Finitely generated modules

A Project Report Submitted

for the Course

MA498 Project I

by

Subhash Atal

(Roll No. 07012321)

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781039, INDIA

November 2014

CERTIFICATE

This is to certify that the work contained in this project report entitled "**Finitely generated modules**" submitted by **Subhash Atal** (**Roll No.: 07012321**) to Department of Mathematics, Indian Institute of Technology Guwahati towards the requirement of the course **MA498 Project I** has been carried out by him/her under my supervision.

Guwahati - 781 039	(Dr. Shyamashree Upadhyay)
November 2014	Project Supervisor

ABSTRACT

In mathematics, we often come across finite dimensional vector spaces and finitely generated abelian groups. These are natural examples of finitely generated modules. The main aim of the project is to understand the structure of finitely generated modules and to prove a basic lemma about them called the Nakayama lemma.

Contents

Chapter 1

Introduction

One of the things which distinguishes the modern approach to Commutative Algebra is the greater emphasis on modules, rather than just on ideals. An ideal $\mathfrak{a}$ and its quotient ring $A/\mathfrak{a}$ are both examples of modules. The collection of all modules over a given ring contains the collection of all ideals of that ring as a subset. The concept of modules is in fact a generalization of the concept of ideals. In this chapter, we give the definition and elementary properties of modules.

Throughout this report, let A denote a commutative ring with unity 1.

1.1 Modules and module homomorphisms

Definition 1.1.1. An A-*module* is an abelian group M (written additively) on which A acts linearly: more precisely, an A-*module* is a pair (M, μ), where M is an abelian group and μ is a mapping from $A \times M$ into M such that for all $a, b \in A$ and for all $x, y \in M$, the following axioms are satisfied:

$$\mu(a, x + y) = \mu(a, x) + \mu(a, y)$$
$$\mu(a + b, x) = \mu(a, x) + \mu(b, x)$$

$$\mu(ab, x) = \mu(a, bx)$$
$$\mu(1, x) = x$$

Equivalently, M is an abelian group together with a ring homomorphism $A \to E(M)$, where $E(M)$ denotes the ring of all endomorphisms of the abelian group M, the ring homomorphism $A \to E(M)$ being given by $a \mapsto \mu(a, .)$.

The notation ax is more generally used for $\mu(a, x)$.

Examples: 1) An ideal $\mathfrak{a}$ of A is an A-module. In particular, A itself is an A-module.
2) If A is a field k, then A-module= k-vector space.
3) If $A = \mathbb{Z}$, then $\mathbb{Z}$-module=abelian group (define nx to be $x + \cdots + x$).
4) If $A = k[X]$ where k is a field, then an A-module M is a k-vector space together with a linear transformation. [Since M is an A-module, there exists a ring homomorphism $A \to E(M)$. The image of the element $X(\in A)$ in $E(M)$ under this ring homomorphism is a linear transformation from M to itself. This is the required linear transformation.]

Definition 1.1.2. Let M, N be A-modules. A mapping $f : M \to N$ is called an A-*module homomorphism* if

$$f(x + y) = f(x) + f(y)$$
$$f(ax) = a.f(x)$$

for all $a \in A$ and all $x, y \in M$.

Thus f is a homomorphism of abelian groups which commutes with the action of each $a \in A$. If A is a field, then a A-module homomorphism is the same thing as a linear transformation of vector spaces.

The composition of A-module homomorphisms is again an A-module homomorphism.

The set of all A-module homomorphisms from M to N can be turned into an A-module as follows: For any two A-module homomorphisms f and g (from M to N), we define $f + g$ and af by the rules

$$(f+g)(x) = f(x) + g(x)$$
$$(af)(x) = a.f(x)$$

for all $x \in M$. It is trivial to check that all the axioms for an A-module are satisfied. This A-module is denoted by $Hom_A(M, N)$ (or just $Hom(M, N)$ if there is no ambiguity about the ring A).

Homomorphisms $u : M' \to M$ and $v : N \to N''$ induce mappings $\bar{u} : Hom(M, N) \to Hom(M', N)$ and $\bar{v} : Hom(M, N) \to Hom(M, N'')$ defined as follows:

$$\bar{u}(f) = f \circ u \text{ and } \bar{v}(f) = v \circ f.$$

These mappings are A-module homomorphisms.

For any A-module M, there is a natural isomorphism $Hom(A, M) \cong M$: any A-module homomorphism $f : A \to M$ is uniquely determined by $f(1)$, which can be any element of M.

1.2 Submodules and quotient modules

Definition 1.2.1. A *submodule* M' of an A-module M is a subgroup of M which is closed under multiplication by elements of A.

Definition 1.2.2. Let M be an A-module and M' be a submodule of M. The abelian group M/M' then inherits an A-module structure from M, defined by

$$a(x + M') = ax + M'$$

The A-module M/M' is called the *quotient* of M by M'.

The natural map π from M onto M/M' given by $x \mapsto x + M'$ is an A-module homomorphism. There is a one-to-one order-preserving correspondence between submodules of M/M' and submodules of M which contain M', given by $U \mapsto \pi^{-1}(U)$ for any submodule U of M/M'.

Definition 1.2.3. If $f : M \to N$ is an A-module homomorphism, the *kernel* of f is the set

$$Ker(f) = \{x \in M : f(x) = 0\}$$

and is a submodule of M.

Definition 1.2.4. If $f : M \to N$ is an A-module homomorphism, the *image* of f is the set

$$Im(f) = \{f(x) : x \in M\} = f(M)$$

and is a submodule of N.

Definition 1.2.5. If $f : M \to N$ is an A-module homomorphism, the *cokernel* of f is

$$Coker(f) = N/Im(f)$$

which is a quotient module of N.

Let M, N be two A-modules and $f : M \to N$ be an A-module homomorphism. If M' is a submodule of M such that $M' \subseteq Ker(f)$, then f gives rise to a A-module homomorphism $\bar{f} : M/M' \to N$, defined as follows:

$$\bar{f}(x + M') := f(x)$$

This map $\bar{f}$ is well-defined because if $x + M' = y + M'$, then $x - y \in M'$ and $M' \subseteq Ker(f)$, hence $f(x - y) = 0$, that is $f(x) = f(y)$, which in turn implies that $\bar{f}(x + M') = \bar{f}(y + M')$. The kernel of $\bar{f}$ is $Ker(f)/M'$ and this map $\bar{f}$ is onto $Im(f)$. The homomorphism $\bar{f}$ is said to be *induced* by f. In particular, taking $M' = Ker(f)$, we have an isomorphism of A-modules

$$M/Ker(f) \cong Im(f). \tag{1.1}$$

The above equation is also known as the *first isomorphism theorem.*

1.3 Operation on submodules

Definition 1.3.1. Let M be an A-module and let $\{M_i\}_{i\in I}$ be a family of submodules of M. Their *sum* $\Sigma_{i\in I} M_i$ is the set of all finite sums $\Sigma_{i\in I} x_i$, where $x_i \in M_i$ for all $i \in I$ and almost all the x_i (that is, all but a finite number) are zero.

It is easy to check that $\Sigma_{i\in I} M_i$ is the smallest submodule of M which contains all the M_i. The *intersection* $\bigcap_{i\in I} M_i$ is again a submodule of M.

Theorem 1.3.2. *(i) If $L \supseteq M \supseteq N$ are A-modules, then*

$$(L/N)/(M/N) \cong L/M.$$

(ii) If M_1, M_2 are submodules of M, then

$$(M_1 + M_2)/M_1 \cong M_2/(M_1 \cap M_2).$$

Proof. (i) Define $\theta : L/N \to L/M$ by $\theta(x + N) = x + M$. Then θ is a well-defined A-module homomorphism of L/N onto L/M and its kernel is M/N. Hence the proof follows from the first isomorphism theorem (equation 1.1).
(ii) Consider the inclusion homomorphism $\iota : M_2 \to M_1 + M_2$ given by $m_2 \mapsto m_2$. Also consider the homomorphism $\pi : M_1 + M_2 \to (M_1 + M_2)/M_1$ given by $x \mapsto x + M_1$. The composition $\pi \circ \iota : M_2 \to (M_1 + M_2)/M_1$ is a module homomorphism. It is surjective, since π is surjective. And the kernel of $\pi \circ \iota$ is $M_1 \cap M_2$. Hence the proof follows from the first isomorphism theorem (equation 1.1). □

Definition 1.3.3. Let M be an A-module and let $\mathfrak{a}$ be an ideal of A. The *product* $\mathfrak{a}M$ is the set of all finite sums $\Sigma a_i x_i$ with $a_i \in \mathfrak{a}$ and $x_i \in M$.

It can be easily checked that $\mathfrak{a}M$ is a submodule of M.

Definition 1.3.4. Let M be an A-module and let N, P be submodules of M. We define $(N : P)$ to be the set of all $a \in A$ such that $aP \subseteq N$.

It can be easily checked that $(N : P)$ is an ideal of A.

Definition 1.3.5. Let M be an A-module. $(0 : M)$ is the set of all $a \in A$ such that $aM = 0$, this ideal is called the *annihilator* of M and is denoted by $Ann(M)$.

Definition 1.3.6. An A-module M is called *faithful* if $Ann(M) = 0$.

If $\mathfrak{a}$ is an ideal of A such that $\mathfrak{a} \subseteq Ann(M)$, then we may regard M as an $A/\mathfrak{a}$-module, as follows:

For any $x + \mathfrak{a} \in A/\mathfrak{a}$ and $m \in M$, define

$$(x + \mathfrak{a})m := xm.$$

Observe here that if $x + \mathfrak{a} = y + \mathfrak{a}$, then $x - y \in \mathfrak{a} \subseteq Ann(M)$. Hence for any $m \in M$, $(x - y)m = 0$, that is, $xm = ym$.
If $Ann(M) = \mathfrak{a}$, then M is faithful as an $A/\mathfrak{a}$-module.

1.4 Direct sum and product

Definition 1.4.1. Let $\{M_i\}_{i\in I}$ be a family of A-modules. Their *direct sum* $\oplus_{i\in I} M_i$ is the set of all tuples $(x_i)_{i\in I}$ such that $x_i \in M_i$ for all $i \in I$ and all but finitely many x_i are 0. This set $\oplus_{i\in I} M_i$ has a natural structure of an A-module given by:

$$(x_i)_{i\in I} + (y_i)_{i\in I} = (x_i + y_i)_{i\in I}$$

$$a(x_i)_{i\in I} = (ax_i)_{i\in I}$$

for all $a \in A$ and for all $(x_i)_{i\in I}, (y_i)_{i\in I} \in \oplus_{i\in I} M_i$.

Definition 1.4.2. Let $\{M_i\}_{i\in I}$ be a family of A-modules. Their *direct product* $\Pi_{i\in I} M_i$ is the set of all tuples $(x_i)_{i\in I}$ such that $x_i \in M_i$ for all $i \in I$. This set $\Pi_{i\in I} M_i$ has a natural structure of an A-module given by:

$$(x_i)_{i\in I} + (y_i)_{i\in I} = (x_i + y_i)_{i\in I}$$

$$a(x_i)_{i\in I} = (ax_i)_{i\in I}$$

for all $a \in A$ and for all $(x_i)_{i\in I}, (y_i)_{i\in I} \in \Pi_{i\in I} M_i$.

Observe that the difference between the above two definitions is that in the definition of the direct product, we have dropped the condition that “all but finitely many x_i are 0". Direct sum and direct product are therefore the same if the index set I is finite, but not otherwise, in general.

Theorem 1.4.3. *Let A be a ring. Then A is isomorphic to a direct sum of finitely many ideals of A if and only if A is isomorphic to a direct product of finitely many rings.*

Proof. Suppose $A \cong \Pi_{i=1}^{n} A_i$ where A_i are rings. For each $i \in \{1, \ldots, n\}$, define

$$\mathfrak{a}_i := \{(0, \ldots, 0, a_i, 0, \ldots, 0) \in \Pi_{i=1}^{n} A_i | a_i \in A_i\}.$$

Then each $\mathfrak{a}_i$ is an ideal of A. The ring $A \cong \Pi_{i=1}^{n} A_i$ can be considered as an A-module. Since each $\mathfrak{a}_i$ is an ideal of A, therefore each $\mathfrak{a}_i$ is an A-module, hence the direct sum $\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_n$ has also the structure of an A-module. Consider the map $\phi : \Pi_{i=1}^{n} A_i \to \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_n$ given by

$$\phi(a_1, \ldots, a_n) := ((a_1, 0, \ldots, 0), \ldots, (0, \ldots, 0, a_i, 0, \ldots, 0), \ldots, (0, \ldots, 0, a_n)).$$

It is easy to verify that ϕ is an isomorphism of A-modules.

Conversely, suppose $A \cong \mathfrak{I}_1 \oplus \cdots \oplus \mathfrak{I}_n$ where $\mathfrak{I}_1, \ldots, \mathfrak{I}_n$ are ideals of A. For each $i \in \{1, \ldots, n\}$, define $\mathfrak{b}_i := \oplus_{j \neq i} \mathfrak{I}_j$. For each $i \in \{1, \ldots, n\}$, Consider the map $f_i : A = \mathfrak{I}_1 \oplus \cdots \oplus \mathfrak{I}_n \to \mathfrak{I}_i$ given by $f_i(x_1, \ldots, x_n) := x_i$. Clearly each f_i is an onto A-module homomorphism, whose kernel is $\mathfrak{b}_i$. Therefore, by the first isomorphism theorem (equation 1.1), we have $A/\mathfrak{b}_i \cong \mathfrak{I}_i$. Since A is a ring, therefore each $A/\mathfrak{b}_i$ is a ring, call it C_i. Now since $A \cong \mathfrak{I}_1 \oplus \cdots \oplus \mathfrak{I}_n$ and $C_i \cong \mathfrak{I}_i$ for each $i \in \{1, \ldots, n\}$, we have $A \cong C_1 \oplus \cdots \oplus C_n$. But for finite indexing sets, direct sums and direct products are the same. Therefore we have $C_1 \oplus \cdots \oplus C_n = \Pi_{i=1}^{n} C_i$ and hence $A \cong \Pi_{i=1}^{n} C_i$. □

Chapter 2

Finitely generated modules

In this chapter, we will study the structure of finitely generated modules and prove a basic lemma about them called the Nakayama lemma.

2.1 Free modules

Definition 2.1.1. A *free* A-module is an A-module which is isomorphic to an A-module of the form $\oplus_{i \in I} M_i$, where each $M_i \cong A$ (as an A-module). Such a module is denoted by $A^{(I)}$.

Definition 2.1.2. Let n be a fixed non-negative integer. A *free* A-module of *rank* n is an A-module that is isomorphic to $A \oplus \cdots \oplus A$ (n summands), which is denoted by A^n. The integer n is called the *rank* of the free module A^n. (Conventionally, A^0 is the zero module.)

Definition 2.1.3. An ideal I in a commutative ring R is called a *principal ideal* if there exists some $a \in R$ such that $I = \{ra | r \in R\}$.

Definition 2.1.4. An integral domain R is called a *principal ideal domain* if every ideal in R is a principal ideal.

Theorem 2.1.5. *Let A be a principal ideal domain and M be a free module of rank n over A. Let N be a submodule of M. Then N is also free of rank $\leq n$.*

Proof. The proof is by induction on n. If $n = 1$, we have $M \cong A$ as an A-module and N is isomorphic to an ideal I of A. Since A is a principal ideal domain, therefore $I = \{ax | a \in A\}$ for some $x \in A$. If $x = 0$, then $I = 0$ and hence I is free of rank $0 \leq 1$. If $x \neq 0$, then $I \cong A$ (the map $\phi : A \to I$ given by $a \mapsto ax$ being an isomorphism), and hence I is free of rank 1. This proves the base case of induction. Assume now that the theorem is true for $n-1$ and consider a submodule N of M. Let $x_1, \ldots, x_n$ be a basis of M. Let

$$J := \{a_n \in A | a_1x_1 + \ldots + a_nx_n \in N, a_i \in A\}.$$

Since N is a submodule of M, therefore J is an ideal of A. Since A is a principal ideal domain, $J = \{ad | a \in A\}$ for some $d \in A$.

If $d = 0$, then $N \subseteq < x_1, \ldots, x_{n-1} >$, the free submodule generated by $x_1, \ldots, x_{n-1}$ and by induction, N is free of rank $\leq n-1$. If $d \neq 0$, choose an element $y \in N$ such that $y = b_1x_1 + b_2x_2 + \cdots + b_{n-1}x_{n-1} + dx_n$. We claim that $N = N\cap < x_1, \ldots, x_{n-1} > \oplus Ay$.

Let $x \in N$ so that $x = c_1x_1 + \ldots + c_nx_n$. Then $c_n = \lambda d$ for some $\lambda \in A$. Then $x - \lambda y = x' \in N\cap < x_1, \ldots, x_{n-1} >$. So $x = x' + \lambda y$. To prove the uniqueness, let $z \in N\cap < x_1, \ldots, x_{n-1} > \cap Ay$ so that $z = a_1x_1 + \ldots + a_{n-1}x_{n-1} = \mu(b_1x_1 + \ldots + b_{n-1}x_{n-1} + dx_n)$ for some $\mu \in A$. Since $x_1, \ldots, x_n$ are linearly independent, the coefficient of x_n, that is, $\mu d = 0$. This implies $\mu = 0$, that is, $z = 0$. Hence $N = N\cap < x_1, \ldots, x_{n-1} > \oplus Ay$. By induction, $N\cap < x_1, \ldots, x_{n-1} >$ is free of rank $\leq n-1$ so that N is free of rank $\leq n$. □

2.2 Structure of finitely generated modules

Definition 2.2.1. Let M be an A-module and let $x \in M$. The set of all multiples $ax(a \in A)$ is a submodule of M, denoted by Ax or $< x >$. This submodule of M is called the *cyclic submodule* of M *generated by* x.

Definition 2.2.2. Let M be an A-module and let $x_i \in M$ for all $i \in I$, where I is some indexing set. If $M = \Sigma_{i \in I} Ax_i$, then the set $\{x_i | i \in I\}$ is called a *set of generators* of M.
This means that every element of M ca be expressed (not necessarily uniquely) as a finite linear combination of the x_i with coefficients in A.

Definition 2.2.3. An A-module M is said to be *finitely generated* if it has a finite set of generators.

Examples: 1) Any finite dimensional vector space over a field k is a finitely generated k-module.
2) Any finitely generated abelian group is a finitely generated $\mathbb{Z}$-module. In particular, finite abelian groups are finitely generated $\mathbb{Z}$-modules.
3) The module of polynomials in one variable x over the ring A of degree at most n is a finitely generated A-module. This module is generated by $1, x, x^2, \ldots, x^n$.

Theorem 2.2.4. *M is a finitely generated A-module if and only if M is isomorphic to a quotient of A^n for some integer $n > 0$.*

Proof. Suppose M is a finitely generated A-module. Let $x_1, \ldots, x_n$ be generators of M. Define $\phi : A^n \to M$ as

$$\phi(a_1, \ldots, a_n) := a_1x_1 + \cdots + a_nx_n.$$

Clearly then ϕ is an A-module homomorphism onto M. Hence by the first isomorphism theorem (equation 1.1), we have $M \cong A^n/Ker(\phi)$. That is, M is isomorphic to a quotient of A^n.

Conversely, suppose $M \cong A^n/B$ for some integer $n > 0$ and for some submodule B of A^n. Then we have a natural A-module homomorphism π of A^n onto $A^n/B \cong M$, given by $\pi(a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_n) + B$. Let $e_i :=$ $(0, \ldots, 0, 1, 0, \ldots, 0)$ (the 1 being at the i-th place). Then the $e_i(1 \leq i \leq n)$ generate A^n. Hence $\pi(e_i)(1 \leq i \leq n)$ generate M over A. That is, M is a finitely generated A-module. □

Theorem 2.2.5. *Let A be a principal ideal domain and M be a finitely-generated A-module. Then any submodule of M is also a finitely generated module over A.*

Proof. Let N be a submodule of M. Since M is a finitely generated A-module, we have from theorem 2.2.4 that M is isomorphic to a quotient of A^n for some integer $n > 0$, say, $M \cong A^n/K$. Hence $N \cong F'/K$ where F' is a submodule of the free module A^n. By theorem 2.1.5, we have that F' is also free over A of rank $\leq n$. Hence N is a quotient of a free A-module by K. Therefore, applying theorem 2.2.4 again, we get that N is finitely generated. □

Theorem 2.2.6. *Let M be a finitely generated A-module, let $\mathfrak{a}$ be an ideal of A, and let ϕ be an A-module endomorphism of M such that $\phi(M) \subseteq \mathfrak{a}M$. Then ϕ satisfies an equation of the form*

$$\phi^n + a_1\phi^{n-1} + \cdots + a_n = 0$$

where the a_i are in $\mathfrak{a}$.

Proof. Let $x_1, \ldots, x_n$ be a set of generators of M. Then since $\phi(M) \subseteq \mathfrak{a}M$, we have $\phi(x_i) \in \mathfrak{a}M$ for each $i \in \{1, \ldots, n\}$. Hence for each $i \in \{, \ldots, n\}$, we have $\phi(x_i) = \Sigma_{j=1}^{n} a_{ij} x_j$ for some $a_{ij} \in \mathfrak{a}$. That is, we have the following system of equations:

$$\Sigma_{j=1}^{n} (\delta_{ij}\phi - a_{ij}) x_j = 0 \;\; \forall i \in \{1, \ldots, n\}$$

where δ_{ij} is the kronecker delta. Let D denote the $n \times n$ matrix whose (i, j)-th entry is $\delta_{ij}\phi - a_{ij}$. Then the above system of equations is the same as the matrix equation $D\overline{X} = \overline{0}$, where $\overline{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\overline{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. By multiplying on the left of the equation $D\overline{X} = \overline{0}$ by the adjoint of the matrix D, we get $det(D)I\overline{X} = \overline{0}$. This implies that the linear operator $det(D)I$ annihilates each x_i. But the x_i generate M, hence $det(D)I$ is the zero endomorphism of M. That is, $det(D) = 0$. Expanding out the determinant of the matrix D, we have an equation of the required form. □

Corollary 2.2.7. *Let M be a finitely generated A-module and let $\mathfrak{a}$ be an ideal of A such that $\mathfrak{a}M = M$. Then there exists $x \equiv 1 (mod\ \mathfrak{a})$ such that $xM = 0$.*

Proof. Take ϕ =identity and $x = 1 + a_1 + \cdots + a_n$ in theorem 2.2.6. □

2.3 The Nakayama lemma

In this section, we will prove the famous Nakayama lemma. But for its proof, we need some preliminaries on ideals, which we provide first.

Definition 2.3.1. An ideal $\mathfrak{m}$ in A is called *maximal* if $\mathfrak{m} \neq A$ and if there is no ideal $\mathfrak{a}$ of A such that $\mathfrak{m} \subset \mathfrak{a} \subset A$ (strict inclusions).

Theorem 2.3.2. *Let $x \in A$ and $Ax := \{ax|a \in A\}$. Then Ax is an ideal in A and $Ax = A$ if and only if x is a unit in A.*

Proof. It is easy to verify that Ax is an ideal in A. If x is a unit in A, then there exists $u \in A$ such that $ux = 1$. So $1 \in Ax$. Now Ax is an ideal in A and $1 \in Ax$ together imply that $A \subseteq Ax$. Hence $Ax = A$.

Conversely, suppose $Ax = A$ for some $x \in A$. Then since $1 \in A$, we have $1 \in Ax$. This implies that 1 equals vx for some $v \in A$, which in turn implies that x is a unit in A. □

Theorem 2.3.3. *Let $\mathfrak{m}$ be a maximal ideal of A and $x \in \mathfrak{m}$. Then x is not a unit in A.*

Proof. Suppose not, that is, suppose x is a unit in A. Then there exists $y \in A$ such that $xy = 1$. Now $x \in \mathfrak{m}$ and $\mathfrak{m}$ is an ideal of A. Therefore, $xy \in \mathfrak{m}$, which implies that $1 \in \mathfrak{m}$. But the facts that $1 \in \mathfrak{m}$ and $\mathfrak{m}$ is an ideal in A, together imply that $\mathfrak{m} = A$, which is absurd. Hence a contradiction. □

The proof of theorem 2.3.8 below requires the use of a well known lemma in set theory called Zorn's lemma, which we need to state first. But for stating Zorn's lemma, we need some definitions, which we provide first:

Definition 2.3.4. Let S be a non-empty set. A binary relation $\prec$ on S is called a *partial order* if $\prec$ is reflexive, antisymmetric and transitive. The set S together with a partial order $\prec$ is called a *partially ordered set.*

Definition 2.3.5. Let $S, \prec$ be a partially ordered set. A subset T of S is called a *chain* of S if for any two elements $a, b \in T$, we have either $a \prec b$ or $b \prec a$.

Definition 2.3.6. Let $S, \prec$ be a partially ordered set and Σ be a subset of S. An element $x \in S$ is called an *upper bound* for Σ if $a \prec x$ for all $a \in \Sigma$.

Definition 2.3.7. Let $S, \prec$ be a partially ordered set. An element $a \in S$ is called a *maximal element* of S if there is no element $b \in S$ such that $a \prec b$.

We are now ready to state the Zorn's lemma.

Zorn's lemma: Let S be a non-empty partially ordered set. If every chain T of S has an upper bound in S, then S has at least one maximal element.

Theorem 2.3.8. *Every non-unit of A is contained in some maximal ideal of A.*

Proof. Let x be an non-unit in A. Then consider the ideal Ax in A. By theorem 2.3.2, we have $Ax \neq A$. Let Σ_x be the set of all ideals $\mathfrak{a}$ in A such that $\mathfrak{a} \neq A$ and $Ax \subseteq \mathfrak{a}$. Then Σ_x is a partially ordered set with respect to the partial order $\subseteq$. The collection Σ_x is also non-empty because $Ax \in \Sigma_x$. The proof will be over if we can show that there exists a maximal ideal of A containing the ideal Ax. In other words, we need to show that the set Σ_x has a maximal element. For showing this, we will apply Zorn's lemma.

To apply Zorn's lemma, we must show that every chain in Σ_x has an upper bound in Σ_x. Let $\{\mathfrak{a}_\alpha\}_{\alpha \in I}$ be an arbitrary chain in Σ_x (where I is an indexing set). Let $\mathfrak{b} := \cup_{\alpha \in I} \mathfrak{a}_\alpha$. Since $\{\mathfrak{a}_\alpha\}_{\alpha \in I}$ is a chain, it follows that $\mathfrak{b}$ is an ideal in A. Also $1 \notin \mathfrak{b}$ since $1 \notin \mathfrak{a}_\alpha$ for all α [This is because if $1 \in \mathfrak{a}_\alpha$ for some α, then since $\mathfrak{a}_\alpha$ is an ideal of A, we will have $\mathfrak{a}_\alpha = A$, which is a contradiction.]. Moreover, $Ax \subseteq \mathfrak{b}$ since $Ax \subseteq \mathfrak{a}_\alpha$ for all $\alpha \in I$. Hence $\mathfrak{b} \in \Sigma_x$, and $\mathfrak{b}$ is an upper bound of the chain $\{\mathfrak{a}_\alpha\}_{\alpha \in I}$. Therefore, by Zorn's lemma, Σ_x has a maximal element. □

Definition 2.3.9. The *Jacobson radical* $\mathfrak{R}$ of A is defined to be the intersection of all the maximal ideals of A.

Theorem 2.3.10. *Let $\mathfrak{R}$ be the Jacobson radical of A. Then $x \in \mathfrak{R}$ if and only if $1 - xy$ is a unit in A for all $y \in A$.*

Proof. Suppose $1 - xy$ is not a unit in A for some $y \in A$. Then by theorem 2.3.8, we have that $1 - xy$ belongs to some maximal ideal $\mathfrak{m}$ of A. But $x \in \mathfrak{R} \subseteq \mathfrak{m}$, hence $xy \in \mathfrak{m}$. Therefore $1 = (1 - xy) + xy \in \mathfrak{m}$, which implies that $\mathfrak{m} = A$. But this is absurd since any maximal ideal is $\neq A$.
Conversely, suppose $x \notin \mathfrak{R}$. Then $x \notin \mathfrak{m}$ for some maximal ideal $\mathfrak{m}$. Then $\mathfrak{m}$ and x together generate the ideal $\mathfrak{m} + Ax$ which contains $\mathfrak{m}$ properly. Hence by maximality of $\mathfrak{m}$, we must have that $\mathfrak{m} + Ax = A$. So we have $1 = u + yx$ for some $u \in \mathfrak{m}$ and some $y \in A$. Hence $1 - yx = 1 - xy = u \in \mathfrak{m}$ and is therefore not a unit (by theorem 2.3.3). □

Theorem 2.3.11. *(Nakayama lemma) Let M be a finitely generated A-module and $\mathfrak{a}$ be an ideal of A contained in the Jacobson radical $\mathfrak{R}$ of A. Then $\mathfrak{a}M = M$ implies $M = 0$.*

Proof. By corollary 2.2.7, we have that $xM = 0$ for some $x \equiv 1(mod\ \mathfrak{R})$. So $x - 1 \in \mathfrak{R}$ and hence $1 - x = -(x - 1) \in \mathfrak{R}$. Now by 2.3.10, $1 - (1 - x)y$ is a unit in A for all $y \in A$. In particular, for $y = 1$, we have that $1 - (1 - x)1 = x$ is a unit in A. Since x is a unit in A, therefore x^{-1} exists in A. Hence $xM = 0$ implies that $x^{-1}xM = 0$, that is, $M = 0$. □

Bibliography

[1] M.F. Atiyah and I.G. Macdonald. Introduction to commutative algebra. pages 17–22. Addison-Wesley Publishing Company, 1969.

[2] N.S. Gopalakrishnan. University algebra. pages 159–166. New Age International (P) Limited, 2013.