique, we must have that \( f(x) = y \in f(V) \).

We now prove Theorem 5.3.12.

**Proof.** We first show that \( V \oplus W \) is a left Hilbert \( \mathcal{N}(G) \)-module. We know by assumption that \( V \) and \( W \) are left \( C[G] \)-modules. By Theorem 2.4.3, it follows that \( V \oplus W \) is a left \( C[G] \)-module. Furthermore, suppose \( (a, b) \in V \oplus W \) and \( k \in G \). Then
\[
\|k(a, b)\|^2_2 = \|(ka, kb)\|^2_2 \\
= \|ka\|^2_2 + \|kb\|^2_2 \\
= \|a\|^2_2 + \|b\|^2_2 \quad (\star) \\
= \|(a, b)\|^2_2,
\] (5.46)
where (\star) follows because \( V \) and \( W \) are each left Hilbert \( \mathcal{N}(G) \)-modules.
By condition (3) of Definition 5.3.5, we know that there exist isometric, left \( \mathbb{C}[G] \)-module monomorphisms \( f : V \to \bigoplus_{i=1}^{n} \ell^2(G) \) and \( g : W \to \bigoplus_{i=1}^{m} \ell^2(G) \). We claim that \( h : V \oplus W \to \left( \bigoplus_{i=1}^{n} \ell^2(G) \right) \oplus \left( \bigoplus_{i=1}^{m} \ell^2(G) \right) \) given by \( h(a, b) = (f(a), g(b)) \) is an isometric left \( \mathbb{C}[G] \)-module monomorphism.

Suppose \((a_1, b_1), (a_2, b_2) \in V \oplus W\). Suppose further that \( h(a_1, b_1) = h(a_2, b_2) \). Then \((f(a_1), g(b_1)) = (f(a_2), g(b_2))\) which implies that \( a_1 = a_2 \) and \( b_1 = b_2 \) since both \( f \) and \( g \) are injective.

Furthermore, let \( c \in \mathbb{C}[G] \). Then

\[
h[c(a_1, b_1) + (a_2, b_2)] = h(ca_1 + a_2, cb_1 + b_2)
\]

\[
= (f(ca_1 + a_2), g(cb_1 + b_2))
\]

\[
= (c f(a_1) + f(a_2), c g(b_1) + g(b_2)) \quad (5.47)
\]

\[
= c(f(a_1), g(b_1)) + (f(a_2), g(b_2))
\]

\[
= c h(a_1, b_1) + h(a_2, b_2).
\]

Finally, we have that

\[
\langle h(a_1, b_1), h(a_2, b_2) \rangle = \langle (f(a_1), g(b_1)), (f(a_2), g(b_2)) \rangle
\]

\[
= \langle f(a_1), f(a_2) \rangle + \langle g(b_1), g(b_2) \rangle \quad (*)
\]

\[
= \langle a_1, a_2 \rangle + \langle b_1, b_2 \rangle \quad (**)
\]

\[
= \langle (a_1, b_1), (a_2, b_2) \rangle, \quad (***)
\]

where (*) and (**) follow by Theorem 4.2.14 and (**) follows since \( f \) and \( g \) are isometries themselves. Thus, \( h \) is an isometry.

Let \( \text{id} : \bigoplus_{i=1}^{n+m} \ell^2(G) \to \bigoplus_{i=1}^{n+m} \ell^2(G) \) be the identity map on \( \bigoplus_{i=1}^{n+m} \ell^2(G) \). Then by Definition 5.3.7 and Lemma 5.3.10, we have \( \overline{\text{id}} = \text{pr}_{h(V \oplus W)} \).

We claim that, for \( z \in \bigoplus_{i=1}^{n+m} \ell^2(G) \), where \( z = (z_1, \ldots, z_n, z_{n+1}, \ldots, z_{n+m}) \),

\[
\text{pr}_{h(V \oplus W)}(z_1, \ldots, z_{n+m}) = (\text{pr}_{f(V)}(z_1, \ldots, z_n), \text{pr}_{g(W)}(z_{n+1}, \ldots, z_{n+m})). \quad (5.49)
\]
We first show that \( f(V) \oplus g(W) = h(V \oplus W) \).

Suppose \((y_1, y_2) \in f(V) \oplus g(W)\). Then \(y_1 = f(x_1)\) and \(y_2 = g(x_2)\) for some \(x_1 \in V\) and \(x_2 \in W\). So we have that

\[
(y_1, y_2) = (f(x_1), g(x_2)) = h(x_1, x_2) \in h(V \oplus W).
\]  

(5.50)

Furthermore, suppose that \(z \in h(V \oplus W)\). Then \(z = h(x_1, x_2)\) for some \((x_1, x_2) \in V \oplus W\). That is,

\[
z = h(x_1, x_2) = (f(x_1), g(x_2)) \in f(V) \oplus g(W).
\]  

(5.51)

By Lemma 5.3.14, \(h(V \oplus W)\) is a closed subspace of \(\bigoplus_{i=1}^{n+m} \ell^2(G)\). So, by Theorem 4.6.8, we must have that

\[
\bigoplus_{i=1}^{n+m} \ell^2(G) = h(V \oplus W) \oplus h(V \oplus W)^\perp
\]

\[
= (f(V) \oplus g(W)) \oplus (f(V) \oplus g(W))^\perp
\]

(5.52)

\[
= (f(V) \oplus g(W)) \oplus f(V)^\perp \oplus g(W)^\perp.
\]

By the same reasoning, since \(f(V)\) is a closed subspace of \(\bigoplus_{i=1}^{n} \ell^2(G)\) and \(g(W)\) is a closed subspace of \(\bigoplus_{j=1}^{m} \ell^2(G)\), we must have that

\[
\bigoplus_{i=1}^{n} \ell^2(G) = f(V) \oplus f(V)^\perp
\]  

(5.53)

and

\[
\bigoplus_{j=1}^{m} \ell^2(G) = g(W) \oplus g(W)^\perp.
\]  

(5.54)

Let \(z \in \bigoplus_{i=1}^{n+m} \ell^2(G)\). Then by Theorem 4.6.8, \(z = x+y\) for some \(x \in h(V \oplus W)\) and \(y \in h(V \oplus W)^\perp\). Since \(h(V \oplus W) = f(V) \oplus g(W)\) and \(h(V \oplus W)^\perp = f(V)^\perp \oplus g(W)^\perp\),
we must have that \( x = (v, w) \) and \( y = (v', w') \) where \( v \in f(V), w \in g(W), v' \in f(V)^\perp, \) and \( w' \in g(W)^\perp. \) Thus, \( z = (v, w) + (v', w') = (v + v', w + w') \), where \( v + v' \in \bigoplus_{i=1}^{m} \ell^2(G) \) and \( w + w' \in \bigoplus_{j=1}^{n} \ell^2(G) \) by (5.53) and (5.54). Furthermore,

\[
\text{pr}_{h(V \oplus W)}(z) = x = (v, w) = (\text{pr}_{f(V)}(v + v'), \text{pr}_{g(W)}(w + w')).
\]

Let \( E_i \in \bigoplus_{i=1}^{n+m} \ell^2(G), E_i' \in \bigoplus_{i=1}^{n} \ell^2(G) \) and \( E_j'' \in \bigoplus_{j=1}^{m} \ell^2(G) \) be defined as in Definition 5.3.7. Then

\[
\dim_{\mathcal{N}(G)}(V \oplus W) = \sum_{i=1}^{n+m} \langle \text{pr}_{h(V \oplus W)}(E_i), E_i \rangle = \sum_{i=1}^{n} \langle \text{pr}_{f(V)}(E_i'), E_i' \rangle + \sum_{j=1}^{m} \langle \text{pr}_{g(W)}(E_j''), E_j'' \rangle = \dim_{\mathcal{N}(G)} V + \dim_{\mathcal{N}(G)} W.
\]

### 5.4 Examples when \( G \) is Finite

**Theorem 5.4.1.** Let \( G \) be a finite multiplicative group. Let \( \iota : \mathbb{C}[G] \to \mathcal{N}[G] \) be given by \( \iota(a) = \lambda_a \), where \( \lambda_a(x) = ax \) for all \( x \in \ell^2(G) \). Let \( f : \mathcal{N}(G) \to \ell^2(G) \) be given by \( f(T) = T(1_G) \). Then \( f \) and \( \iota \) are bijections.

**Proof.** By Theorem 5.3.2, we have that \( \iota : \mathbb{C}[G] \to \mathcal{N}(G) \) given by \( \iota(a) = \lambda_a \) is injective. Suppose \( T \in \mathcal{N}(G) \). Then by definition of \( \mathcal{N}(G) \), (Definition 5.3.1), we know that \( T(1_G) \in \ell^2(G) = \mathbb{C}[G] \), since \( G \) is finite. Then by Corollary 5.3.4, there exists \( a \in \mathbb{C}[G] \) such that \( T(x) = ax \). That is, there exists \( a \in \mathbb{C}[G] \) such that \( \iota(a) = T \). Thus, \( \iota \) is a bijection.
We now show that $f$ is injective. Suppose $S, T \in \mathcal{N}(G)$ and that $f(S) = f(T)$. Then $S(1_G) = T(1_G)$ which implies, by Theorem 5.3.3, that $S = T$.

Now, $f \circ \iota(a) = f(\lambda_a) = \lambda_a(1_G) = a1_G = a$. It follows that $f \circ \iota$ is the identity map. The identity map is bijective, as is $\iota^{-1}$. Since the composition of two bijective functions is bijective, and function composition is associative, we have that $f = (f \circ \iota) \circ \iota^{-1}$ is a bijection. \hfill \Box

The following definition is taken from Serre [Ser77, p. 7].

**Definition 5.4.2.** Let $V$ be a left $\mathbb{C}[G]$-module. Then $V$ is **irreducible** if $V$ has no proper non-trivial submodules.

**Theorem 5.4.3.** Let $G$ be a finite group, let $k$ be the number of conjugacy classes of elements in $G$, and let $M_n(\mathbb{C})$ be the ring of $n \times n$ complex matrices. Then

(1) Any finite-dimensional left $\mathbb{C}[G]$-module is the direct sum of irreducible left $\mathbb{C}[G]$-modules.

(2) There exist exactly $k$ isomorphism classes $V_1, \ldots, V_k$ of finite-dimensional irreducible left $\mathbb{C}[G]$-modules. Additionally, for $1 \leq i \leq k$, we may choose $V_i$ to be a Hilbert space such that $\|ga\|_2 = \|a\|_2$ for all $a \in V_i$ and $g \in G$.

(3) Let $n_i = \dim_{\mathbb{C}} V_i$. Then there exists a ring isomorphism

$$
\mu : \mathbb{C}[G] \to \bigoplus_{i=1}^k M_{n_i}(\mathbb{C})
$$

such that $\mu(1_G) = (I_{n_1}, \ldots, I_{n_k})$, and $\mu$ is an isometry of Hilbert spaces.

(4) There exists an isometric injective left $\mathbb{C}[G]$-module homomorphism $f : V_i \to \mathbb{C}[G]$ such that for $A_i \in M_{n_i}(\mathbb{C})$,

$$
pr_{\mu(f)(V_i)}(A_1, \ldots, A_k) = (0, \ldots, 0, \hat{A}_i, 0, \ldots, 0),
$$

(5.58)
where $\tilde{A}_i$ is the matrix whose first column is the first column of $A$ and the rest of whose entries are 0.

Note that by (4), we have that $V_i \cong f(V_i)$ where $f(V_i)$ is a submodule of $\mathbb{C}[G]$.

**Proof.** By Theorem 2.7.6, any $\mathbb{C}[G]$-module is a representation of $G$. By Serre [Ser77, Theorem 2], every representation is a direct sum of irreducible representations. For item (2), see Serre [Ser77, Theorem 7] and Isaacs [Isa94, Theorem 4.17]. Items (3) and (4) follow from Isaacs [Isa94, Lemma 1.13 and Theorem 1.15]. □

**Lemma 5.4.4.** Let $G$ be a finite group of order $n$. Let $\lambda$ be as in Theorem 5.2.1. Then for $g \in G$,

$$\text{tr}([\lambda_g]_G) = \begin{cases} |G| & \text{if } g = 1_G \\ 0 & \text{if } g \neq 1_G \end{cases} \quad (5.59)$$

**Proof.** Let $G = \{g_1, g_2, \ldots, g_n\}$. Recall that $\lambda : G \to GL(\ell^2(G))$ is defined to be $\lambda(g) = \lambda_g$ for all $g \in G$, where $\lambda_g(a) = ga$ for all $a \in \ell^2(G)$. Note that since $G$ is finite, $\mathbb{C}[G] = \ell^2(G)$, and thus, $\lambda_g : \mathbb{C}[G] \to \mathbb{C}[G]$ where $G$ is a basis for $\mathbb{C}[G]$.

Suppose $g = 1_G$. Then by Definition 3.3.4, we have

$$[\lambda_{1_G}]_G = [\lambda_{1_G}(g_1)]_G \cdots [\lambda_{1_G}(g_n)]_G = [g_1]_G \cdots [g_n]_G = [e_{g_1} \cdots e_{g_n}] = I_n. \quad (5.60)$$

Thus, $\text{tr}([\lambda_{1_G}]_G) = \text{tr}(I_n) = n = |G|$. On the other hand, suppose that $g \neq 1_G$. Then

$$[\lambda_g]_G = [\lambda_g(g_1)]_G \cdots [\lambda_g(g_n)]_G = [gg_1]_G \cdots [gg_n]_G = [e_{gg_1} \cdots e_{gg_n}] \quad (5.61)$$

We claim that none of the diagonal entries are 1. That is, we claim $e_{gg_i} \neq e_{g_i}$. Suppose on the contrary that $e_{gg_i} = e_{g_i}$. Then $gg_i = g_i$ which implies that $g = 1_G$. But this contradicts our assumption and thus we must have that $e_{gg_i} \neq e_{g_i}$ for all $i$. Since the only possible entries in the matrix are 0 and 1, we must have that the diagonal entries are 0 and thus, $\text{tr}([\lambda_g]_G) = 0$. □

Lemma 5.4.5. Let $G$ be a group. Then

$$[\lambda^{-1}_g]_G = [\lambda_g]_G^t = [\lambda_g]^*_G.$$ (5.62)

Proof. We immediately have that $[\lambda_g]_G^t = [\lambda_g]^*_G$ since all of the entries in $[\lambda_g]_G$ are real (they are either 1 or 0). It remains to show that $[\lambda^{-1}_g]_G = [\lambda_g]_G$. Now,

$$[\lambda^{-1}_g]_G = \left[\begin{array}{ccc} g^{-1} g_1 & \cdots & g^{-1} g_n \end{array}\right] = \left[\begin{array}{ccc} e_{g^{-1} g_1} & \cdots & e_{g^{-1} g_n} \end{array}\right].$$ (5.63)

If we let $a_{ij}$ denote the entries of $[\lambda^{-1}_g]_G$, then

$$a_{ij} = \begin{cases} 1 & \text{if } g_i = g^{-1} g_j \\ 0 & \text{otherwise.} \end{cases}$$ (5.64)

We have that

$$[\lambda_g]_G = \left[\begin{array}{ccc} \lambda_g(g_1) & \cdots & \lambda_g(g_n) \end{array}\right] = \left[\begin{array}{ccc} e_{g g_1} & \cdots & e_{g g_n} \end{array}\right].$$ (5.65)

If we let $b_{ij}$ denote the entries of $[\lambda_g]_G$, then

$$b_{ij} = \begin{cases} 1 & \text{if } g_i = g g_j \\ 0 & \text{otherwise.} \end{cases}$$ (5.66)

So,

$$b_{ji} = \begin{cases} 1 & \text{if } g_j = g^{-1} g_i \\ 0 & \text{otherwise.} \end{cases}$$ (5.67)

$$= a_{ij}. \quad \square$$
Theorem 5.4.6. Let $G$ be a finite group. Let $a = \sum_{g \in G} a_g g$ and $b = \sum_{h \in G} b_h h$ be in $\ell^2(G)$.

Let $\lambda$ be as in Theorem 5.2.1. Then the inner product $(a, b) = \frac{1}{|G|} \text{tr} ([\lambda_a]_G [\lambda_b]_G^*).

Note that in the proof below $\lambda_g$ will be denoted as $\lambda(g)$.

Proof. We have that

\[
\text{tr}([\lambda(a)]_G [\lambda(b)]_G^*) = \text{tr} \left( \lambda \left( \sum_{g \in G} a_g g \right)_G \left( \sum_{h \in G} b_h h \right)_G^* \right).
\]

(5.68)

\[
= \text{tr} \left( \sum_{g \in G} a_g \lambda(g)_G \sum_{h \in G} b_h \lambda(h)_G^* \right).
\]

(5.69)

\[
= \text{tr} \left( \sum_{g \in G} \sum_{h \in G} a_g b_h [\lambda(g)]_G [\lambda(h)]_G^* \right).
\]

(5.70)

\[
= \sum_{g \in G} \sum_{h \in G} a_g b_h \text{tr} ([\lambda(g)]_G [\lambda(h^{-1})]_G).
\]

(5.71)

\[
= \sum_{g \in G} \sum_{h \in G} a_g b_h \text{tr} ([\lambda(g) \lambda(h^{-1})]_G).
\]

(5.72)

\[
= \sum_{g \in G} \sum_{h \in G} a_g b_h \text{tr} ([\lambda(gh^{-1})]_G).
\]

(5.73)

\[
= \sum_{g \in G} \sum_{h \in G} a_g b_h \text{tr} ([\lambda(gh^{-1})]_G).
\]

(5.74)

where (5.69) follows because $\lambda$ is a homomorphism and (5.70) follows because $[-]_G$ is a homomorphism. Furthermore, in (5.71) the summations may be rearranged since $G$ is finite, (5.72) follows because tr is linear, (5.73) follows because $[-]_G$ is a homomorphism, and (5.74) follows because $\lambda$ is a homomorphism.
By Lemma 5.4.5,

\[
\text{tr}([\lambda(gh^{-1})]_{G}) = \begin{cases} |G| & \text{if } gh^{-1} = 1_G \\ 0 & \text{otherwise} \end{cases}
\]

\[
= \begin{cases} |G| & \text{if } g = h \\ 0 & \text{otherwise.} \end{cases}
\]

(5.75)

Thus,

\[
\text{tr}([\lambda(a)]_{G}[\lambda(b)]_{G}^*) = \sum_{g \in G} \sum_{h \in G} a_g b_h \text{tr}([\lambda(gh^{-1})]_{G}) = \sum_{g \in G} a_g b_g |G|.
\]

Recall that \(\langle a, b \rangle = \sum_{g \in G} a_g b_g\). It follows that \(\langle a, b \rangle = \frac{1}{|G|} \text{tr}([\lambda(a)]_{G}[\lambda(b)]_{G}^*)\).

\(\square\)

**Theorem 5.4.7.** Let \(G\) be a finite group and \(V\) a left \(\mathbb{C}[G]\)-module such that \(\dim \mathbb{C} V < \infty\). Then \(V\) is a left Hilbert \(N(G)\)-module and \(\dim_{N(G)} V = \frac{1}{|G|} \dim \mathbb{C} V\).

Note that the latter part of Theorem 5.4.7, \(\dim_{N(G)} V = \frac{1}{|G|} \dim \mathbb{C} V\), is taken from Luck [Lüc02, Note after definition 1.10].

**Proof.** Let \(k\) be the number of conjugacy classes of elements in \(G\). By Theorem 5.4.3 (1), since \(V\) is a finite dimensional left \(\mathbb{C}[G]\)-module, it can be expressed as the direct sum of irreducible left \(\mathbb{C}[G]\)-modules, that is, \(V = V_1 \oplus \cdots \oplus V_s\) for some \(s\), \(1 \leq s \leq k\). As such, it is enough to prove the theorem for an irreducible left \(\mathbb{C}[G]\)-module \(V_i\), for fixed \(i\) such that \(1 \leq i \leq k\), because

\[
\dim_{N(G)} V = \dim_{N(G)} (V_1 \oplus \cdots \oplus V_s)
\]

\[
= \dim_{N(G)} V_1 + \cdots + \dim_{N(G)} V_s, \quad (*)
\]

where (\(*)\) follows by Theorem 5.3.12.

We first show that \(V_i\) is a left Hilbert \(N(G)\)-module. By assumption, \(V_i\) is a left \(\mathbb{C}[G]\)-module. By Theorem 5.4.3 (2), we have that \(\|ga\|_2 = \|a\|_2\) for all \(a \in V_i\).
and $g \in G$. Finally, by Theorem 5.4.3 (4), we know that there exists an isometric injective left $\mathbb{C}[G]$-module homomorphism $f : V_i \to \mathbb{C}[G] = \ell^2(G)$.

Let $\lambda : \mathbb{C}[G] \to \mathbb{C}[G]$ be as in Theorem 5.4.6. Then by Lemma 5.3.10, we have

$$\dim_{\mathbb{C}[G]} V_i = \langle \text{pr}_{f(V_i)}(E_1), E_1 \rangle$$

$$= \langle \text{pr}_{f(V_i)}(1_G), 1_G \rangle$$

$$= \frac{1}{|G|} \text{tr} (\lambda_{\text{pr}_{f(V_i)}(1_G)} |_{\mathbb{C}[G]})$$

$$= \frac{1}{|G|} \text{tr} (\lambda_{1_G})$$

(5.78)

where (*) follows by Theorem 5.4.6, and (**) follows because $[\lambda_{1_G}]_G$ is the identity matrix.

Let $n_j$ be the complex dimension of $V_j$, $1 \leq j \leq k$. Then by Theorem 5.4.3 (3), we have that $\mathbb{C}[G]$ is ring isomorphic to $\bigoplus_{j=1}^k M_{n_j}(\mathbb{C})$ via an isomorphism $\mu$, where $\mu(1_G) = (I_{n_1}, \ldots, I_{n_k})$. Furthermore, for $A_j \in M_{n_j}(\mathbb{C})$,

$$\text{pr}_{(\mu(f))} (V_j) (A_1, \ldots, A_k) = (0, \ldots, 0, \hat{A}_j, 0, \ldots, 0),$$

(5.79)

where $\hat{A}_j$ is the matrix whose first column is the first column of $A_j$ and the rest of whose entries are 0.

Let $c = \text{pr}_{f(V_i)} (1_G) \in \mathbb{C}[G]$. Using the fact that $\mu$ is an isomorphism and that $\mu(1_G) = (I_{n_1}, \ldots, I_{n_k})$, we can define

$$a = \mu(c) = \text{pr}_{\mu(f(V_i))} (\mu(1_G)) = \text{pr}_{\mu(f(V_i))} (I_{n_1}, \ldots, I_{n_k})$$

$$= (0, \ldots, 0, \underbrace{E_{11}}_{\text{th position}}, 0, \ldots, 0).$$

(5.80)

Furthermore, since $\mu$ is an isomorphism,

$$\text{tr} ([\lambda_c]_G) = \text{tr} ([\lambda_{\mu(c)}]_{\mu(G)}) = \text{tr} ([\lambda_a]_{\mu(G)}).$$

(5.81)
Note that we can choose another basis for $\bigoplus_{j=1}^{k} M_{n_j}(\mathbb{C})$ (aside from $\mu(G)$) to calculate the matrix representation in (5.81) because the trace will remain the same regardless of which basis is used to calculate the matrix.

We can define a basis for $\bigoplus_{j=1}^{k} M_{n_j}(\mathbb{C})$ in the following manner. Let $0$ represent the $n_t \times n_t$ zero matrix, $1 \leq t \leq k$, $t \neq j$, and let $E_{\ell m}$ be the $n_j \times n_j$ matrix with a $1$ in the $\ell m$ position. Let

$$b_{\ell m} = (0, \ldots, 0, E_{\ell m}, 0, \ldots, 0). \quad (5.82)$$

Then $\beta = \{b_{\ell m}\}$ is a basis for $\bigoplus_{j=1}^{k} M_{n_j}(\mathbb{C})$.

Additionally,

$$\lambda_a(b_{\ell m}) = a(b_{\ell m})$$

$$= (0, \ldots, 0, E_{11}, 0, \ldots, 0)(0, \ldots, 0, E_{\ell m}, 0, \ldots, 0)$$

$$= \begin{cases} b_{\ell m} & \text{if } j = i, \ell = 1 \\ 0 & \text{otherwise (\star)} \end{cases} \quad (5.83)$$

where the $0$ in (\star) represents the zero $k$-tuple, $(0, 0, \ldots, 0) \in \bigoplus_{j=1}^{k} M_{n_j}(\mathbb{C})$.

If we let $w = |G| = |\beta|$, then the resulting matrix, $[\lambda_a]_{\beta}$, will be a $w \times w$ matrix. Now,

$$[\lambda_a]_{\beta} = \begin{bmatrix} [\lambda_a(b_{111})]_{\beta} & \ldots & [\lambda_a(b_{k,n_k,n_k})]_{\beta} \end{bmatrix}. \quad (5.84)$$

The column vector, $[\lambda_a(b_{\ell m})]_{\beta}$, will be non-zero if and only if $\lambda_a(b_{\ell m}) = b_{\ell m}$, in which case,

$$[\lambda_a(b_{\ell m})]_{\beta} = [b_{\ell m}]_{\beta} = e_{b_{\ell m}}, \quad (5.85)$$

where $e_{b_{\ell m}} \subseteq \{e_1, \ldots, e_w\}$ is the standard basis vector as per Definition 2.5.6.

Per (5.83), $\lambda_a(b_{\ell m}) = b_{\ell m}$ if and only if $j = i$ and $\ell = 1$. In other words, (5.85)
holds true for the set \( \{b_{1m}\} \subseteq \{b_{j\ell m}\} \), where \( 1 \leq m \leq n_i \). Therefore, we must have that \( |\{b_{1m}\}| = n_i \).

So, in conclusion, the matrix \([\lambda_a]_B\) will have exactly \( n_i \) diagonal entries that are equal to 1, while the rest of the diagonal entries will be 0. Thus,

\[
\text{tr}([\lambda_a]_B) = n_i = \dim_\mathbb{C} V_i,
\]

and hence, by (5.78),

\[
\dim_{\mathcal{N}(G)} V_i = \frac{1}{|G|} \text{tr}([\lambda_a]_B) = \frac{1}{|G|} \dim_\mathbb{C} V_i.
\]

The results of Theorem 5.4.7 shows that it is possible to have left Hilbert \( \mathcal{N}(G) \)-modules of fractional dimension. For example, let

\[
G = S_3 = \langle \sigma, \tau \mid 1 = \sigma^3 = \tau^2, \tau^{-1}\sigma \tau = \sigma^{-1} \rangle.
\]

We claim without proof that \( \mathbb{C}[S_3] \cong V_1 \oplus V_2 \oplus V_3 \oplus V_3 \), where each \( V_i \) is an irreducible \( \mathbb{C}[S_3] \)-module, and \( \dim_\mathbb{C} V_1 = 1 \), \( \dim_\mathbb{C} V_2 = 1 \), and \( \dim_\mathbb{C} V_3 = 2 \). (See Isaacs [Isa94, Corollary 1.17].)

We specifically examine the case of \( V_3 \). Define a representation \( \lambda : S_3 \to GL(V_3) \) by \( \lambda(\sigma) = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix} \) and \( \lambda(\tau) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \). Then by Theorem 2.7.4, \( \lambda \) gives \( V_3 \) the structure of a \( \mathbb{C}[S_3] \)-module. By Theorem 5.4.7 we have

\[
\dim_{\mathcal{N}(S_3)} V_3 = \frac{1}{|S_3|} \dim_\mathbb{C} V_3 = \frac{2}{6} = \frac{1}{3}.
\]

By Theorem 5.4.1, since \( |S_3| \) is finite, \( \mathcal{N}(S_3) \) may be used interchangeably with \( \mathbb{C}[S_3] \cong \ell^2(S_3) \). By the note after Theorem 5.4.3, \( V_3 \cong f(V_3) \) is a submodule of \( \mathbb{C}[S_3] \). Thus, if \( \dim_{\mathcal{N}[S_3]}(V_3) = \frac{1}{3} \), one possible interpretation of von Neumann dimension may be to think of \( V_3 \cong f(V_3) \) as "comprising" \( \frac{1}{3} \) of \( \mathcal{N}(S_3) \cong \mathbb{C}[S_3] = \ell^2(S_3) \).
5.5 Applications of Chapter 2

In section 5.4, we proved the existence of Hilbert \( \mathcal{N}(G) \)-modules with fractional dimension. In this final section, we use the results from section 3.6 of Chapter 3 to demonstrate the conditions under which modules over a ring \( \mathcal{U}(G) \) have integral dimension.

The following definition refers to Hungerford [Hun74, Chapter IV, section 3].

**Definition 5.5.1.** Let \( R \) be a ring, and let \( P \) be a left \( R \)-module. Then \( P \) is said to be **projective** if for all left \( R \)-modules \( A \) and \( B \) such that \( f : P \to B \), \( g : A \to B \) are left \( R \)-module homomorphisms and \( g \) is surjective, there exists a left \( R \)-module homomorphism \( h : P \to A \) such that \( gh = f \).

**Theorem 5.5.2.** Every free module over a ring \( R \) is projective.

*Proof.* See Hungerford [Hun74, Chapter IV, Theorem 3.2]. \( \square \)

**Theorem 5.5.3.** Let \( G \) be a countable group. Let \( \mathcal{P}(\mathcal{N}(G)) \) be the set of all finitely generated projective left \( \mathcal{N}(G) \)-modules. Then there exists a function

\[
\overline{\dim}_{\mathcal{N}(G)} : \mathcal{P}(\mathcal{N}(G)) \to [0, \infty)
\]

such that \( \overline{\dim}_{\mathcal{N}(G)}(P \oplus Q) = \overline{\dim}_{\mathcal{N}(G)}P + \overline{\dim}_{\mathcal{N}(G)}Q \) for every \( P, Q \) in the set \( \mathcal{P}(\mathcal{N}(G)) \).

*Proof.* See Lück [Lüc02, Theorem 6.5]. \( \square \)

**Theorem 5.5.4.** Let \( \mathcal{H}(\mathcal{N}(G)) \) be the set of all finitely generated left Hilbert \( \mathcal{N}(G) \)-modules. Then there exists a bijection \( F : \mathcal{P}(\mathcal{N}(G)) \to \mathcal{H}(\mathcal{N}(G)) \) such that \( \overline{\dim}_{\mathcal{N}(G)}P = \dim_{\mathcal{N}(G)}(F(P)) \). In particular, \( \overline{\dim}_{\mathcal{N}(G)}(\mathcal{N}(G)^n) = n \).

*Proof.* See Lück [Lüc02, Theorem 6.24].
We know that \( \overline{\text{dim}} \) is defined for \((\mathcal{N}(G))^n\) because \((\mathcal{N}(G))^n\) is a finitely generated left free \(\mathcal{N}(G)\)-module and therefore by Theorem 5.5.2, a projective module. Then, by Luck [Lüc02, (6.22)] and Theorem 5.3.11, we have that \( \overline{\text{dim}}_{\mathcal{N}(G)}(\mathcal{N}(G)^n) = n \).

The following definition refers to Hungerford [Hun74, Chapter IV, section 5].

**Definition 5.5.5.** Let \( R \) and \( S \) be rings. Let \( M \) be an \( R-S \) bimodule and let \( N \) be a left \( S \)-module. Let \( F \) be the free abelian group on the set \( M \times N \). Let \( K \) be the subgroup of \( F \) generated by elements of the following forms for all \( m, m' \in M, n, n' \in N, \) and \( r \in S \):

1. \( (m + m', n) - (m, n) - (m, n) \);
2. \( (m, n + n') - (m, n) - (m, n') \); and
3. \( (mr, b) - (m, rb) \).

Then the quotient group \( F/K \) is called the tensor product of \( M \) and \( N \) and is denoted \( M \otimes_S N \).

**Theorem 5.5.6.** Let \( R \) and \( S \) be rings, \( M \) an \( R-S \) bimodule, and \( N \) a left \( S \)-module. Then \( M \otimes_S N \) is a left \( R \)-module.

**Proof.** See Hungerford [Hun74, Chapter IV, Theorem 5.5]. □

**Theorem 5.5.7.** Let \( R \) be a ring and \( S \) a subring of \( R \). Then \( R \otimes_S S^n \cong R^n \).

**Proof.** See Hungerford [Hun74, Chapter IV, Proof of Theorem 5.11]. □

The following theorem refers to a ring \( \mathcal{U}(G) \). Note that the actual definition of the ring \( \mathcal{U}(G) \) is too technical to be included in this paper; however, Linnell [Lin06, Section 4], refers to \( \mathcal{U}(G) \) as the “ring of unbounded operators on \( \ell^2(G) \) affiliated to \( \mathbb{C}[G] \)".
Theorem 5.5.8. Let $G$ be a countable group. Then there exists a ring $\mathcal{U}(G)$ such that $\mathcal{N}(G) \subseteq \mathcal{U}(G)$. If $G$ is a finitely generated free group, then there exists a division ring $D(G)$ such that

$$\mathbb{C}[G] \leq D(G) \leq \mathcal{U}(G).$$

(5.91)

Proof. See Linnell [Lin06, Section 4] and Linnell [Lin93, Theorem 1.3].

Theorem 5.5.9. Let $\mathcal{U}(G)$ be defined as in Theorem 5.5.8. Then there exists a function

$$\dim_{\mathcal{U}(G)} : \{\text{finitely generated projective left } \mathcal{U}(G)\text{-modules}\} \to [0, \infty),$$

(5.92)

such that $\dim_{\mathcal{U}(G)}$ is additive. If $P$ is a finitely generated projective left $\mathcal{N}(G)$-module, then

$$\dim_{\mathcal{U}(G)}(\mathcal{U}(G) \otimes_{\mathcal{N}(G)} P) = \overline{\dim_{\mathcal{N}(G)}} P.$$  

(5.93)

In particular,

$$\dim_{\mathcal{U}(G)}(\mathcal{U}(G)^n) = n.$$  

(5.94)

Proof. See Luck [Lüc02, Lemma 8.27 and Theorem 8.29]. For (5.94), by Theorem 5.5.7, we have that $\mathcal{U}(G) \otimes_{\mathcal{N}(G)} \mathcal{N}(G)^n \cong \mathcal{U}(G)^n$. Therefore, by (5.93) and Theorem 5.5.4, we have

$$\dim_{\mathcal{U}(G)}(\mathcal{U}(G)^n) = \dim_{\mathcal{U}(G)}(\mathcal{U}(G) \otimes_{\mathcal{N}(G)} \mathcal{N}(G)^n) = \overline{\dim_{\mathcal{N}(G)}} \mathcal{N}(G)^n = n.$$  

(5.95)

To summarize our results thus far, in section 5.4, we proved the existence of a Hilbert $\mathcal{N}(G)$-module $M$ with fractional (non-integral) dimension. In Theorem 5.5.4, we established the existence of a projective $\mathcal{N}(G)$-module $P$ with the same dimension as $M$. Finally, by Theorem 5.5.9, $\mathcal{U}(G) \otimes_{\mathcal{N}(G)} P$, which is itself a left $\mathcal{U}(G)$-module, has the same (fractional) dimension as $P$. Thus, we have established the existence of $\mathcal{U}(G)$-modules of fractional dimension.
The following corollaries establish the conditions under which finitely generated left $\mathcal{U}(G)$-modules have integral dimension (over $\mathcal{U}(G)$).

**Corollary 5.5.10.** Let $G$ be a finitely generated free group. Let $M$ be a free $\mathcal{U}(G)$-module with basis $\beta = \{u_1, \ldots, u_m\}$, and $N = \langle v_1, \ldots, v_n \rangle$ be a submodule of $M$. Suppose that

$$\mathcal{A} = \begin{bmatrix} [v_1]_{\beta} \\ \vdots \\ [v_n]_{\beta} \end{bmatrix} \in M_{n \times m}(\mathbb{C}[G]) \subseteq M_{n \times m}(D(G)).$$

(5.96)

Then

(1) $N$ is a finitely generated free left $\mathcal{U}(G)$-module.

(2) $M/N$ is a finitely generated free left $\mathcal{U}(G)$-module.

**Proof.** Since $\mathcal{A} \in M_{n \times m}(D(G))$ and $D(G)$ is a division ring, by Theorem 3.4.10, we have that $\mathcal{A}$ is $D$-row reducible and therefore $\mathcal{U}(G)$-row reducible. The result follows immediately by Theorem 3.6.1. □

**Corollary 5.5.11.** Let $G$ be a finitely generated free group. Let $T : M \rightarrow N$ be a left $\mathcal{U}(G)$-module homomorphism, and let $\mathcal{A} = [T]^{\beta}_{\alpha}$. Assume that $\mathcal{A} \in M_{n \times m}(\mathbb{C}[G]) \subseteq M_{n \times m}(D(G))$. Then $\ker(T)$ is a finitely generated free left $\mathcal{U}(G)$-module.

**Proof.** Since $\mathcal{A} \in M_{n \times m}(D(G))$ and $D(G)$ is a division ring, by Theorem 3.5.9, we have that $\mathcal{A}$ is $D$-column reducible and therefore $\mathcal{U}(G)$-column reducible. The result follows immediately from Theorem 3.6.3. □

Note that since $N$, $M/N$, and $\ker(T)$, (as defined in Corollaries 5.5.10 and 5.5.11), are finitely generated free left $\mathcal{U}(G)$-modules, it follows by Theorem 2.5.4 that $N$, $M/N$, and $\ker(T)$ are each left $\mathcal{U}(G)$-module isomorphic to a direct sum of
copies of $U(G)$. That is, $N \cong (U(G))^k$, $M/N \cong (U(G))^\ell$, and Ker($T$) $\cong (U(G))^s$, where $k, \ell, s \in \mathbb{Z}$ represent the number of elements in a basis for $N$, $M/N$, and Ker($T$), respectively. By Theorem 5.5.9, since $\dim(U(G)^n) = n$, it follows that $N$, $M/N$, and Ker($T$) have integral dimension over $U(G)$. 


