Reg no----------

(16MMU403)

KARPAGAM ACADEMY OF HIGHER EDUCATION
Coimbatore-21
DEPARTMENT OF MATHEMATICS
Fourth Semester
I Internal Test - Jan'2018
Ring Theory and Linear Algebra-II

Date: -01-2018 **Time: 2 Hours**
Class: II-Bsc Mathematics **Maximum Marks:50**

PART-A(20X1=20 Marks)

Answer all the Questions:

1. A Polynomial $f(x) = x^2+(i+2)x+25$ is in_____
   (a) N[x] (b) R[x] (c) Q[x] **(d) C[x]**
2. Which one of the below is constant polynomial?
   (a) f(x)in F[x] (b) f(x) = 5x **(c) f(x) = 5** (d) f(x) = 2x
3. The greatest common divisor of all the coefficients of f(x) is called _____
   (a) **content** (b) primitive
   (c) zero polynomial (d) constant polynomial
4. If a polynomial is called primitive then _____
   (a) **coefficients GCD value is 1** (b) Deg(f(x)) = 1
   (c) Deg(f(x)) = 1 (d) coefficients GCD value is 0
5. The degree of $f(x) = (3x^2+5x-2)(x+8)$ is _____
   (a) 3 (b) 1 (c) 2 (d) 8
6. If deg(f(x)) = 5 and deg(g(x)) = 5 then deg (f(x).g(x)) is _____
   (a) 5 **(b) 10** (c) 25 (d) 5^5
7. Root of the polynomial $f(x) = x^2+2$ is in _____
   (a) N ` (b) Z (c) R (d) all the above
8. If F: $Z[x] \rightarrow Z_3[x]$ then the polynomial $f(x) = 5x^2+3x-8$ in Z[x] is transform to _____ in $Z_3[x]$
   (a) same **(b) $2(x^2-1)$** (c) $2(x^2+1)$ (d) not exist
9. An element a in F is a zero of f(x) in F[x] iff x-a is a _____ of f(x) in F[x].
   (a) factor (b) content (c) primitive (d) solution
10. The polynomial $f(x) = x^4+3x^3+2x+4$ in $Z_5[x]$ has a zero____ in Z_5
    (a) **1** (b) 6 (c) -1 (d) 11
11. If the polynomial $f(x) = x^2+3x+2$ is reducible over _____
    (a) Q ` (b) Z (c) R (d) **all the above**
12. The polynomial f(x) = a(x).b(x) is called irreducible then _____
    (a) Neither a(x) =1 nor b(x) = 1 (b) **either a(x) =1 or b(x) = 1**
    (c) both a(x) and b(x) non unit (d) deg(a(x)) > deg(f(x))
13. A polynomial f(x) is the product of two lower degree polynomial then f(x) is called _____
    (a) primitive (b) content **(c) reducible** (d) irreducible
14. From the division algorithm, f(x) = g(x) q(x) + r(x) where _____
    (a) r(x) =0
    (b) **r(x)=0 or deg (r(x)) < deg(g(x))**
    (c) deg (r(x)) < deg(g(x))
    (d) deg (r(x)) < deg(f(x))
15. Every polynomial can be expressed as the irreducible or product of irreducible polynomial in _____
    (a) unique way (b) multiple way
    (c) finite number of way (d) no way
16. In $Z_2[x]$ _____ is possible
    (a) $(x+1)^2 = x^2+1$ (b) x+5 = x+1 (c) 8=0 **(d) all the above**
17. A non zero polynomial f(x) in F[x] of degree n has _____zeros.
    (a) at least n (b) atmost n
    (c) equal to n (d) all of the three posible
18. If p(x) is irreducible and p(x)divides r(x) and s(x) then ___
    (a) either p(x) divides r(x) or p(x) divides s(x)
    (b) neither p(x) divides r(x) nor p(x) divides s(x)
    (c) p(x) divides r(x)
    (d) p(x) divides s(x)

19. If $f(x) = x^3+x+2$ then the polynomial is _____
    (a) **primitive** (b) degree 2 (c) content 3 (d) not in Q
20. Eisenstein criterion is the reducibility test for _____
    (a) rational (b) irrational (c) real numbers (d) all the above

PART-B (3X2=6 Marks)

Answer all the Questions:

21. Define Principal ideal domain.
22. Find all the zeros of $f(x) = x^4+4$ in Z_5
23. Define irreducible polynomial with example.

PART-C (3X8=24 Marks)

Answer all the Questions:

24. (a) State and prove division algorithm for polynomial.

    (OR)

    (b) The polynomial x^3+2x^2+2x+1 factored into the linear factor in $Z_7[x]$. Find this factorization.
25. (a) State and prove Eisenstein criterion.

    (OR)

    (b) State and prove factor theorem.

26. (a) State and prove Unique factorization theorem

    (OR)

    (b) Let f(x) and g(x) be any two polynomial then prove that deg(f(x).g(x)) = deg (f(x)) + deg (g(x)).

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

LECTURE PLAN
DEPARTMENT OF MATHEMATICS

STAFF NAME: U.R.RAMAKRISHNAN
SUBJECT NAME: Ring Theory and Linear Algebra-II SUB.CODE:16MMU403
SEMESTER: IV CLASS: I B.Sc Mathematics

S. No	Lecture Duration Hour	Topics To Be Covered	Support Materials
		UNIT-I	
S.NO	**DURATION HOURS**	**TOPICS TO BE COVERED**	**SUPPORT MATERIAL**
1	1	Definition and Examples of Ring	T1: chap-IV Pg.No:167-168
2	1	Polynomial rings over commutative rings	T1: chap-IV Pg.No:198-209
3	1	Tutorial	
4	1	Division algorithm for polynomial Ring.	T1: chap-IV Pg.No:210-212
5	1	Principal ideal domains	T1: chap-IX Pg.No:391-393
6	1	Tutorial	
7	1	Factorization of polynomials	T1: chap-V Pg.No:237-239
8	1	Theorems on Factorization of Polynomials	T1: chap-V Pg.No:242-243
9	1	Reducibility tests	T1: chap-IV Pg.No:214-216
10	1	Tutorial	
11	1	Irreducibility tests	T1: chap-IV Pg.No:216-218
12	1	Theorems for Reducibility and irreducibility	T1: chap-IV Pg.No:214-218
13	1	Continuous of theorems on irreducibility	T1: chap-IV Pg.No:218-220
14	1	Tutorial	

15	1	Eisenstein criterion	T1: chap-IV Pg.No:215-217
16	1	Unique factorization in Z[x].	T1: chap-IV Pg.No:217-218
17	1	Tutorial	
18	1	Recapitulation and Discussion of possible questions	
	Total No of Hours Planned For Unit 1=18		
		UNIT-II	
1	1	Divisibility in integral domains	T1: chap-IX Pg.No:388-389
2	1	Primes	T1: chap-IX Pg.No:394
3	1	Irreducible	T1: chap-IX Pg.No:389-390
4	1	Tutorial	
5	1	Examples of Prime and irreducible.	T1: chap-IX Pg.No:394-395
6	1	Theorems on prime and irreducible.	R3: chap-17 Pg.No:1063
7	1	Principle Ideal Domain	T1: chap-IX Pg.No:391-393
8	1	Tutorial	
9	1	Unique Factorization Domain	T1: chap-IX Pg.No:390-391
10	1	Theorems on PID and UFD	T1: chap-IX Pg.No:391-398
11	1	Corollary and Example of above theorem.	T1: chap-IX Pg.No:341-398
12	1	Tutorial	
13	1	Euclidean Domain definition with examples	T1: chap-IX Pg.No:401
14	1	Theorems on ED and PID	T1: chap-IX Pg.No:402-404
15	1	Tutorial	
16	1	Corollary on ED and UFD	T1: chap-IX Pg.No:404-406
17	1	Example of Guassin integer	T1: chap-IX Pg.No:407-411
18	1	Tutorial	
19	1	Recapitulation and Discussion of possible questions	
	Total No of Hours Planned For Unit 2=19		
		UNIT-III	
1	1	Introduction to Dual Spaces	R1: chap-16 Pg.No:987-990
2	1	Definition and concepts of Dual basis	R1: chap-16 Pg.No:990-

			994
3	1	dual basis, double dual	R1: chap-16 Pg.No:112-994
4	1	Tutorial	
5	1	Transpose of a linear transformation and its matrix in the dual basis	R3: chap-16 Pg.No:1010-1015
6	1	Transpose of a linear transformation and its matrix in the dual basis-Problems	R3: chap-16 Pg.No:1026-1030
8	1	Tutorial	
9	1	Annihilator-Definition and basic concepts	R3: chap-16 Pg.No:1030-1032
10	1	Theorems on Annihilator	R3: chap-16 Pg.No:1032-1035
11	1	Eigen spaces of a linear operator	R3: chap-16 Pg.No:1041-1048
12	1	Tutorial	
13	1	Diagonalizability	R3: chap-16 Pg.No:1041-1048
14	1	invariant subspaces and Cayley-Hamilton theorem	R3: chap-16 Pg.No:1048-1055
15	1	Tutorial	
16	1	Problems on Cayley-Hamilton theorem	R3: chap-16 Pg.No:1050-1055
17	1	Minimal polynomial for a linear operator	R3: chap-16 Pg.No:1050-1055
18	1	Tutorial	
19	1	Recapitulation and Discussion of possible questions	
	Total No of Hours Planned For Unit 3 =19		
UNIT-IV			
1	1	Inner product spaces	R3: chap-17 Pg.No:1081-1083
2	1	Norm of the inner product spaces-Definitions	R3: chap-17 Pg.No:1070-1080
3	1	Examples For Inner Product Spaces And Norms	R3: chap-17 Pg.No:1070-1080
4	1	Tutorial	
5	1	Gram-Schmidt orthogonalisation process-theorem	R1:Chap-13 Pg.No:1034-1035
6	1	Continuation of the theorem on Gram-Schmidt orthogonalisation process	R1:Chap- Pg.No:1035-1037
7	1	Problems on Gram-Schmidt orthogonalisation process	R3: chap-17 Pg.No:1083-1090
8	1	Tutorial	
9	1	Orthogonal Complements	R3: chap-17 Pg.No:1083-

			1090
10	1	Bessel's inequality	R3: chap-17 Pg.No:1083-1090
11	1	The adjoint of a linear operator	R2: chap-10 Pg.No:1061-1090
12	1	Tutorial	
13	1	Theorems on the adjoint of a linear operator	
14	1	Continuation of theorems on adjoint of a linear operator	
15	1	Tutorial	
16	1	Recapitulation and Discussion of possible questions	
	Total No of Hours Planned For Unit 4=16		
UNIT-V			
1	1	Least Squares Approximation	R3: chap-17 Pg.No:1091-1097
2	1	Theorem on Least Squares Approximation	R3: chap-17 Pg.No:1117-1120
3	1	Problems on Least Squares Approximation	R3: chap-17 Pg.No:1120-1128
4	1	Tutorial	
5	1	Systems of linear equations	R1: chap- Pg.No:
6	1	Minimal solutions to systems of linear equations	R1: chap- Pg.No:
7	1	Tutorial	
8	1	Normaloperator-Problems	R3: chap-17 Pg.No:1128-1129
9	1	self-adjoint operators-Problems	R3: chap-17 Pg.No:1129-1133
10	1	Tutorial	
11	1	Orthogonal projections	R3: chap-17 Pg.No:1135-1136
12	1	Problems on Orthogonal projections	R3: chap-17 Pg.No:1136-1138
13	1	Spectral theorem	R3: chap-17 Pg.No:1136-1138
14	1	Tutorial	
15	1	Recapitulation and Discussion of possible questions	
16	1	Discussion on Previous ESE Question Papers	
17	1	Discussion on Previous ESE Question Papers	
18	1	Discussion on Previous ESE Question Papers	
	Total No of Hours Planned For Unit 5=18		

TEXT BOOKS

1. Fraleigh.J.B., (2004). A First Course in Abstract Algebra , Seventh Edition , Pearson EducationLtd, Singapore.

REFERENCES

1. Stephen H. Friedberg., Arnold J. Insel., Lawrence E. Spence, (2004) . Linear Algebra, Fourth Edition., Prentice- Hall of India Pvt. Ltd., New Delhi.
2. S. Lang, (2005). Introduction to Linear Algebra, Second Edition, Springer.2. Herstein.I.N.,(2010). Topics in Algebra ,Second Edition, Willey and sons Pvt Ltd, Singapore.
3. Joseph A. Gallian., (2001). Contemporary Abstract Algebra, Fourth Edition.,Narosa Publishing House, New Delhi.
4. Artin.M., (2008).Algebra, Prentice - Hall of India, New Delhi.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

SYLLABUS

Semester – IV

		L	T	P	C
16MMU403	**RING THEORY AND LINEAR ALGEBRA II**	**6**	**2**	**0**	**6**

Scope: On successful completion of course the learners gain about the behavior of polynomials and operators.

Objectives: To enable the students to learn and gain knowledge about polynomial rings over commutative rings, dual spaces, dual basis, double dual, minimal solutions to systems of linear equations, normal and self-adjoint operators.

UNIT I

Polynomial rings over commutative rings, division algorithm and consequences, principal ideal domains, factorization of polynomials, reducibility tests, irreducibility tests, Eisenstein criterion, unique factorization in Z[x]**.**

UNIT II

Divisibility in integral domains, irreducibles, primes, unique factorization domains, Euclidean domains.

UNIT III

Dual spaces, dual basis, double dual, transpose of a linear transformation and its matrix in the dual basis, annihilators, Eigen spaces of a linear operator, diagonalizability, invariant subspaces and Cayley-Hamilton theorem, the minimal polynomial for a linear operator.

UNIT IV

Inner product spaces and norms, Gram-Schmidt orthogonalisation process, orthogonal complements, Bessel's inequality, the adjoint of a linear operator.

UNIT V

Least Squares Approximation, minimal solutions to systems of linear equations, Normal and self-adjoint operators, Orthogonal projections and Spectral theorem.

SUGGESTED READINGS

TEXT BOOKS

1. Fraleigh.J.B., (2004). A First Course in Abstract Algebra , Seventh Edition , Pearson EducationLtd, Singapore.

REFERENCES

1. Stephen H. Friedberg., Arnold J. Insel., Lawrence E. Spence, (2004) . Linear Algebra, Fourth Edition., Prentice- Hall of India Pvt. Ltd., New Delhi.

2. S. Lang, (2005). Introduction to Linear Algebra, Second Edition, Springer.2. Herstein.I.N.,(2010). Topics in Algebra ,Second Edition, Willey and sons Pvt Ltd, Singapore.
3. Joseph A. Gallian., (2001). Contemporary Abstract Algebra, Fourth Edition.,Narosa Publishing House, New Delhi.
4. Artin.M., (2008).Algebra, Prentice - Hall of India, New Delhi.

UNIT-V

SYLLABUS

Least Squares Approximation, minimal solutions to systems of linear equations, Normal and self-adjoint operators, Orthogonal projections and Spectral theorem.

KAHE

Definition. *If* A *is an* n × n *matrix over the field* F, *a* **characteristic value of** A **in** F *is a scalar* c *in* F *such that the matrix* (A − cI) *is singular* (*not invertible*).

Since c is a characteristic value of A if and only if $\det(A - cI) = 0$, or equivalently if and only if $\det(cI - A) = 0$, we form the matrix $(xI - A)$ with polynomial entries, and consider the polynomial $f = \det(xI - A)$. Clearly the characteristic values of A in F are just the scalars c in F such that $f(c) = 0$. For this reason f is called the **characteristic polynomial** of A. It is important to note that f is a monic polynomial which has degree exactly n. This is easily seen from the formula for the determinant of a matrix in terms of its entries.

Lemma. *Similar matrices have the same characteristic polynomial.*

Proof. If $B = P^{-1}AP$, then

$$\begin{aligned}\det(xI - B) &= \det(xI - P^{-1}AP) \\ &= \det(P^{-1}(xI - A)P) \\ &= \det P^{-1} \cdot \det(xI - A) \cdot \det P \\ &= \det(xI - A). \quad \blacksquare\end{aligned}$$

This lemma enables us to define sensibly the characteristic polynomial of the operator T as the characteristic polynomial of any $n \times n$ matrix which represents T in some ordered basis for V. Just as for matrices, the characteristic values of T will be the roots of the characteristic polynomial for T. In particular, this shows us that T cannot have more than n distinct characteristic values. It is important to point out that T may not have any characteristic values.

EXAMPLE Let T be the linear operator on R^2 which is represented

in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

The characteristic polynomial for T (or for A) is

$$\det (xI - A) = \begin{vmatrix} x & 1 \\ -1 & x \end{vmatrix} = x^2 + 1.$$

Since this polynomial has no real roots, T has no characteristic values. If U is the linear operator on C^2 which is represented by A in the standard ordered basis, then U has two characteristic values, i and $-i$. Here we see a subtle point. In discussing the characteristic values of a matrix A, we must be careful to stipulate the field involved. The matrix A above has no characteristic values in R, but has the two characteristic values i and $-i$ in C.

EXAMPLE Let A be the (real) 3×3 matrix

$$\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}.$$

Then the characteristic polynomial for A is

$$\begin{vmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{vmatrix} = x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

Thus the characteristic values of A are 1 and 2.

Suppose that T is the linear operator on R^3 which is represented by A in the standard basis. Let us find the characteristic vectors of T associated with the characteristic values, 1 and 2. Now

$$A - I = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix}.$$

It is obvious at a glance that $A - I$ has rank equal to 2 (and hence $T - I$ has nullity equal to 1). So the space of characteristic vectors associated with the characteristic value 1 is one-dimensional. The vector $\alpha_1 = (1, 0, 2)$ spans the null space of $T - I$. Thus $T\alpha = \alpha$ if and only if α is a scalar multiple of α_1. Now consider

$$A - 2I = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix}.$$

Evidently $A - 2I$ also has rank 2, so that the space of characteristic vectors associated with the characteristic value 2 has dimension 1. Evidently $T\alpha = 2\alpha$ if and only if α is a scalar multiple of $\alpha_2 = (1, 1, 2)$.

Definition. *Let* T *be a linear operator on the finite-dimensional space* V. *We say that* T *is* **diagonalizable** *if there is a basis for* V *each vector of which is a characteristic vector of* T.

The reason for the name should be apparent; for, if there is an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ for V in which each α_i is a characteristic vector of T, then the matrix of T in the ordered basis $\mathcal{B}$ is diagonal. If $T\alpha_i = c_i\alpha_i$, then

$$[T]_{\mathcal{B}} = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}.$$

We certainly do not require that the scalars $c_1, \ldots, c_n$ be distinct; indeed, they may all be the same scalar (when T is a scalar multiple of the identity operator).

One could also define T to be diagonalizable when the characteristic vectors of T span V. This is only superficially different from our definition, since we can select a basis out of any spanning set of vectors.

Lemma. *Let* T *be a linear operator on the finite-dimensional space* V. *Let* $c_1, \ldots, c_k$ *be the distinct characteristic values of* T *and let* W_i *be the space of characteristic vectors associated with the characteristic value* c_i. *If* $W = W_1 + \cdots + W_k$, *then*

$$\dim W = \dim W_1 + \cdots + \dim W_k.$$

In fact, if $\mathcal{B}_i$ *is an ordered basis for* W_i, *then* $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ *is an ordered basis for* W.

Proof. The space $W = W_1 + \cdots + W_k$ is the subspace spanned by all of the characteristic vectors of T. Usually when one forms the sum W of subspaces W_i, one expects that $\dim W < \dim W_1 + \cdots + \dim W_k$ because of linear relations which may exist between vectors in the various spaces. This lemma states that the characteristic spaces associated with different characteristic values are independent of one another.

Suppose that (for each i) we have a vector β_i in W_i, and assume that $\beta_1 + \cdots + \beta_k = 0$. We shall show that $\beta_i = 0$ for each i. Let f be any polynomial. Since $T\beta_i = c_i\beta_i$, the preceding lemma tells us that

$$\begin{aligned} 0 = f(T)0 &= f(T)\beta_1 + \cdots + f(T)\beta_k \\ &= f(c_1)\beta_1 + \cdots + f(c_k)\beta_k. \end{aligned}$$

Choose polynomials $f_1, \ldots, f_k$ such that

$$f_i(c_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Then

$$\begin{aligned} 0 = f_i(T)0 &= \sum_j \delta_{ij}\beta_j \\ &= \beta_i. \end{aligned}$$

Now, let $\mathcal{B}_i$ be an ordered basis for W_i, and let $\mathcal{B}$ be the sequence $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$. Then $\mathcal{B}$ spans the subspace $W = W_1 + \cdots + W_k$. Also, $\mathcal{B}$ is a linearly independent sequence of vectors, for the following reason. Any linear relation between the vectors in $\mathcal{B}$ will have the form $\beta_1 + \cdots + \beta_k = 0$, where β_i is some linear combination of the vectors in $\mathcal{B}_i$. From what we just did, we know that $\beta_i = 0$ for each i. Since each $\mathcal{B}_i$ is linearly independent, we see that we have only the trivial linear relation between the vectors in $\mathcal{B}$. ∎

Theorem *Let* T *be a linear operator on a finite-dimensional space* V.

Let $c_1, \ldots, c_k$ *be the distinct characteristic values of* T *and let* W_i *be the null space of* $(T - c_iI)$. *The following are equivalent.*

(i) T *is diagonalizable.*

(ii) *The characteristic polynomial for* T *is*

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$$

and dim $W_i = d_i$, $i = 1, \ldots, k$.

(iii) *dim* $W_1 + \cdots + dim\ W_k = dim\ V$.

Proof. We have observed that (i) implies (ii). If the characteristic polynomial f is the product of linear factors, as in (ii), then $d_1 + \cdots + d_k = \dim V$. For, the sum of the d_i's is the degree of the characteristic polynomial, and that degree is dim V. Therefore (ii) implies (iii). Suppose (iii) holds. By the lemma, we must have $V = W_1 + \cdots + W_k$, i.e., the characteristic vectors of T span V. ∎

The matrix analogue of Theorem 2 may be formulated as follows. Let A be an $n \times n$ matrix with entries in a field F, and let $c_1, \ldots, c_k$ be the distinct characteristic values of A in F. For each i, let W_i be the space of column matrices X (with entries in F) such that

$$(A - c_iI)X = 0,$$

and let $\mathcal{B}_i$ be an ordered basis for W_i. The bases $\mathcal{B}_1, \ldots, \mathcal{B}_k$ collectively string together to form the sequence of columns of a matrix P:

$$P = [P_1, P_2, \ldots] = (\mathcal{B}_1, \ldots, \mathcal{B}_k).$$

The matrix A is similar over F to a diagonal matrix if and only if P is a square matrix. When P is square, P is invertible and $P^{-1}AP$ is diagonal.

EXAMPLE 3. Let T be the linear operator on R^3 which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

Let us indicate how one might compute the characteristic polynomial, using various row and column operations:

$$\begin{vmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{vmatrix} = \begin{vmatrix} x-5 & 0 & 6 \\ 1 & x-2 & -2 \\ -3 & 2-x & x+4 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-5 & 0 & 6 \\ 1 & 1 & -2 \\ -3 & -1 & x+4 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-5 & 0 & 6 \\ 1 & 1 & -2 \\ -2 & 0 & x+2 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-5 & 6 \\ -2 & x+2 \end{vmatrix}$$

$$= (x-2)(x^2-3x+2)$$
$$= (x-2)^2(x-1).$$

What are the dimensions of the spaces of characteristic vectors associated with the two characteristic values? We have

$$A - I = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}$$

$$A - 2I = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix}.$$

We know that $A - I$ is singular and obviously rank $(A - I) \geq 2$. Therefore, rank $(A - I) = 2$. It is evident that rank $(A - 2I) = 1$.

Let W_1, W_2 be the spaces of characteristic vectors associated with the characteristic values 1, 2. We know that dim $W_1 = 1$ and dim $W_2 = 2$. By Theorem 2, T is diagonalizable. It is easy to exhibit a basis for R^3 in which T is represented by a diagonal matrix. The null space of $(T - I)$ is spanned

by the vector $\alpha_1 = (3, -1, 3)$ and so $\{\alpha_1\}$ is a basis for W_1. The null space of $T - 2I$ (i.e., the space W_2) consists of the vectors (x_1, x_2, x_3) with $x_1 = 2x_2 + 2x_3$. Thus, one example of a basis for W_2 is

$$\alpha_2 = (2, 1, 0)$$
$$\alpha_3 = (2, 0, 1).$$

If $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$, then $[T]_{\mathcal{B}}$ is the diagonal matrix

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The fact that T is diagonalizable means that the original matrix A is similar (over R) to the diagonal matrix D. The matrix P which enables us to change coordinates from the basis $\mathcal{B}$ to the standard basis is (of course) the matrix which has the transposes of α_1, α_2, α_3 as its column vectors:

$$P = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

Furthermore, $AP = PD$, so that

$$P^{-1}AP = D.$$

Definition. *Let* T *be a linear operator on a finite-dimensional vector space* V *over the field* F. *The* **minimal polynomial** *for* T *is the (unique) monic generator of the ideal of polynomials over* F *which annihilate* T.

The name 'minimal polynomial' stems from the fact that the generator of a polynomial ideal is characterized by being the monic polynomial of minimum degree in the ideal. That means that the minimal polynomial p for the linear operator T is uniquely determined by these three properties:

(1) p is a monic polynomial over the scalar field F.
(2) $p(T) = 0$.
(3) No polynomial over F which annihilates T has smaller degree than p has.

If A is an $n \times n$ matrix over F, we define the **minimal polynomial** for A in an analogous way, as the unique monic generator of the ideal of all polynomials over F which annihilate A. If the operator T is represented in

some ordered basis by the matrix A, then T and A have the same minimal polynomial. That is because $f(T)$ is represented in the basis by the matrix $f(A)$, so that $f(T) = 0$ if and only if $f(A) = 0$.

From the last remark about operators and matrices it follows that similar matrices have the same minimal polynomial. That fact is also clear from the definitions because

$$f(P^{-1}AP) = P^{-1}f(A)P$$

for every polynomial f.

Theorem *Let* T *be a linear operator on an* n*-dimensional vector space* V [*or, let* A *be an* n × n *matrix*]. *The characteristic and minimal polynomials for* T [*for* A] *have the same roots, except for multiplicities.*

Proof. Let p be the minimal polynomial for T. Let c be a scalar. What we want to show is that $p(c) = 0$ if and only if c is a characteristic value of T.

First, suppose $p(c) = 0$. Then

$$p = (x - c)q$$

where q is a polynomial. Since deg $q <$ deg p, the definition of the minimal polynomial p tells us that $q(T) \neq 0$. Choose a vector β such that $q(T)\beta \neq 0$. Let $\alpha = q(T)\beta$. Then

$$\begin{aligned} 0 &= p(T)\beta \\ &= (T - cI)q(T)\beta \\ &= (T - cI)\alpha \end{aligned}$$

and thus, c is a characteristic value of T.

Now, suppose that c is a characteristic value of T, say, $T\alpha = c\alpha$ with $\alpha \neq 0$. As we noted in a previous lemma,

$$p(T)\alpha = p(c)\alpha.$$

Since $p(T) = 0$ and $\alpha \neq 0$, we have $p(c) = 0$. ∎

Let T be a diagonalizable linear operator and let $c_1, \ldots, c_k$ be the distinct characteristic values of T. Then it is easy to see that the minimal polynomial for T is the polynomial

$$p = (x - c_1) \cdots (x - c_k).$$

If α is a characteristic vector, then one of the operators $T - c_1 I, \ldots, T - c_k I$ sends α into 0. Therefore

$$(T - c_1 I) \cdots (T - c_k I)\alpha = 0$$

for every characteristic vector α. There is a basis for the underlying space which consists of characteristic vectors of T; hence

$$p(T) = (T - c_1 I) \cdots (T - c_k I) = 0.$$

What we have concluded is this. If T is a diagonalizable linear operator, then the minimal polynomial for T is a product of distinct linear factors. As we shall soon see, that property characterizes diagonalizable operators.

EXAMPLE Let's try to find the minimal polynomials for the operators in Examples 1, 2, and 3. We shall discuss them in reverse order. The operator in Example 3 was found to be diagonalizable with characteristic polynomial

$$f = (x - 1)(x - 2)^2.$$

From the preceding paragraph, we know that the minimal polynomial for T is

$$p = (x - 1)(x - 2).$$

The reader might find it reassuring to verify directly that

$$(A - I)(A - 2I) = 0.$$

In Example 2, the operator T also had the characteristic polynomial $f = (x - 1)(x - 2)^2$. But, this T is not diagonalizable, so we don't know that the minimal polynomial is $(x - 1)(x - 2)$. What do we know about the minimal polynomial in this case? From Theorem 3 we know that its roots are 1 and 2, with some multiplicities allowed. Thus we search for p among polynomials of the form $(x - 1)^k(x - 2)^l$, $k \geq 1, l \geq 1$. Try $(x - 1)(x - 2)$:

$$(A - I)(A - 2I) = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{bmatrix}.$$

Thus, the minimal polynomial has degree at least 3. So, next we should try either $(x-1)^2(x-2)$ or $(x-1)(x-2)^2$. The second, being the characteristic polynomial, would seem a less random choice. One can readily compute that $(A-I)(A-2I)^2=0$. Thus the minimal polynomial for T is its characteristic polynomial.

In Example 1 we discussed the linear operator T on R^2 which is represented in the standard basis by the matrix

$$A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}.$$

The characteristic polynomial is x^2+1, which has no real roots. To determine the minimal polynomial, forget about T and concentrate on A. As a complex 2×2 matrix, A has the characteristic values i and $-i$. Both roots must appear in the minimal polynomial. Thus the minimal polynomial is divisible by x^2+1. It is trivial to verify that $A^2+I=0$. So the minimal polynomial is x^2+1.

Theorem 4 (Cayley-Hamilton). *Let* T *be a linear operator on a finite dimensional vector space* V. *If* f *is the characteristic polynomial for* T, *then* $f(T)=0$; *in other words, the minimal polynomial divides the characteristic polynomial for* T.

Proof. Later on we shall give two proofs of this result independent of the one to be given here. The present proof, although short, may be difficult to understand. Aside from brevity, it has the virtue of providing

an illuminating and far from trivial application of the general theory of determinants developed in Chapter 5.

Let K be the commutative ring with identity consisting of all polynomials in T. Of course, K is actually a commutative algebra with identity over the scalar field. Choose an ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ for V, and let A be the matrix which represents T in the given basis. Then

$$T\alpha_i = \sum_{j=1}^{n} A_{ji}\alpha_j, \qquad 1 \leq i \leq n.$$

These equations may be written in the equivalent form

$$\sum_{j=1}^{n} (\delta_{ij}T - A_{ji}I)\alpha_j = 0, \qquad 1 \leq i \leq n.$$

Let B denote the element of $K^{n\times n}$ with entries

$$B_{ij} = \delta_{ij}T - A_{ji}I.$$

When $n = 2$

$$B = \begin{bmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{bmatrix}$$

and

$$\begin{aligned} \det B &= (T - A_{11}I)(T - A_{22}I) - A_{12}A_{21}I \\ &= T^2 - (A_{11} + A_{22})T + (A_{11}A_{22} - A_{12}A_{21})I \\ &= f(T) \end{aligned}$$

where f is the characteristic polynomial:

$$f = x^2 - (\text{trace } A)x + \det A.$$

For the case $n > 2$, it is also clear that

$$\det B = f(T)$$

since f is the determinant of the matrix $xI - A$ whose entries are the polynomials

$$(xI - A)_{ij} = \delta_{ij}x - A_{ji}.$$

We wish to show that $f(T) = 0$. In order that $f(T)$ be the zero operator, it is necessary and sufficient that $(\det B)\alpha_k = 0$ for $k = 1, \ldots, n$. By the definition of B, the vectors $\alpha_1, \ldots, \alpha_n$ satisfy the equations

$$\sum_{j=1}^{n} B_{ij}\alpha_j = 0, \qquad 1 \leq i \leq n. \tag{6-6}$$

When $n = 2$, it is suggestive to write (6-6) in the form

$$\begin{bmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In this case, the classical adjoint, adj B is the matrix

$$\tilde{B} = \begin{bmatrix} T - A_{22}I & A_{21}I \\ A_{12}I & T - A_{11}I \end{bmatrix}$$

and

$$\tilde{B}B = \begin{bmatrix} \det B & 0 \\ 0 & \det B \end{bmatrix}.$$

Hence, we have

$$\begin{aligned} (\det B) \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} &= (\tilde{B}B) \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \\ &= \tilde{B} \left(B \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \right) \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

In the general case, let $\tilde{B} = \text{adj } B$. Then by (6-6)

$$\sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j = 0$$

for each pair k, i, and summing on i, we have

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \tilde{B}_{ki}B_{ij} \right) \alpha_j. \end{aligned}$$

Now $\tilde{B}B = (\det B)I$, so that

$$\sum_{i=1}^{n} \tilde{B}_{ki}B_{ij} = \delta_{kj} \det B.$$

Therefore

$$\begin{aligned} 0 &= \sum_{j=1}^{n} \delta_{kj}(\det B)\alpha_j \\ &= (\det B)\alpha_k, \qquad 1 \leq k \leq n. \end{aligned}$$ ∎

The Cayley-Hamilton theorem is useful to us at this point primarily because it narrows down the search for the minimal polynomials of various operators. If we know the matrix A which represents T in some ordered basis, then we can compute the characteristic polynomial f. We know that the minimal polynomial p divides f and that the two polynomials have the same roots. There is no method for computing precisely the roots of a polynomial (unless its degree is small); however, if f factors

$$(6\text{-}7) \qquad f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}, \qquad c_1, \ldots, c_k \text{ distinct}, d_i \geq 1$$

then

$$(6\text{-}8) \qquad p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}, \qquad 1 \leq r_j \leq d_j.$$

That is all we can say in general. If f is the polynomial (6-7) and has degree n, then for every polynomial p as in (6-8) we can find an $n \times n$ matrix which has f as its characteristic polynomial and p as its minimal polynomial. We shall not prove this now. But, we want to emphasize the fact that the knowledge that the characteristic polynomial has the form (6-7) tells us that the minimal polynomial has the form (6-8), and it tells us nothing else about p.

EXAMPLE 5. Let A be the 4×4 (rational) matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

The powers of A are easy to compute:

$$A^2 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{bmatrix}.$$

Thus $A^3 = 4A$, i.e., if $p = x^3 - 4x = x(x+2)(x-2)$, then $p(A) = 0$. The minimal polynomial for A must divide p. That minimal polynomial is obviously not of degree 1, since that would mean that A was a scalar multiple of the identity. Hence, the candidates for the minimal polynomial are: p, $x(x+2)$, $x(x-2)$, $x^2 - 4$. The three quadratic polynomials can be eliminated because it is obvious at a glance that $A^2 \neq -2A$, $A^2 \neq 2A$, $A^2 \neq 4I$. Therefore p is the minimal polynomial for A. In particular 0, 2, and -2 are the characteristic values of A. One of the factors x, $x-2$, $x+2$ must be repeated twice in the characteristic polynomial. Evidently, rank $(A) = 2$. Consequently there is a two-dimensional space of characteristic vectors associated with the characteristic value 0. From Theorem 2, it should now be clear that the characteristic polynomial is $x^2(x^2-4)$ and that A is similar over the field of rational numbers to the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

Invariant Subspaces

Definition. *Let* V *be a vector space and* T *a linear operator on* V. *If* W *is a subspace of* V, *we say that* W *is* **invariant under** T *if for each vector* α *in* W *the vector* Tα *is in* W, *i.e., if* T(W) *is contained in* W.

EXAMPLE 6. If T is any linear operator on V, then V is invariant under T, as is the zero subspace. The range of T and the null space of T are also invariant under T.

EXAMPLE 7. Let F be a field and let D be the differentiation operator on the space $F[x]$ of polynomials over F. Let n be a positive integer and let W be the subspace of polynomials of degree not greater than n. Then W is invariant under D. This is just another way of saying that D is 'degree decreasing.'

EXAMPLE Let T be any linear operator on a finite-dimensional space V. Let W be the subspace spanned by *all* of the characteristic vectors

space V. Let W be the subspace spanned by *all* of the characteristic vectors of T. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T. For each i, let W_i be the space of characteristic vectors associated with the characteristic value c_i, and let $\mathcal{B}_i$ be an ordered basis for W_i. The lemma before Theorem 2 tells us that $\mathcal{B}' = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is an ordered basis for W. In particular,

$$\dim W = \dim W_1 + \cdots + \dim W_k.$$

Let $\mathcal{B}' = \{\alpha_1, \ldots, \alpha_r\}$ so that the first few α's form the basis $\mathcal{B}_1$, the next few $\mathcal{B}_2$, and so on. Then

$$T\alpha_i = t_i\alpha_i, \qquad i = 1, \ldots, r$$

where $(t_1, \ldots, t_r) = (c_1, c_1, \ldots, c_1, \ldots, c_k, c_k, \ldots, c_k)$ with c_i repeated $\dim W_i$ times.

Now W is invariant under T, since for each α in W we have

$$\alpha = x_1\alpha_1 + \cdots + x_r\alpha_r$$
$$T\alpha = t_1x_1\alpha_1 + \cdots + t_rx_r\alpha_r.$$

Choose any other vectors $\alpha_{r+1}, \ldots, \alpha_n$ in V such that $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. The matrix of T relative to $\mathcal{B}$ has the block form (6-10), and the matrix of the restriction operator T_W relative to the basis $\mathcal{B}'$ is

$$B = \begin{bmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & t_r \end{bmatrix}.$$

The characteristic polynomial of B (i.e., of T_W) is

$$g = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$$

where $e_i = \dim W_i$. Furthermore, g divides f, the characteristic polynomial for T. Therefore, the multiplicity of c_i as a root of f is at least $\dim W_i$.

All of this should make Theorem 2 transparent. It merely says that T is diagonalizable if and only if $r = n$, if and only if $e_1 + \cdots + e_k = n$. It does not help us too much with the non-diagonalizable case, since we don't know the matrices C and D of (6-10).

Definition. *Let* W *be an invariant subspace for* T *and let* α *be a vector in* V. *The* **T-conductor of α into** W *is the set* $S_T(\alpha; W)$, *which consists of all polynomials* g *(over the scalar field) such that* $g(T)\alpha$ *is in* W.

Since the operator T will be fixed throughout most discussions, we shall usually drop the subscript T and write $S(\alpha; W)$. The authors usually call that collection of polynomials the 'stuffer' (*das einstopfende Ideal*). 'Conductor' is the more standard term, preferred by those who envision a less aggressive operator $g(T)$, gently leading the vector α into W. In the special case $W = \{0\}$ the conductor is called the **T-annihilator of α.**

Lemma. *If* W *is an invariant subspace for* T, *then* W *is invariant under every polynomial in* T. *Thus, for each* α *in* V, *the conductor* $S(\alpha; W)$ *is an ideal in the polynomial algebra* F[x].

Proof. If β is in W, then $T\beta$ is in W. Consequently, $T(T\beta) = T^2\beta$ is in W. By induction, $T^k\beta$ is in W for each k. Take linear combinations to see that $f(T)\beta$ is in W for every polynomial f.

The definition of $S(\alpha; W)$ makes sense if W is any subset of V. If W is a subspace, then $S(\alpha; W)$ is a subspace of $F[x]$, because

$$(cf + g)(T) = cf(T) + g(T).$$

If W is also invariant under T, let g be a polynomial in $S(\alpha; W)$, i.e., let $g(T)\alpha$ be in W. If f is any polynomial, then $f(T)[g(T)\alpha]$ will be in W. Since

$$(fg)(T) = f(T)g(T)$$

fg is in $S(\alpha; W)$. Thus the conductor absorbs multiplication by any polynomial. ∎

The unique monic generator of the ideal $S(\alpha; W)$ is also called the **T-conductor of α into W** (the **T-annihilator** in case $W = \{0\}$). The T-conductor of α into W is the monic polynomial g of least degree such that $g(T)\alpha$ is in W. A polynomial f is in $S(\alpha; W)$ if and only if g divides f. Note that the conductor $S(\alpha; W)$ always contains the minimal polynomial for T; hence, *every* T*-conductor divides the minimal polynomial for* T.

As the first illustration of how to use the conductor $S(\alpha; W)$, we shall characterize triangulable operators. The linear operator T is called **triangulable** if there is an ordered basis in which T is represented by a triangular matrix.

Lemma. *Let* V *be a finite-dimensional vector space over the field* F. *Let* T *be a linear operator on* V *such that the minimal polynomial for* T *is a product of linear factors*

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}, \qquad c_i \text{ in } F.$$

Let W *be a proper* ($W \neq V$) *subspace of* V *which is invariant under* T. *There exists a vector* α *in* V *such that*

(a) α *is not in* W;
(b) $(T - cI)\alpha$ *is in* W, *for some characteristic value* c *of the operator* T.

Proof. What (a) and (b) say is that the T-conductor of α into W is a linear polynomial. Let β be any vector in V which is not in W. Let g be the T-conductor of β into W. Then g divides p, the minimal polynomial for T. Since β is not in W, the polynomial g is not constant. Therefore,

$$g = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$$

where at least one of the integers e_i is positive. Choose j so that $e_j > 0$. Then $(x - c_j)$ divides g:

$$g = (x - c_j)h.$$

By the definition of g, the vector $\alpha = h(T)\beta$ cannot be in W. But

$$\begin{aligned}(T - c_jI)\alpha &= (T - c_jI)h(T)\beta \\ &= g(T)\beta\end{aligned}$$

is in W. ∎

Theorem *Let* V *be a finite-dimensional vector space over the field* F

and let T *be a linear operator on* V. *Then* T *is triangulable if and only if the minimal polynomial for* T *is a product of linear polynomials over* F.

Proof. Suppose that the minimal polynomial factors

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}.$$

By repeated application of the lemma above, we shall arrive at an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ in which the matrix representing T is upper-triangular:

$$[T]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}. \tag{6-11}$$

Now (6-11) merely says that

$$T\alpha_j = a_{1j}\alpha_1 + \cdots + a_{jj}\alpha_j, \qquad 1 \leq j \leq n \tag{6-12}$$

that is, $T\alpha_j$ is in the subspace spanned by $\alpha_1, \ldots, \alpha_j$. To find $\alpha_1, \ldots, \alpha_n$, we start by applying the lemma to the subspace $W = \{0\}$, to obtain the vector α_1. Then apply the lemma to W_1, the space spanned by α_1, and we obtain α_2. Next apply the lemma to W_2, the space spanned by α_1 and α_2. Continue in that way. One point deserves comment. After $\alpha_1, \ldots, \alpha_i$ have been found, it is the triangular-type relations (6-12) for $j = 1, \ldots, i$ which ensure that the subspace spanned by $\alpha_1, \ldots, \alpha_i$ is invariant under T.

If T is triangulable, it is evident that the characteristic polynomial for T has the form

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}, \qquad c_i \text{ in } F.$$

Just look at the triangular matrix (6-11). The diagonal entries $a_{11}, \ldots, a_{1n}$ are the characteristic values, with c_i repeated d_i times. But, if f can be so factored, so can the minimal polynomial p, because it divides f. ∎

Corollary. *Let* F *be an algebraically closed field, e.g., the complex number field. Every* n × n *matrix over* F *is similar over* F *to a triangular matrix.*

Theorem *Let* V *be a finite-dimensional vector space over the field* F

and let T *be a linear operator on* V. *Then* T *is diagonalizable if and only if the minimal polynomial for* T *has the form*

$$p = (x - c_1) \cdots (x - c_k)$$

where $c_1, \ldots, c_k$ *are distinct elements of* F.

Proof. We have noted earlier that, if T is diagonalizable, its minimal polynomial is a product of distinct linear factors (see the discussion prior to Example 4). To prove the converse, let W be the subspace spanned by all of the characteristic vectors of T, and suppose $W \neq V$. By the lemma used in the proof of Theorem 5, there is a vector α not in W and a characteristic value c_j of T such that the vector

$$\beta = (T - c_j I)\alpha$$

lies in W. Since β is in W,

$$\beta = \beta_1 + \cdots + \beta_k$$

where $T\beta_i = c_i\beta_i$, $1 \leq i \leq k$, and therefore the vector

$$h(T)\beta = h(c_1)\beta_1 + \cdots + h(c_k)\beta_k$$

is in W, for every polynomial h.

Now $p = (x - c_j)q$, for some polynomial q. Also

$$q - q(c_j) = (x - c_j)h.$$

We have

$$q(T)\alpha - q(c_j)\alpha = h(T)(T - c_j I)\alpha = h(T)\beta$$

But $h(T)\beta$ is in W and, since

$$0 = p(T)\alpha = (T - c_j I)q(T)\alpha$$

the vector $q(T)\alpha$ is in W. Therefore, $q(c_j)\alpha$ is in W. Since α is not in W, we have $q(c_j) = 0$. That contradicts the fact that p has distinct roots. ∎

POSSIBLE QUESTIONS

PART-B (5 x 2 =10 Marks)

Answer all the questions

1. Define Homomorphism.
2. Define Isomorphism.
3. Define Normalizer.
4. Define centre of a group.
5. Write about Kernel of a group.

PART-C (5 x 6 =30 Marks)

Answer all the questions

1. Let G be a group of all non zero real number under multiplication and $\bar{G}$={1, -1}be group under multiplication. Prove that φ is a homomorphism.
2. If $\varphi$is a homomorphism of G into $\bar{G}$, then prove that
   (i) $\varphi(e) = \bar{e}$, the unit element in $\bar{G}$.
   (ii) $\varphi(x^{-1}) = \varphi(x)^{-1} \forall\, x \in G$.
3. Let φ be a homomorphism of G into $\bar{G}$ with kernel K, then prove that K is a normal subgroup of G.
4. If φ is a homomorphism of G onto $\bar{G}$ with kernel K, then show that the set of all inverse images $\bar{g} \in \bar{G}$ under φ in G is given by kx where x is any particular inverse image of $\bar{g}$ in G.
5. State and prove fundamental theorem of homomorphism.
6. State and prove first isomorphism theorem.
7. Let φ be the homomorphism of G onto $\bar{G}$with kernel K. For $\bar{H}$ is a subgroup of $\bar{G}$, let H be defined by H = $\{x \in G : \varphi(x) \in \bar{H}\}$. Then prove that H is a subgroup of G and H$\supset$ K. If $\bar{H}$ is normal in $\bar{G}$, then H is normal in G.
8. State and prove third isomorphism theorem.
9. State and prove second isomorphism theorem.
10. State and prove Cayley's theorem.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari (Po),
Coimbatore –641 021

Subject: RING THEORY AND LINEAR ALGEBRA-II **Subject Code: 16MMU403**

Class : II B.Sc Mathematics **Semester : IV**

UNIT -I

PART A (20x1=20 Marks)

(Question Nos. 1 to 20 Online Examinations)

Possible Questions

Question	Choice 1	Choice 2	Choice 3	Choice 4	Answer
A commutative ring with unity without zero divisors is called -------	integral domain	zero	identity	commutative ring	integral domain
A commutative ring with unity ---------- is called integral domain	divisors	with zero divisors	zero	identity	without zero divisors
A commutative division ring is ---------	ring	Field	integral domain	zero	Field
Another name of division ring is----------	Field	integral domain	skew Field	group	Field
Every ---------- is a field	integral domain	domain	domain	ring	finite integral domain
An element a of a ring R is said to be idempotent if -------	a=1	$a^2=1$	$a^2=a$	$a^2=0$	$a^2=a$
An element a of a ring R is said to be ---------- if $a^2=a$	idempotent	nilpotent	identity	none	idempotent
An element a of a ring R is said to be ---------- if $a^2=0$	idempotent	nilpotent	identity	none	nilpotent
A commutative ring is an ---------- if it has no zero divisors	Division ring	field	integral domain	Eucledian ring	integral domain
A ring is said to be ---------- if its nonzero elements form a group under multiplication	Division ring	field	integral domain	Eucledian ring	Division ring
A ring is said to be division ring if its nonzero elements form a ---------- under multiplication	Division ring	group	integral domain	Eucledian ring	group
A commutative ring is an integral domain if it has ----------	Division ring	field	no zero divisiors	zero divisiors	no zero divisiors
A finite integral domain is a ----------	Division ring	field	integral domain	Eucledian ring	field
A ---------- is a field-	Division ring	domain	integral domain	ring	finite integral domain
A homomorphism of R into R' is said to be an ---------- if it is a one-to one mapping	isomorphism	automorphism	homomorphism	monomorphism	isomorphism
A homomorphism of R into R' is said to be an isomrphism if it is a ----------mapping	one-one	onto	into	into & onto	one-one
A homomorphism of R into R' is said to be an isomrphism if and only if I(Φ)= -------	one	zero	two	three	zero
A ---------- ring is an integral domain if it has no zero divisors	Division ring	field	commutative ring	Eucledian ring	commutative ring
A---------- possesses a unit element	Division ring	field	integral domain	Eucledian ring	Eucledian ring
A non-empty set I is called ---------- if it is both left and right ideal K	one-sided ideal	two-sided ideal	field	integral domain	two-sided ideal
A non-empty set I is called two sided ideal if it is ----------	left ideal	right ideal	field	ideal	ideal
The polynomial is said to be --------- if the G.C.D is one	primitive	field	integral domain	Eucledian ring	primitive
The polynomial is said to be primitive if the G.C.D is ----------	two	one	zero	four	one
A polynomial is said to be integer monic if all its coefficients are ----------	integers	rational	real	complex	integers
A polynomial is said to be ---------- if all its coefficients are integers	integer monic	rational monic	real monic	complex monic	integer monic
J(i) is a ----------	integral domain	Euclidean ring	Field	skew field	Euclidean ring
.---------- is a Eucledian ring.	F(i)	J(i)	M(i)	A(i)	J(i)
If a∈R is an ---------- and a/bc, then a/b or a/c	zero divisor	primitive	irreducible	integers	irreducible
_____ is a commutative ring with unit element	(R, +,.)	(Z, *,.)	(R, *,.)	(R, +,*)	(R, +,.)
(R, +,.) is a with unit element	field	commutative ring	Eucledian ring	ring	commutative ring
If in a ring R there is an element 1 in R such that a.1=1.a=a then R is --------	element	commutative ring	zero	none	ring with unit element
If the multiplication of R such that a.b=b.a then R is --------	element	commutative ring	zero	none	commutative ring

UNIT-I

SYLLABUS

Polynomial rings over commutative rings, division algorithm and consequences, principal idealdomains, factorization of polynomials, reducibility tests, irreducibility tests, Eisenstein criterion, unique factorization in Z[x].

KAHE

RINGS OF POLYNOMIALS

Let R be a ring. A **polynomial $f(x)$ with coefficients in R** is an infinite formal sum

$$\sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n + \cdots,$$

where $a_i \in R$ and $a_i = 0$ for all but a finite number of values of i. The a_i are **coefficients of $f(x)$**. If for some $i \geq 0$ it is true that $a_i \neq 0$, the largest such value of i is the **degree of $f(x)$**. If all $a_i = 0$, then the degree of $f(x)$ is undefined.† ■

To simplify working with polynomials, let us agree that if $f(x) = a_0 + a_1 x + \cdots + a_n x^n + \cdots$ has $a_i = 0$ for $i > n$, then we may denote $f(x)$ by $a_0 + a_1 x + \cdots + a_n x^n$. Also, if R has unity $1 \neq 0$, we will write a term $1x^k$ in such a sum as x^k. For example, in $\mathbb{Z}[x]$, we will write the polynomial $2 + 1x$ as $2 + x$. Finally, we shall agree that we may omit altogether from the formal sum any term $0x^i$, or a_0 if $a_0 = 0$ but not all $a_i = 0$. Thus 0, 2, x, and $2 + x^2$ are polynomials with coefficients in $\mathbb{Z}$. An element of R is a **constant polynomial.**

Addition and multiplication of polynomials with coefficients in a ring R are defined in a way familiar to us. If

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n + \cdots$$

and

$$g(x) = b_0 + b_1 x + \cdots + b_n x^n + \cdots,$$

then for polynomial addition, we have

$$f(x) + g(x) = c_0 + c_1 x + \cdots + c_n x^n + \cdots \text{ where } c_n = a_n + b_n.$$

Theorem The set $R[x]$ of all polynomials in an indeterminate x with coefficients in a ring R is a ring under polynomial addition and multiplication. If R is commutative, then so is $R[x]$, and if R has unity $1 \neq 0$, then 1 is also unity for $R[x]$.

Proof That $\langle R[x], + \rangle$ is an abelian group is apparent. The associative law for multiplication and the distributive laws are straightforward, but slightly cumbersome, computations. We illustrate by proving the associative law.

Applying ring axioms to $a_i, b_j, c_k \in R$, we obtain

$$\begin{aligned}\left[\left(\sum_{i=0}^{\infty} a_i x^i\right)\left(\sum_{j=0}^{\infty} b_j x^j\right)\right]\left(\sum_{k=0}^{\infty} c_k x^k\right) &= \left[\sum_{n=0}^{\infty}\left(\sum_{i=0}^{n} a_i b_{n-i}\right)x^n\right]\left(\sum_{k=0}^{\infty} c_k x^k\right)\\ &= \sum_{s=0}^{\infty}\left[\sum_{n=0}^{s}\left(\sum_{i=0}^{n} a_i b_{n-i}\right)c_{s-n}\right]x^s\\ &= \sum_{s=0}^{\infty}\left(\sum_{i+j+k=s} a_i b_j c_k\right)x^s\\ &= \sum_{s=0}^{\infty}\left[\sum_{m=0}^{s} a_{s-m}\left(\sum_{j=0}^{m} b_j c_{m-j}\right)\right]x^s\\ &= \left(\sum_{i=0}^{\infty} a_i x^i\right)\left[\sum_{m=0}^{\infty}\left(\sum_{j=0}^{m} b_j c_{m-j}\right)x^m\right]\\ &= \left(\sum_{i=0}^{\infty} a_i x^i\right)\left[\left(\sum_{j=0}^{\infty} b_j x^j\right)\left(\sum_{k=0}^{\infty} c_k x^k\right)\right].\end{aligned}$$

The Evaluation Homomorphisms

Theorem **(The Evaluation Homomorphisms for Field Theory)** Let F be a subfield of a field E, let α be any element of E, and let x be an indeterminate. The map $\phi_\alpha : F[x] \to E$ defined by

$$\phi_\alpha(a_0 + a_1 x + \cdots + a_n x^n) = a_0 + a_1\alpha + \cdots + a_n\alpha^n$$

for $(a_0 + a_1 x + \cdots + a_n x^n) \in F[x]$ is a homomorphism of $F[x]$ into E. Also, $\phi_\alpha(x) = \alpha$, and ϕ_α maps F isomorphically by the identity map; that is, $\phi_\alpha(a) = a$ for $a \in F$. The homomorphism ϕ_α is **evaluation at** α.

Proof The subfield and mapping diagram in Fig. 22.5 may help us to visualize this situation. The dashed lines indicate an element of the set. The theorem is really an immediate

consequence of our definitions of addition and multiplication in $F[x]$. The map ϕ_α is well defined, that is, independent of our representation of $f(x) \in F[x]$ as a finite sum

$$a_0 + a_1x + \cdots + a_nx^n,$$

since such a finite sum representing $f(x)$ can be changed only by insertion or deletion of terms $0x^i$, which does not affect the value of $\phi_\alpha(f(x))$.

If $f(x) = a_0 + a_1x + \cdots + a_nx^n$, $g(x) = b_0 + b_1x + \cdots + b_mx^m$, and $h(x) = f(x) + g(x) = c_0 + c_1x + \cdots + c_rx^r$, then

$$\phi_\alpha(f(x) + g(x)) = \phi_\alpha(h(x)) = c_0 + c_1\alpha + \cdots + c_r\alpha^r,$$

while

$$\phi_\alpha(f(x)) + \phi_\alpha(g(x)) = \left(a_0 + a_1\alpha + \cdots + a_n\alpha^n\right) + \left(b_0 + b_1\alpha + \cdots + b_m\alpha^m\right).$$

Since by definition of polynomial addition we have $c_i = a_i + b_i$, we see that

$$\phi_\alpha(f(x) + g(x)) = \phi_\alpha(f(x)) + \phi_\alpha(g(x)).$$

Turning to multiplication, we see that if

$$f(x)g(x) = d_0 + d_1x + \cdots + d_sx^s,$$

then

$$\phi_\alpha(f(x)g(x)) = d_0 + d_1\alpha + \cdots + d_s\alpha^s,$$

while

$$[\phi_\alpha(f(x))][\phi_\alpha(g(x))] = \left(a_0 + a_1\alpha + \cdots + \alpha_n\alpha^n\right)\left(b_0 + b_1\alpha + \cdots + b_m\alpha^m\right).$$

Since by definition of polynomial multiplication $d_j = \sum_{i=0}^{j} a_ib_{j-i}$, we see that

$$\phi_\alpha(f(x)g(x)) = [\phi_\alpha(f(x))][\phi_\alpha(g(x))].$$

Thus ϕ_α is a homomorphism.

The very definition of ϕ_α applied to a constant polynomial $a \in F[x]$, where $a \in F$, gives $\phi_\alpha(a) = a$, so ϕ_α maps F isomorphically by the identity map. Again by definition of ϕ_α, we have $\phi_\alpha(x) = \phi_\alpha(1x) = 1\alpha = \alpha$. ◆

Example Let F be $\mathbb{Q}$ and E be $\mathbb{R}$ in Theorem 22.4, and consider the evaluation homomorphism $\phi_0 : \mathbb{Q}[x] \to \mathbb{R}$. Here

$$\phi_0(a_0 + a_1x + \cdots + a_nx^n) = a_0 + a_10 + \cdots + a_n0^n = a_0.$$

Thus every polynomial is mapped onto its constant term. ▲

Example Let F be $\mathbb{Q}$ and E be $\mathbb{R}$ in Theorem 22.4 and consider the evaluation homomorphism $\phi_2 : \mathbb{Q}[x] \to \mathbb{R}$. Here

$$\phi_2(a_0 + a_1x + \cdots + a_nx^n) = a_0 + a_12 + \cdots + a_n2^n.$$

Note that

$$\phi_2(x^2 + x - 6) = 2^2 + 2 - 6 = 0.$$

Thus $x^2 + x - 6$ is in the kernel N of ϕ_2. Of course,

$$x^2 + x - 6 = (x - 2)(x + 3),$$

and the reason that $\phi_2(x^2 + x - 6) = 0$ is that $\phi_2(x - 2) = 2 - 2 = 0$. ▲

Example Let F be $\mathbb{Q}$ and E be $\mathbb{C}$ in Theorem 22.4 and consider the evaluation homomorphism $\phi_i : \mathbb{Q}[x] \to \mathbb{C}$. Here

$$\phi_i(a_0 + a_1x + \cdots + a_nx^n) = a_0 + a_1i + \cdots + a_ni^n$$

and $\phi_i(x) = i$. Note that

$$\phi_i(x^2 + 1) = i^2 + 1 = 0,$$

so $x^2 + 1$ is in the kernel N of ϕ_i. ▲

Example Let F be $\mathbb{Q}$ and let E be $\mathbb{R}$ in Theorem 22.4 and consider the evaluation homomorphism $\phi_\pi : \mathbb{Q}[x] \to \mathbb{R}$. Here

$$\phi_\pi(a_0 + a_1x + \cdots + a_nx^n) = a_0 + a_1\pi + \cdots + a_n\pi^n.$$

It can be proved that $a_0 + a_1\pi + \cdots + a_n\pi^n = 0$ if and only if $a_i = 0$ for $i = 0, 1, \cdots, n$. Thus the kernel of ϕ_π is $\{0\}$, and ϕ_π is a one-to-one map. This shows that all *formal polynomials in π with rational coefficients* form a ring isomorphic to $\mathbb{Q}[x]$ in a natural way with $\phi_\pi(x) = \pi$. ▲

Definition Let F be a subfield of a field E, and let α be an element of E. Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ be in $F[x]$, and let $\phi_\alpha : F[x] \to E$ be the evaluation homomorphism of Theorem 22.4. Let $f(\alpha)$ denote

$$\phi_\alpha(f(x)) = a_0 + a_1\alpha + \cdots + a_n\alpha^n.$$

If $f(\alpha) = 0$, then α is a **zero of** $f(x)$. ■

In terms of this definition, we can rephrase the classical problem of finding all real numbers r such that $r^2 + r - 6 = 0$ by letting $F = \mathbb{Q}$ and $E = \mathbb{R}$ and *finding all* $\alpha \in \mathbb{R}$ *such that*

$$\phi_\alpha(x^2 + x - 6) = 0,$$

that is, finding all zeros of $x^2 + x - 6$ in $\mathbb{R}$. Both problems have the same answer, since

$$\{\alpha \in \mathbb{R} \mid \phi_\alpha(x^2 + x - 6) = 0\} = \{r \in \mathbb{R} \mid r^2 + r - 6 = 0\} = \{2, -3\}.$$

Theorem The polynomial $x^2 - 2$ has no zeros in the rational numbers. Thus $\sqrt{2}$ is not a rational number.

Proof Suppose that m/n for $m, n \in \mathbb{Z}$ is a rational number such that $(m/n)^2 = 2$. We assume that we have canceled any factors common to m and n, so that the fraction m/n is in lowest terms with $\gcd(m, n) = 1$. Then

$$m^2 = 2n^2,$$

where both m^2 and $2n^2$ are integers. Since m^2 and $2n^2$ are the same integer, and since 2 is a factor of $2n^2$, we see that 2 must be one of the factors of m^2. But as a square, m^2 has as factors the factors of m repeated twice. Thus m^2 must have two factors 2. Then $2n^2$ must have two factors 2, so n^2 must have 2 as a factor, and consequently n has 2 as a factor. We have deduced from $m^2 = 2n^2$ that both m and n must be divisible by 2, contradicting the fact that the fraction m/n is in lowest terms. Thus we have $2 \neq (m/n)^2$ for any $m, n \in \mathbb{Z}$. ◆

Thus the Pythagoreans ran right into the question of a solution of a polynomial equation, $x^2 - 2 = 0$. We refer the student to Shanks [36, Chapter 3], for a lively and totally delightful account of this Pythagorean dilemma and its significance in mathematics.

Factorization of Polynomials over a Field

The Division Algorithm in $F[x]$

The following theorem is the basic tool for our work in this section. Note the similarity with the division algorithm for $\mathbb{Z}$ given in Theorem 6.3, the importance of which has been amply demonstrated.

Theorem **(Division Algorithm for $F[x]$)** Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

and

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0$$

be two elements of $F[x]$, with a_n and b_m both nonzero elements of F and $m > 0$. Then there are unique polynomials $q(x)$ and $r(x)$ in $F[x]$ such that $f(x) = g(x)q(x) + r(x)$, where either $r(x) = 0$ or the degree of $r(x)$ is less than the degree m of $g(x)$.

Proof Consider the set $S = \{f(x) - g(x)s(x) \mid s(x) \in F[x]\}$. If $0 \in S$ then there exists an $s(x)$ such that $f(x) - g(x)s(x) = 0$, so $f(x) = g(x)s(x)$. Taking $q(x) = s(x)$ and $r(x) = 0$, we are done. Otherwise, let $r(x)$ be an element of minimal degree in S. Then

$$f(x) = g(x)q(x) + r(x)$$

for some $q(x) \in F[x]$. We must show that the degree of $r(x)$ is less than m. Suppose that

$$r(x) = c_t x^t + c_{t-1}x^{t-1} + \cdots + c_0,$$

with $c_j \in F$ and $c_t \neq 0$. If $t \geq m$, then

$$f(x) - q(x)g(x) - (c_t/b_m)x^{t-m}g(x) = r(x) - (c_t/b_m)x^{t-m}g(x), \qquad \textbf{(1)}$$

and the latter is of the form

$$r(x) - (c_t x^t + \text{terms of lower degree}),$$

which is a polynomial of degree lower than t, the degree of $r(x)$. However, the polynomial in Eq. (1) can be written in the form

$$f(x) - g(x)\left[q(x) + (c_t/b_m)x^{t-m}\right],$$

so it is in S, contradicting the fact that $r(x)$ was selected to have minimal degree in S. Thus the degree of $r(x)$ is less than the degree m of $g(x)$.

For uniqueness, if

$$f(x) = g(x)q_1(x) + r_1(x)$$

and

$$f(x) = g(x)q_2(x) + r_2(x),$$

then subtracting we have

$$g(x)[q_1(x) - q_2(x)] = r_2(x) - r_1(x).$$

Because either $r_2(x) - r_1(x) = 0$ or the degree of $r_2(x) - r_1(x)$ is less than the degree of $g(x)$, this can hold only if $q_1(x) - q_2(x) = 0$ so $q_1(x) = q_2(x)$. Then we must have $r_2(x) - r_1(x) = 0$ so $r_1(x) = r_2(x)$. ◆

Example Let us work with polynomials in $\mathbb{Z}_5[x]$ and divide

$$f(x) = x^4 - 3x^3 + 2x^3 + 4x - 1$$

by $g(x) = x^2 - 2x + 3$ to find $q(x)$ and $r(x)$ of Theorem 23.1. The long division should be easy to follow, but remember that we are in $\mathbb{Z}_5[x]$, so, for example, $4x - (-3x) = 2x$.

$$\begin{array}{r|l} & x^2 - x - 3 \\ \hline x^2 - 2x + 3 & x^4 - 3x^3 + 2x^2 + 4x - 1 \\ & \underline{x^4 - 2x^3 + 3x^2} \\ & \quad - x^3 - x^2 + 4x \\ & \quad \underline{- x^3 + 2x^2 - 3x} \\ & \qquad - 3x^2 + 2x - 1 \\ & \qquad \underline{- 3x^2 + x - 4} \\ & \qquad\qquad x + 3 \end{array}$$

Thus

$$q(x) = x^2 - x - 3, \quad \text{and} \quad r(x) = x + 3. \qquad \blacktriangle$$

Corollary **(Factor Theorem)** An element $a \in F$ is a zero of $f(x) \in F[x]$ if and only if $x - a$ is a factor of $f(x)$ in $F[x]$.

Proof Suppose that for $a \in F$ we have $f(a) = 0$. By Theorem 23.1, there exist $q(x)$, $r(x) \in F[x]$ such that

$$f(x) = (x - a)q(x) + r(x),$$

where either $r(x) = 0$ or the degree of $r(x)$ is less than 1. Thus we must have $r(x) = c$ for $c \in F$, so

$$f(x) = (x - a)q(x) + c.$$

Applying our evaluation homomorphism, $\phi_a : F[x] \to F$ of Theorem 22.4, we find

$$0 = f(a) = 0q(a) + c,$$

so it must be that $c = 0$. Then $f(x) = (x - a)q(x)$, so $x - a$ is a factor of $f(x)$.

Conversely, if $x - a$ is a factor of $f(x)$ in $F[x]$, where $a \in F$, then applying our evaluation homomorpohism ϕ_a to $f(x) = (x - a)q(x)$, we have $f(a) = 0q(a) = 0$. ◆

Example Working again in $\mathbb{Z}_5[x]$, note that 1 is a zero of

$$(x^4 + 3x^3 + 2x + 4) \in \mathbb{Z}_5[x].$$

Thus by Corollary 23.3, we should be able to factor $x^4 + 3x^3 + 2x + 4$ into $(x - 1)q(x)$ in $\mathbb{Z}_5[x]$. Let us find the factorization by long division.

$$\begin{array}{r|l} & x^3 + 4x^2 + 4x + 1 \\ \hline x - 1 & x^4 + 3x^3 + \quad\quad 2x + 4 \\ & \underline{x^4 - x^3} \\ & \quad 4x^3 \\ & \quad \underline{4x^3 - 4x^2} \\ & \quad\quad 4x^2 + 2x \\ & \quad\quad \underline{4x^2 - 4x} \\ & \quad\quad\quad x + 4 \\ & \quad\quad\quad \underline{x - 1} \\ & \quad\quad\quad\quad 0 \end{array}$$

Thus $x^4 + 3x^3 + 2x + 4 = (x - 1)(x^3 + 4x^2 + 4x + 1)$ in $\mathbb{Z}_5[x]$. Since 1 is seen to be a zero of $x^3 + 4x^2 + 4x + 1$ also, we can divide this polynomial by $x - 1$ and get

$$\begin{array}{r|l} & x^2 + 4 \\ \hline x - 1 & x^3 + 4x^2 + 4x + 1 \\ & \underline{x^3 - x^2} \\ & \quad 0 \quad + 4x + 1 \\ & \quad\quad\quad \underline{4x - 4} \\ & \quad\quad\quad\quad 0 \end{array}$$

Since $x^2 + 4$ still has 1 as a zero, we can divide again by $x - 1$ and get

$$\begin{array}{r|l} & x + 1 \\ \hline x - 1 & x^2 \quad\quad + 4 \\ & \underline{x^2 - x} \\ & \quad\quad x + 4 \\ & \quad\quad \underline{x - 1} \\ & \quad\quad\quad 0 \end{array}$$

Thus $x^4 + 3x^3 + 2x + 4 = (x - 1)^3(x + 1)$ in $\mathbb{Z}_5[x]$. ▲

Corollary A nonzero polynomial $f(x) \in F[x]$ of degree n can have at most n zeros in a field F.

Proof The preceding corollary shows that if $a_1 \in F$ is a zero of $f(x)$, then

$$f(x) = (x - a_1)q_1(x),$$

where, of course, the degree of $q_1(x)$ is $n - 1$. A zero $a_2 \in F$ of $q_1(x)$ then results in a factorization

$$f(x) = (x - a_1)(x - a_2)q_2(x).$$

Continuing this process, we arrive at

$$f(x) = (x - a_1) \cdots (x - a_r)q_r(x),$$

where $q_r(x)$ has no further zeros in F. Since the degree of $f(x)$ is n, at most n factors $(x - a_i)$ can appear on the right-hand side of the preceding equation, so $r \leq n$. Also, if $b \neq a_i$ for $i = 1, \cdots, r$ and $b \in F$, then

$$f(b) = (b - a_1) \cdots (b - a_r)q_r(b) \neq 0,$$

since F has no divisors of 0 and none of $b - a_i$ or $q_r(b)$ are 0 by construction. Hence the a_i for $i = 1, \cdots, r \leq n$ are all the zeros in F of $f(x)$. ◆

Corollary If G is a finite subgroup of the multiplicative group $\langle F^*, \cdot \rangle$ of a field F, then G is cyclic. In particular, the multiplicative group of all nonzero elements of a finite field is cyclic.

Proof By Theorem 11.12 as a finite abelian group, G is isomorphic to a direct product $\mathbb{Z}_{d_1} \times \mathbb{Z}_{d_2} \times \cdots \times \mathbb{Z}_{d_r}$, where each d_i is a power of a prime. Let us think of each of the $\mathbb{Z}_{d_i}$ as a cyclic group of order d_i in *multiplicative* notation. Let m be the least common multiple of all the d_i for $i = 1, 2, \cdots, r$; note that $m \leq d_1 d_2 \cdots d_r$. If $a_i \in \mathbb{Z}_{d_i}$, then ${a_i}^{d_i} = 1$, so ${a_i}^m = 1$ since d_i divides m. Thus for all $\alpha \in G$, we have $\alpha^m = 1$, so every element of G is zero of $x^m - 1$. But G has $d_1 d_2 \cdots d_r$ elements, while $x^m - 1$ can have at most m zeros in the field F by Corollary 23.5, so $m \geq d_1 d_2 \cdots d_r$. Hence $m = d_1 d_2 \cdots d_r$, so the primes involved in the prime powers $d_1, d_2, \cdots, d_r$ are distinct, and the group G is isomorphic to the cyclic group $\mathbb{Z}_m$. ◆

Irreducible Polynomials

Definition A nonconstant polynomial $f(x) \in F[x]$ is **irreducible over** F or is an **irreducible polynomial in** $F[x]$ if $f(x)$ cannot be expressed as a product $g(x)h(x)$ of two polynomials $g(x)$ and $h(x)$ in $F[x]$ both of lower degree than the degree of $f(x)$. If $f(x) \in F[x]$ is a nonconstant polynomial that is not irreducible over F, then $f(x)$ is **reducible over** F. ■

Example Theorem 22.11 shows that $x^2 - 2$ viewed in $\mathbb{Q}[x]$ has no zeros in $\mathbb{Q}$. This shows that $x^2 - 2$ is irreducible over $\mathbb{Q}$, for a factorization $x^2 - 2 = (ax + b)(cx + d)$ for $a, b, c, d \in \mathbb{Q}$ would give rise to zeros of $x^2 - 2$ in $\mathbb{Q}$. However, $x^2 - 2$ viewed in $\mathbb{R}[x]$ is not irreducible over $\mathbb{R}$, because $x^2 - 2$ factors in $\mathbb{R}[x]$ into $(x - \sqrt{2})(x + \sqrt{2})$. ▲

Example Let us show that $f(x) = x^3 + 3x + 2$ viewed in $\mathbb{Z}_5[x]$ is irreducible over $\mathbb{Z}_5$. If $x^3 + 3x + 2$ factored in $\mathbb{Z}_5[x]$ into polynomials of lower degree then there would exist at least one linear factor of $f(x)$ of the form $x - a$ for some $a \in \mathbb{Z}_5$. But then $f(a)$ would be 0, by Corollary 23.3. However, $f(0) = 2$, $f(1) = 1$, $f(-1) = -2$, $f(2) = 1$, and $f(-2) = -2$, showing that $f(x)$ has no zeros in $\mathbb{Z}_5$. Thus $f(x)$ is irreducible over $\mathbb{Z}_5$. This test for irreducibility by finding zeros works nicely for quadratic and cubic polynomials over a finite field with a small number of elements. ▲

Theorem Let $f(x) \in F[x]$, and let $f(x)$ be of degree 2 or 3. Then $f(x)$ is reducible over F if and only if it has a zero in F.

Proof If $f(x)$ is reducible so that $f(x) = g(x)h(x)$, where the degree of $g(x)$ and the degree of $h(x)$ are both less than the degree of $f(x)$, then since $f(x)$ is either quadratic or cubic, either $g(x)$ or $h(x)$ is of degree 1. If, say, $g(x)$ is of degree 1, then except for a possible factor in F, $g(x)$ is of the form $x - a$. Then $g(a) = 0$, which implies that $f(a) = 0$, so $f(x)$ has a zero in F.

Conversely, Corollary 23.3 shows that if $f(a) = 0$ for $a \in F$, then $x - a$ is a factor of $f(x)$, so $f(x)$ is reducible. ◆

Theorem If $f(x) \in \mathbb{Z}[x]$, then $f(x)$ factors into a product of two polynomials of lower degrees r and s in $\mathbb{Q}[x]$ if and only if it has such a factorization with polynomials of the same degrees r and s in $\mathbb{Z}[x]$.

Proof The proof is omitted here. ◆

Corollary If $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ is in $\mathbb{Z}[x]$ with $a_0 \neq 0$, and if $f(x)$ has a zero in $\mathbb{Q}$, then it has a zero m in $\mathbb{Z}$, and m must divide a_0.

Proof If $f(x)$ has a zero a in $\mathbb{Q}$, then $f(x)$ has a linear factor $x - a$ in $\mathbb{Q}[x]$ by Corollary 23.3. But then by Theorem 23.11, $f(x)$ has a factorization with a linear factor in $\mathbb{Z}[x]$, so for some $m \in \mathbb{Z}$ we must have

$$f(x) = (x - m)\left(x^{n-1} + \cdots - a_0/m\right).$$

Thus a_0/m is in $\mathbb{Z}$, so m divides a_0. ◆

Theorem **(Eisenstein Criterion)** Let $p \in \mathbb{Z}$ be a prime. Suppose that $f(x) = a_n x^n + \cdots + a_0$ is in $\mathbb{Z}[x]$, and $a_n \not\equiv 0 \pmod{p}$, but $a_i \equiv 0 \pmod{p}$ for all $i < n$, with $a_0 \not\equiv 0 \pmod{p^2}$. Then $f(x)$ is irreducible over $\mathbb{Q}$.

Proof By Theorem 23.11 we need only show that $f(x)$ does not factor into polynomials of lower degree in $\mathbb{Z}[x]$. If

$$f(x) = (b_r x^r + \cdots + b_0)(c_s x^s + \cdots + c_0)$$

is a factorization in $\mathbb{Z}[x]$, with $b_r \not\equiv 0$, $c_s \not\equiv 0$ and $r, s < n$, then $a_0 \not\equiv 0 \pmod{p^2}$ implies that b_0 and c_0 are not both congruent to 0 modulo p. Suppose that $b_0 \not\equiv 0 \pmod{p}$ and $c_0 \equiv 0 \pmod{p}$. Now $a_n \not\equiv 0 \pmod{p}$ implies that $b_r, c_s \not\equiv 0 \pmod{p}$, since $a_n = b_r c_s$. Let m be the smallest value of k such that $c_k \not\equiv 0 \pmod{p}$. Then

$$a_m = b_0 c_m + b_1 c_{m-1} + \cdots + \begin{cases} b_m c_0 \text{ if } r \geq m, \\ b_r c_{m-r} \text{ if } r < m. \end{cases}$$

The fact that neither b_0 nor c_m are congruent to 0 modulo p while $c_{m-1}, \cdots, c_0$ are all congruent to 0 modulo p implies that $a_m \not\equiv 0$ modulo p, so $m = n$. Consequently, $s = n$, contradicting our assumption that $s < n$; that is, that our factorization was nontrivial. ◆

Corollary The polynomial

$$\Phi_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1$$

is irreducible over $\mathbb{Q}$ for any prime p.

Proof Again by Theorem 23.11, we need only consider factorizations in $\mathbb{Z}[x]$. We remarked following Theorem 22.5 that its proof actually shows that evaluation homomorphims can be used for commutative rings. Here we want to use the evaluation homomorphism $\phi_{x+1} : \mathbb{Q}[x] \to \mathbb{Q}[x]$. It is natural for us to denote $\phi_{x+1}(f(x))$ by $f(x+1)$ for $f(x) \in \mathbb{Q}[x]$. Let

$$g(x) = \Phi_p(x+1) = \frac{(x+1)^p - 1}{(x+1) - 1} = \frac{x^p + \binom{p}{1} x^{p-1} + \cdots + px}{x}.$$

The coefficient of x^{p-r} for $0 < r < p$ is the binomial coefficient $p!/[r!(p-r)!]$ which is divisible by p because p divides $p!$ but does not divide either $r!$ or $(p-r)!$ when $0 < r < p$. Thus

$$g(x) = x^{p-1} + \binom{p}{1} x^{p-2} + \cdots + p$$

satisifies the Eisenstein criterion for the prime p and is thus irreducible over $\mathbb{Q}$. But if $\Phi_p(x) = h(x)r(x)$ were a nontrivial factorization of $\Phi_p(x)$ in $\mathbb{Z}[x]$, then

$$\Phi_p(x+1) = g(x) = h(x+1)r(x+1)$$

would give a nontrivial factorization of $g(x)$ in $\mathbb{Z}[x]$. Thus $\Phi_p(x)$ must also be irreducible over $\mathbb{Q}$. ◆

Theorem Let $p(x)$ be an irreducible polynomial in $F[x]$. If $p(x)$ divides $r(x)s(x)$ for $r(x), s(x) \in F[x]$, then either $p(x)$ divides $r(x)$ or $p(x)$ divides $s(x)$.

Corollary If $p(x)$ is irreducible in $F[x]$ and $p(x)$ divides the product $r_1(x) \cdots r_n(x)$ for $r_i(x) \in F[x]$, then $p(x)$ divides $r_i(x)$ for at least one i.

Theorem If F is a field, then every nonconstant polynomial $f(x) \in F[x]$ can be factored in $F[x]$ into a product of irreducible polynomials, the irreducible polynomials being unique except for order and for unit (that is, nonzero constant) factors in F.

Proof Let $f(x) \in F[x]$ be a nonconstant polynomial. If $f(x)$ is not irreducible, then $f(x) = g(x)h(x)$, with the degree of $g(x)$ and the degree of $h(x)$ both less than the degree of $f(x)$. If $g(x)$ and $h(x)$ are both irreducible, we stop here. If not, at least one of them factors into polynomials of lower degree. Continuing this process, we arrive at a factorization

$$f(x) = p_1(x)p_2(x)\cdots p_r(x),$$

where $p_i(x)$ is irreducible for $i = 1, 2, \cdots, r$.

It remains for us to show uniqueness. Suppose that

$$f(x) = p_1(x)p_2(x)\cdots p_r(x) = q_1(x)q_2(x)\cdots q_s(x)$$

are two factorizations of $f(x)$ into irreducible polynomials. Then by Corollary 23.19, $p_1(x)$ divides some $q_j(x)$, let us assume $q_1(x)$. Since $q_1(x)$ is irreducible,

$$q_1(x) = u_1 p_1(x),$$

where $u_1 \neq 0$, but u_1 is in F and thus is a unit. Then substituting $u_1 p_1(x)$ for $q_1(x)$ and canceling, we get

$$p_2(x)\cdots p_r(x) = u_1 q_2(x)\cdots q_s(x).$$

By a similar argument, say $q_2(x) = u_2 p_2(x)$, so

$$p_3(x)\cdots p_r(x) = u_1 u_2 q_3(x)\cdots q_s(x).$$

Continuing in this manner, we eventually arrive at

$$1 = u_1 u_2 \cdots u_r q_{r+1}(x)\cdots q_s(x).$$

This is only possible if $s = r$, so that this equation is actually $1 = u_1 u_2 \cdots u_r$. Thus the irreducible factors $p_i(x)$ and $q_j(x)$ were the same except possibly for order and unit factors. ◆

POSSIBLE QUESTIONS

2 Mark Questions:

1. Find all the zeros of $f(x) = x^4+4$ in Z_5.
2. Define content of the polynomial.
3. Define Monic polynomial.
4. Define primitive polynomial.
5. Define product of two polynomial with example.

6 Mark Questions:

6. State and prove Eisenstein criterion.
7. State and prove Factor theorem
8. Let f(x) and g(x) be any two polynomial then prove that
   deg(f(x).g(x)) = deg (f(x)) + deg (g(x)).
9. The polynomial x^3+2x^2+2x+1 factored into the linear factor in $Z_7[x]$. Find this factorization.
10. State and prove division algorithm over polynomial ring.
11. Prove that the polynomial $\phi_p(x) = \frac{x^{p-1}}{x-1} = x^{p-1} + x^{p-2} + ... + x + 1$ is irreducible over Q for any prime p.
12. State and prove Evaluation homomorphism for field theory.
13. prove that the set of all polynomials over Z with addition and multiplication is a ring.
14. State and prove Gauss lemma.
15. Let F be a field. If f(x) in F[x] and deg (f(x)) = 2 or 3 then prove that f(x) is reducible over F if and only if f(x) has a zero in F.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari (Po),
Coimbatore –641 021

Code: 16MMU403

Semester : IV

UNIT -II

PART A (20x1=20 Marks)

(Question Nos. 1 to 20 Online Examinations)

Possible Questions

Question	Choice 1	Choice 2	Choice 3	Choice 4	Answer
_____ is an Integral domain.	skew Field	Field	ring	group	Field
Field is an _____	integer monic	Eucledian ring	integers	integral domain	integral domain
The smallest such positive integer n is called _____ if no positive integer then r is said to be a characteristic zero or infinite.	Euclidean ring R	the characteristics of a ring R	infinite integral domain	Division ring R	the characteristics of a ring R
The smallest such positive integer n is called the characteristics of a ring R if no _____ integer then r is said to be a characteristic zero or infinite.	positive	real	rational	complex	positive
The smallest such positive integer n is called the characteristics of a ring R if no positive integer then r is said to be a _____	characteristic zero or infinite	characteristic one	characteristic finite	characteristic ring	characteristic zero or infinite
A_____ has no proper ideals	field	group	Field	ring	field
A field has no _____	right ideal	proper ideals	one-sided ideal	two-sided ideal	proper ideals
An _____ generated by a single element of itself it called a principle ideal	group	ideal	Field	ring	ideal
An ideal generated by a _____ element of itself it called a principle ideal	two-sided ideal	one-sided ideal	double	single	single
An ideal generated by a single element of itself it called a _____	integral domain	principle ideal	ideal	Eucledian ring	principle ideal
An ---------- possess a unit element.	integer monic	Division ring	Euclidean ring	integral domain	Euclidean ring
An Euclidean ring possess a _____ element.	field	unit	double	no	unit
An _____ is said to be of characteristics zero if the relation Ma = 0, where a ≠0 is in D and where m is an integer can hold only if m=0	skew Field	Integral domain D	Division ring R	the characteristics of a ring R	Integral domain D
A ---------- of R into R' is said to be an isomorphism if it is one- one mapping.	homomorphism	isomorphism	automorphism	monomorphism	homomorphism
A homomorphism of R into R' is said to be an ---------- if it is one- one mapping.	isomorphism	identity	integral domain	Eucledian ring	isomorphism
A homomorphism of R into R' is said to be an isomorphism if it is --------- mapping.	onto	one- one	into	into & onto	one- one
We cannot define the ---------- of the zero polynomial.	sum	degree	order	power	degree
A---------- is a constant if it degree is zero.	monomial	trinomial	polynomial	binomial	polynomial

UNIT-II

SYLLABUS

Divisibility in integral domains, irreducibles, primes, unique factorization domains, Euclidean domains.

KAHE

Divisibility in Integral Domains

Irreducibles, Primes

In the previous two chapters, we focused on factoring polynomials over the integers or a field. Several of those results, unique factorization in $Z[x]$ and the division algorithm for $F[x]$, for instance, are natural counterparts to theorems about the integers. In this chapter and the next, we examine factoring in a more abstract setting.

DEFINITION Associates, Irreducibles, Primes
Elements a and b of an integral domain D are called *associates* if $a = ub$ where u is a unit of D. A nonzero element a of an integral domain D is called *irreducible* if a is not a unit and, whenever $b, c \in D$ with $a = bc$, then b or c is a unit. A nonzero element a of an integral domain D is called *prime* if a is not a unit and $a \mid bc$ implies $a \mid b$ or $a \mid c$.

Example 1 We exhibit an irreducible in $Z[\sqrt{-3}]$ that is not prime. Here $N(a + b\sqrt{-3}) = a^2 + 3b^2$. Consider $1 + \sqrt{-3}$. Suppose we can factor this as xy where neither x nor y is a unit. Then $N(xy) = N(x)N(y) = N(1 + \sqrt{-3}) = 4$, and it follows that $N(x) = 2$. But there are no integers a and b satisfying $a^2 + 3b^2 = 2$. Thus, x or y is a unit and $1 + \sqrt{-3}$ is an irreducible. To verify that it is not prime, we observe that $(1 + \sqrt{-3})(1 - \sqrt{-3}) = 4 = 2 \cdot 2$ so that $1 + \sqrt{-3}$ divides $2 \cdot 2$. On the other hand, for integers a and b to exist so that $2 = (1 + \sqrt{-3})(a + b\sqrt{-3}) = (a - 3b) + (a + b)\sqrt{-3}$, we must have $a - 3b = 2$ and $a + b = 0$, which is impossible. □

Theorem 20.1 *Prime Implies Irreducible*
In an integral domain, every prime is irreducible.

Proof. Suppose a is a prime in an integral domain and $a = bc$. We must show that b or c is a unit. By definition of prime, we know $a \mid b$ or $a \mid c$. Say, $at = b$. Then $b \cdot 1 = b = at = (bc)t = b(ct)$ and, by cancellation, $1 = ct$. Thus, c is a unit. ■

Recall that a principal ideal domain is an integral domain in which every ideal has the form $\langle a \rangle$. The next theorem reveals a circumstance in which primes and irreducibles are equivalent.

Theorem 20.2 *PID Implies Irreducible Equals Prime*
In a principal ideal domain, an element is irreducible if and only if it is prime.

Proof. Theorem 20.1 shows primes are irreducible. To prove the converse, let a be an irreducible element of a principal ideal domain D and suppose $a \mid bc$. We must show $a \mid b$ or $a \mid c$. Consider the ideal $I = \{ax + by \mid x,$

$y \in D\}$ and let $\langle d \rangle = I$. Since $a \in I$, we can write $a = dr$, and because a is irreducible, d is a unit or r is a unit. If d is a unit, then $I = D$ and we may write $1 = ax + by$. Then $c = acx + bcy$, and since a divides both terms on the right, a also divides c.

On the other hand, if r is a unit, then $\langle a \rangle = \langle d \rangle = I$, and because $b \in I$, there is an element t in D such that $at = b$. Thus, a divides b. ■

Example 2 We show $Z[x]$ is not a principal ideal domain. Consider the ideal $I = \{ax + 2b \mid a, b \in Z\}$. We claim I is not of the form $\langle h(x) \rangle$. If this were so, there would be $f(x)$ and $g(x)$ in $Z[x]$ such that $2 = h(x)f(x)$ and $x = h(x)g(x)$, since both 2 and x belong to I. By the degree rule (exercise 16 of Chapter 18), $0 = \deg 2 = \deg h(x) + \deg f(x)$ so that $h(x)$ is a constant polynomial. To determine which constant, we observe $2 = h(1)f(1)$. Thus, $h(1) = \pm 1$ or ± 2. Since 1 is not in I, we must have $h(x) = \pm 2$. But then $x = \pm 2g(x)$, which is nonsense. □

Unique Factorization Domains

We now have the necessary terminology to formalize the idea of unique factorization.

DEFINITION Unique Factorization Domain (UFD)
An integral domain D is a *unique factorization domain* if

1. every nonzero element of D that is not a unit can be written as a product of irreducibles of D, and
2. the factorization into irreducibles is unique up to associates and the order in which the factors appear.

Another way to formulate part 2 of this definition is the following. If $p_1^{n_1}p_2^{n_2} \cdots p_r^{n_r}$ and $q_1^{m_1}q_2^{m_2} \cdots q_s^{m_s}$ are two factorizations of some element as a product of irreducibles, where none of the p_i's are associates and none of the q_j's are associates, then $r = s$, and each $p_i^{n_i}$ is an associate of one and only one $q_j^{m_j}$.

Lemma *Ascending Chain Condition for a PID*
In a principal ideal domain, any strictly increasing chain of ideals $I_1 \subset I_2 \subset \cdots$ *must be finite in length.*

Proof. Let $I_1 \subset I_2 \subset \cdots$ be a chain of strictly increasing ideals in an integral domain D, and let I be the union of all the ideals in this chain. We leave it as an exercise to verify that I is an ideal of D.

Then, since D is a principal ideal domain, there is an element a in D such that $I = \langle a \rangle$. Because $a \in I$ and $I = \cup_k I_k$, a belongs to some member of the chain, say, $a \in I_n$. Clearly, then, for any member I_i of the chain, we have $I_i \subseteq I = \langle a \rangle \subseteq I_n$ so that I_n must be the last member of the chain. ■

Theorem 20.3 *PID Implies UFD*
Every principal ideal domain is a unique factorization domain.

Proof. Let D be a principal ideal domain. We first show that any nonzero element of D that is not a unit is a product of irreducibles (the product could consist of only one factor). To do this, let a_0 be a nonzero, nonunit that is not irreducible. Then we may write $a_0 = b_1 a_1$ where neither b_1 nor a_1 is a unit. Now, if both b_1 and a_1 can be written as a product of irreducibles, then so can a_0. Thus, we may assume one of b_1 or a_1 cannot be written as a product of irreducibles, say a_1. Then, as before, we may write $a_1 = b_2 a_2$ where neither b_2 nor a_2 is a unit. Continuing in this fashion, we obtain an infinite sequence $b_1, b_2, \ldots$ of elements that are not units in D and an infinite sequence $a_0, a_1, a_2, \ldots$ of nonzero elements of D, with $a_n = b_{n+1} a_{n+1}$ for each n. Since b_{n+1} is not a unit, we have $\langle a_n \rangle \subset \langle a_{n+1} \rangle$ for each n (see exercise 3). Hence, $\langle a_0 \rangle \subset \langle a_1 \rangle \subset \cdots$ is an infinite strictly increasing chain of ideals. This contradicts the preceding lemma, so we conclude a_0 is, indeed, a product of irreducibles.

It remains to show that the factorization is unique up to associates and the order in which the factors appear. To do this, suppose some element a of D can be written

$$a = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s,$$

where the p's and q's are irreducible and repetition is permitted. We induct on r. If $r = 1$, then a is irreducible and, clearly, $s = 1$ and $p_1 = q_1$. So

we may assume that any element that can be expressed as a product of fewer than r irreducible factors can be done so in only one way (up to order and associates). Since p_1 divides $q_1q_2 \cdots q_s$, it must divide some q_i (see exercise 24), say, $p_1 \mid q_1$. Then, $q_1 = up_1$, where u is a unit of D. Thus,

$$ua = up_1p_2 \cdots p_r = q_1(uq_2) \cdots q_s$$

and, by cancellation,

$$p_2 \cdots p_r = (uq_2) \cdots q_s.$$

The induction hypothesis now tells us that these two factorizations are identical up to associates and the order in which the factors appear. Hence, the same is true about the two factorizations of a. ■

Corollary *$F[x]$ Is a UFD*
Let F be a field. Then $F[x]$ is a unique factorization domain.

Proof. By Theorem 18.3, $F[x]$ is a principal ideal domain. So, $F[x]$ is a unique factorization domain, as well. ■

Example 3 Let

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0 \in Z[x],$$

and suppose p is prime such that

$$p \nmid a_n,\ p \mid a_{n-1},\ \ldots,\ p \mid a_0,$$

and $p^2 \nmid a_0$. We will prove that $f(x)$ is irreducible over Q. If $f(x)$ is reducible over Q, we know there exist elements $g(x)$ and $h(x)$ in $Z[x]$ such that $f(x) = g(x)h(x)$ and $1 \leq \deg g(x), \deg h(x) < n$. Let $\bar{f}(x)$, $\bar{g}(x)$, and $\bar{h}(x)$ be the polynomials in $Z_p[x]$ obtained from $f(x)$, $g(x)$, and $h(x)$ by reducing all coefficients modulo p. Then, since p divides all the coefficients of $f(x)$ except a_n, we have $\bar{a}_nx^n = \bar{f}(x) = \bar{g}(x)\bar{h}(x)$. Since Z_p is a field, $Z_p[x]$ is a unique factorization domain. Thus, $x \mid \bar{g}(x)$ and $x \mid \bar{h}(x)$. So, $\bar{g}(0) = \bar{h}(0) = 0$ and, therefore, $p \mid g(0)$ and $p \mid h(0)$. But, then, $p^2 \mid g(0)h(0) = f(0) = a_0$, a contradiction. □

Euclidean Domains

Another important kind of integral domain is a Euclidean domain.

DEFINITION Euclidean Domain
An integral domain D is called a *Euclidean domain* if there is a function d from the nonzero elements of D to the nonnegative integers such that

1. $d(a) \leq d(ab)$ for all nonzero a, b in D; and
2. if $a, b \in D$, $b \neq 0$, then there exist elements q and r in D such that $a = bq + r$ where $r = 0$ or $d(r) < d(b)$.

Example 4 The ring Z is a Euclidean domain with $d(a) = |a|$ (the absolute value of a). □

Example 5 Let F be a field. Then $F[x]$ is a Euclidean domain with $d(f(x)) = \deg f(x)$ (see Theorem 18.2). □

Example 6 The ring of Gaussian integers

$$Z[i] = \{a + bi \mid a, b \in Z\}$$

is a Euclidean domain with $d(a + bi) = a^2 + b^2$. Unlike the previous two examples, the function d does not obviously satisfy the necessary conditions. That $d(x) \leq d(xy)$ for $x, y \in Z[i]$ follows directly from the fact that $d(xy) = d(x)d(y)$ (exercise 5). If $x, y \in Z[i]$ and $y \neq 0$ then $xy^{-1} \in Q[i]$, the field of quotients of $Z[i]$ (exercise 41 of Chapter 17). Say, $xy^{-1} = s + ti$ where $s, t \in Q$. Now let m be the integer nearest s, and let n be the integer nearest t. (These integers may not be uniquely determined but that does not matter.) Thus, $|m - s| \leq \frac{1}{2}$ and $|n - t| \leq \frac{1}{2}$. Then

$$\begin{aligned} xy^{-1} = s + ti &= (m - m + s) + (n - n + t)i \\ &= (m + ni) + [(s - m) + (t - n)i]. \end{aligned}$$

So,

$$x = (m + ni)y + [(s - m) + (t - n)i]y.$$

Theorem 20.4 *ED (Euclidean Domain) Implies PID*
Every Euclidean domain is a principal ideal domain.

Proof. Let D be a Euclidean domain and I a nonzero ideal of D. Among all the elements of I, let a be such that $d(a)$ is minimum. Then $I = \langle a \rangle$. For, if $b \in I$, there are elements q and r such that $b = aq + r$ where $r = 0$ or $d(r) < d(a)$. But $r = b - aq \in I$, so $d(r)$ cannot be less than $d(a)$. Thus, $r = 0$ and $b \in \langle a \rangle$. ■

Corollary *ED Implies UFD*
Every Euclidean domain is a unique factorization domain.

We may summarize our theorems and remarks as follows:

$$\text{ED} \Rightarrow \text{PID} \Rightarrow \text{UFD}$$
$$\text{UFD} \not\Rightarrow \text{PID} \not\Rightarrow \text{ED}$$

Theorem 20.5 *D a UFD Implies D[x] a UFD*
If D is a unique factorization domain, then D[x] is a unique factorization.

We conclude this chapter with an example of an integral domain that is not a unique factorization domain.

Example 7 The ring $Z[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in Z\}$ is an integral domain but not a unique factorization domain. It is straightforward that $Z[\sqrt{-5}]$ is an integral domain (see exercise 9 of Chapter 15). To verify that unique factorization does not hold, we mimic the method used in Example 1 with $N(a + b\sqrt{-5}) = a^2 + 5b^2$. Since $N(xy) = N(x)N(y)$ and $N(x) = 1$ if and only if x is a unit (see exercise 14), it follows that the only units of $Z[\sqrt{-5}]$ are ± 1.

Now consider the following factorizations:

$$46 = 2 \cdot 23,$$
$$46 = (1 + 3\sqrt{-5})(1 - 3\sqrt{-5}).$$

We claim that each of these four factors is irreducible over $Z[\sqrt{-5}]$. Suppose, say, $2 = xy$ where $x, y \in Z[\sqrt{-5}]$ and neither is a unit. Then $4 = N(2) = N(x)N(y)$ and, therefore, $N(x) = N(y) = 2$, which is impossible. Likewise, if $23 = xy$ were a nontrivial factorization, then $N(x) = 23$. Thus, there would be integers a and b such that $a^2 + 5b^2 = 23$. Clearly, no such integers exist. The same argument applies to $1 \pm 3\sqrt{-5}$. □

POSSIBLE QUESTIONS

2 Mark Questions:

1. Prove that in the integral domain every prime is irreducible.
2. Define Principal ideal domain.
3. Define Euclidean domain.
4. Define irreducible.
5. Define prime.

6 Mark Questions:

6. Let F be a field and let p(x) in F[x] be irreducible polynomial over F then prove $F[x]/\langle p(x)\rangle$ is a field.
7. Prove that in a principal ideal domain, any strictly increasing ideals $I_1 \subset I_2 \subset \ldots\ldots$must be finite in lengh.
8. Prove if F is a field then every non constant polynomial f(x) in F[x] can be factored as a irreducible polynomial or product of irreducible polynomials being unique factor in F.
9. Let F be a field and let p(x) in F[x] then prove $\langle p(x)\rangle$ is a maximal ideal in F[x] if and only if p(x) is irreducible polynomial over F.
10. Prove if F is a field then F[x] is an unique factorization domain.
11. Write bout the Gaussian integer with example.
12. Let F be a field and let p(x), a(x) and b(x) in F[x]. Prove that if p(x) be a irreducible polynomial over F then p(x) divides a(x) or p(x) divides b(x).
13. Prove that every Euclidean domain is a Principal ideal domain.
14. State and prove Unique factorization theorem.
15. Prove that every principal ideal domian is an unique factorization domain.

UNIT-III

SYLLABUS

Dual spaces, dual basis, double dual, transpose of a linear transformation and its matrix in the dual basis, annihilators, Eigen spaces of a linear operator, diagonalizability, invariant subspaces and Cayley-Hamilton theorem, the minimal polynomial for a linear operator.

KAHE

Eigen space of a Linear operator:

Characteristic Values

Definition. *Let* V *be a vector space over the field* F *and let* T *be a linear operator on* V. *A* **characteristic value** *of* T *is a scalar* c *in* F *such that there is a non-zero vector* α *in* V *with* $T\alpha = c\alpha$. *If* c *is a characteristic value of* T, *then*

(a) *any* α *such that* $T\alpha = c\alpha$ *is called a* **characteristic vector** *of* T *associated with the characteristic value* c;

(b) *the collection of all* α *such that* $T\alpha = c\alpha$ *is called the* **characteristic space** *associated with* c.

Characteristic values are often called characteristic roots, latent roots, eigenvalues, proper values, or spectral values. In this book we shall use only the name 'characteristic values.'

If T is any linear operator and c is any scalar, the set of vectors α such that $T\alpha = c\alpha$ is a subspace of V. It is the null space of the linear transformation $(T - cI)$. We call c a characteristic value of T if this subspace is different from the zero subspace, i.e., if $(T - cI)$ fails to be 1:1. If the underlying space V is finite-dimensional, $(T - cI)$ fails to be 1:1 precisely when its determinant is different from 0. Let us summarize.

Theorem

Let T *be a linear operator on a finite-dimensional space* V *and let* c *be a scalar. The following are equivalent.*

(i) c *is a characteristic value of* T.
(ii) *The operator* $(T - cI)$ *is singular (not invertible).*
(iii) *det* $(T - cI) = 0$.

The determinant criterion (iii) is very important because it tells us where to look for the characteristic values of T. Since $\det(T - cI)$ is a polynomial of degree n in the variable c, we will find the characteristic values as the roots of that polynomial. Let us explain carefully.

If $\mathcal{B}$ is any ordered basis for V and $A = [T]_{\mathcal{B}}$, then $(T - cI)$ is invertible if and only if the matrix $(A - cI)$ is invertible. Accordingly, we make the following definition.

Definition. *If* A *is an* n × n *matrix over the field* F, *a* **characteristic value of** A **in** F *is a scalar* c *in* F *such that the matrix* (A − cI) *is singular* (*not invertible*).

Since c is a characteristic value of A if and only if $\det(A - cI) = 0$, or equivalently if and only if $\det(cI - A) = 0$, we form the matrix $(xI - A)$ with polynomial entries, and consider the polynomial $f = \det(xI - A)$. Clearly the characteristic values of A in F are just the scalars c in F such that $f(c) = 0$. For this reason f is called the **characteristic polynomial** of A. It is important to note that f is a monic polynomial which has degree exactly n. This is easily seen from the formula for the determinant of a matrix in terms of its entries.

Lemma. *Similar matrices have the same characteristic polynomial.*

Proof. If $B = P^{-1}AP$, then

$$\begin{aligned}\det(xI - B) &= \det(xI - P^{-1}AP)\\ &= \det(P^{-1}(xI - A)P)\\ &= \det P^{-1} \cdot \det(xI - A) \cdot \det P\\ &= \det(xI - A). \quad \blacksquare\end{aligned}$$

This lemma enables us to define sensibly the characteristic polynomial of the operator T as the characteristic polynomial of any $n \times n$ matrix which represents T in some ordered basis for V. Just as for matrices, the characteristic values of T will be the roots of the characteristic polynomial for T. In particular, this shows us that T cannot have more than n distinct characteristic values. It is important to point out that T may not have any characteristic values.

EXAMPLE Let T be the linear operator on R^2 which is represented

in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

The characteristic polynomial for T (or for A) is

$$\det (xI - A) = \begin{vmatrix} x & 1 \\ -1 & x \end{vmatrix} = x^2 + 1.$$

Since this polynomial has no real roots, T has no characteristic values. If U is the linear operator on C^2 which is represented by A in the standard ordered basis, then U has two characteristic values, i and $-i$. Here we see a subtle point. In discussing the characteristic values of a matrix A, we must be careful to stipulate the field involved. The matrix A above has no characteristic values in R, but has the two characteristic values i and $-i$ in C.

EXAMPLE Let A be the (real) 3×3 matrix

$$\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}.$$

Then the characteristic polynomial for A is

$$\begin{vmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{vmatrix} = x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

Thus the characteristic values of A are 1 and 2.

Suppose that T is the linear operator on R^3 which is represented by A in the standard basis. Let us find the characteristic vectors of T associated with the characteristic values, 1 and 2. Now

$$A - I = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix}.$$

It is obvious at a glance that $A - I$ has rank equal to 2 (and hence $T - I$ has nullity equal to 1). So the space of characteristic vectors associated with the characteristic value 1 is one-dimensional. The vector $\alpha_1 = (1, 0, 2)$ spans the null space of $T - I$. Thus $T\alpha = \alpha$ if and only if α is a scalar multiple of α_1. Now consider

$$A - 2I = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix}.$$

Evidently $A - 2I$ also has rank 2, so that the space of characteristic vectors associated with the characteristic value 2 has dimension 1. Evidently $T\alpha = 2\alpha$ if and only if α is a scalar multiple of $\alpha_2 = (1, 1, 2)$.

Definition. *Let* T *be a linear operator on the finite-dimensional space* V. *We say that* T *is* **diagonalizable** *if there is a basis for* V *each vector of which is a characteristic vector of* T.

The reason for the name should be apparent; for, if there is an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ for V in which each α_i is a characteristic vector of T, then the matrix of T in the ordered basis $\mathcal{B}$ is diagonal. If $T\alpha_i = c_i\alpha_i$, then

$$[T]_{\mathcal{B}} = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}.$$

We certainly do not require that the scalars $c_1, \ldots, c_n$ be distinct; indeed, they may all be the same scalar (when T is a scalar multiple of the identity operator).

One could also define T to be diagonalizable when the characteristic vectors of T span V. This is only superficially different from our definition, since we can select a basis out of any spanning set of vectors.

Lemma. *Let* T *be a linear operator on the finite-dimensional space* V. *Let* $c_1, \ldots, c_k$ *be the distinct characteristic values of* T *and let* W_i *be the space of characteristic vectors associated with the characteristic value* c_i. *If* $W = W_1 + \cdots + W_k$, *then*

$$\dim W = \dim W_1 + \cdots + \dim W_k.$$

In fact, if $\mathcal{B}_i$ *is an ordered basis for* W_i, *then* $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ *is an ordered basis for* W.

Proof. The space $W = W_1 + \cdots + W_k$ is the subspace spanned by all of the characteristic vectors of T. Usually when one forms the sum W of subspaces W_i, one expects that dim $W <$ dim $W_1 + \cdots +$ dim W_k because of linear relations which may exist between vectors in the various spaces. This lemma states that the characteristic spaces associated with different characteristic values are independent of one another.

Suppose that (for each i) we have a vector β_i in W_i, and assume that $\beta_1 + \cdots + \beta_k = 0$. We shall show that $\beta_i = 0$ for each i. Let f be any polynomial. Since $T\beta_i = c_i\beta_i$, the preceding lemma tells us that

$$\begin{aligned} 0 = f(T)0 &= f(T)\beta_1 + \cdots + f(T)\beta_k \\ &= f(c_1)\beta_1 + \cdots + f(c_k)\beta_k. \end{aligned}$$

Choose polynomials $f_1, \ldots, f_k$ such that

$$f_i(c_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Then

$$\begin{aligned} 0 = f_i(T)0 &= \sum_j \delta_{ij}\beta_j \\ &= \beta_i. \end{aligned}$$

Now, let $\mathcal{B}_i$ be an ordered basis for W_i, and let $\mathcal{B}$ be the sequence $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$. Then $\mathcal{B}$ spans the subspace $W = W_1 + \cdots + W_k$. Also, $\mathcal{B}$ is a linearly independent sequence of vectors, for the following reason. Any linear relation between the vectors in $\mathcal{B}$ will have the form $\beta_1 + \cdots + \beta_k = 0$, where β_i is some linear combination of the vectors in $\mathcal{B}_i$. From what we just did, we know that $\beta_i = 0$ for each i. Since each $\mathcal{B}_i$ is linearly independent, we see that we have only the trivial linear relation between the vectors in $\mathcal{B}$. ∎

Theorem

Let T *be a linear operator on a finite-dimensional space* V.

Let $c_1, \ldots, c_k$ *be the distinct characteristic values of* T *and let* W_i *be the null space of* $(T - c_iI)$. *The following are equivalent.*

(i) T *is diagonalizable.*

(ii) *The characteristic polynomial for* T *is*

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$$

and dim $W_i = d_i$, $i = 1, \ldots, k$.

(iii) *dim* $W_1 + \cdots + dim\ W_k = dim\ V$.

Proof. We have observed that (i) implies (ii). If the characteristic polynomial f is the product of linear factors, as in (ii), then $d_1 + \cdots + d_k = \dim V$. For, the sum of the d_i's is the degree of the characteristic polynomial, and that degree is dim V. Therefore (ii) implies (iii). Suppose (iii) holds. By the lemma, we must have $V = W_1 + \cdots + W_k$, i.e., the characteristic vectors of T span V. ∎

The matrix analogue of Theorem 2 may be formulated as follows. Let A be an $n \times n$ matrix with entries in a field F, and let $c_1, \ldots, c_k$ be the distinct characteristic values of A in F. For each i, let W_i be the space of column matrices X (with entries in F) such that

$$(A - c_iI)X = 0,$$

and let $\mathcal{B}_i$ be an ordered basis for W_i. The bases $\mathcal{B}_1, \ldots, \mathcal{B}_k$ collectively string together to form the sequence of columns of a matrix P:

$$P = [P_1, P_2, \ldots] = (\mathcal{B}_1, \ldots, \mathcal{B}_k).$$

The matrix A is similar over F to a diagonal matrix if and only if P is a square matrix. When P is square, P is invertible and $P^{-1}AP$ is diagonal.

EXAMPLE 3. Let T be the linear operator on R^3 which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

Let us indicate how one might compute the characteristic polynomial, using various row and column operations:

$$\begin{vmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{vmatrix} = \begin{vmatrix} x-5 & 0 & 6 \\ 1 & x-2 & -2 \\ -3 & 2-x & x+4 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-5 & 0 & 6 \\ 1 & 1 & -2 \\ -3 & -1 & x+4 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-5 & 0 & 6 \\ 1 & 1 & -2 \\ -2 & 0 & x+2 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-5 & 6 \\ -2 & x+2 \end{vmatrix}$$

$$= (x-2)(x^2-3x+2)$$
$$= (x-2)^2(x-1).$$

What are the dimensions of the spaces of characteristic vectors associated with the two characteristic values? We have

$$A - I = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}$$

$$A - 2I = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix}.$$

We know that $A - I$ is singular and obviously rank $(A - I) \geq 2$. Therefore, rank $(A - I) = 2$. It is evident that rank $(A - 2I) = 1$.

Let W_1, W_2 be the spaces of characteristic vectors associated with the characteristic values 1, 2. We know that dim $W_1 = 1$ and dim $W_2 = 2$. By Theorem 2, T is diagonalizable. It is easy to exhibit a basis for R^3 in which T is represented by a diagonal matrix. The null space of $(T - I)$ is spanned

by the vector $\alpha_1 = (3, -1, 3)$ and so $\{\alpha_1\}$ is a basis for W_1. The null space of $T - 2I$ (i.e., the space W_2) consists of the vectors (x_1, x_2, x_3) with $x_1 = 2x_2 + 2x_3$. Thus, one example of a basis for W_2 is

$$\alpha_2 = (2, 1, 0)$$
$$\alpha_3 = (2, 0, 1).$$

If $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$, then $[T]_{\mathcal{B}}$ is the diagonal matrix

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The fact that T is diagonalizable means that the original matrix A is similar (over R) to the diagonal matrix D. The matrix P which enables us to change coordinates from the basis $\mathcal{B}$ to the standard basis is (of course) the matrix which has the transposes of α_1, α_2, α_3 as its column vectors:

$$P = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

Furthermore, $AP = PD$, so that

$$P^{-1}AP = D.$$

Definition. *Let* T *be a linear operator on a finite-dimensional vector space* V *over the field* F. *The* **minimal polynomial** *for* T *is the (unique) monic generator of the ideal of polynomials over* F *which annihilate* T.

The name 'minimal polynomial' stems from the fact that the generator of a polynomial ideal is characterized by being the monic polynomial of minimum degree in the ideal. That means that the minimal polynomial p for the linear operator T is uniquely determined by these three properties:

(1) p is a monic polynomial over the scalar field F.
(2) $p(T) = 0$.
(3) No polynomial over F which annihilates T has smaller degree than p has.

If A is an $n \times n$ matrix over F, we define the **minimal polynomial** for A in an analogous way, as the unique monic generator of the ideal of all polynomials over F which annihilate A. If the operator T is represented in

some ordered basis by the matrix A, then T and A have the same minimal polynomial. That is because $f(T)$ is represented in the basis by the matrix $f(A)$, so that $f(T) = 0$ if and only if $f(A) = 0$.

From the last remark about operators and matrices it follows that similar matrices have the same minimal polynomial. That fact is also clear from the definitions because

$$f(P^{-1}AP) = P^{-1}f(A)P$$

for every polynomial f.

Theorem *Let* T *be a linear operator on an* n*-dimensional vector space* V [*or, let* A *be an* n × n *matrix*]. *The characteristic and minimal polynomials for* T [*for* A] *have the same roots, except for multiplicities.*

Proof. Let p be the minimal polynomial for T. Let c be a scalar. What we want to show is that $p(c) = 0$ if and only if c is a characteristic value of T.

First, suppose $p(c) = 0$. Then

$$p = (x - c)q$$

where q is a polynomial. Since deg $q <$ deg p, the definition of the minimal polynomial p tells us that $q(T) \neq 0$. Choose a vector β such that $q(T)\beta \neq 0$. Let $\alpha = q(T)\beta$. Then

$$\begin{aligned} 0 &= p(T)\beta \\ &= (T - cI)q(T)\beta \\ &= (T - cI)\alpha \end{aligned}$$

and thus, c is a characteristic value of T.

Now, suppose that c is a characteristic value of T, say, $T\alpha = c\alpha$ with $\alpha \neq 0$. As we noted in a previous lemma,

$$p(T)\alpha = p(c)\alpha.$$

Since $p(T) = 0$ and $\alpha \neq 0$, we have $p(c) = 0$. ∎

Let T be a diagonalizable linear operator and let $c_1, \ldots, c_k$ be the distinct characteristic values of T. Then it is easy to see that the minimal polynomial for T is the polynomial

$$p = (x - c_1) \cdots (x - c_k).$$

If α is a characteristic vector, then one of the operators $T - c_1 I, \ldots, T - c_k I$ sends α into 0. Therefore

$$(T - c_1 I) \cdots (T - c_k I)\alpha = 0$$

for every characteristic vector α. There is a basis for the underlying space which consists of characteristic vectors of T; hence

$$p(T) = (T - c_1 I) \cdots (T - c_k I) = 0.$$

What we have concluded is this. If T is a diagonalizable linear operator, then the minimal polynomial for T is a product of distinct linear factors. As we shall soon see, that property characterizes diagonalizable operators.

Theorem 4 (Cayley-Hamilton). *Let* T *be a linear operator on a finite dimensional vector space* V. *If* f *is the characteristic polynomial for* T, *then* f(T) = 0; *in other words, the minimal polynomial divides the characteristic polynomial for* T.

Proof. Later on we shall give two proofs of this result independent of the one to be given here. The present proof, although short, may be difficult to understand. Aside from brevity, it has the virtue of providing

an illuminating and far from trivial application of the general theory of determinants developed in Chapter 5.

Let K be the commutative ring with identity consisting of all polynomials in T. Of course, K is actually a commutative algebra with identity over the scalar field. Choose an ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ for V, and let A be the matrix which represents T in the given basis. Then

$$T\alpha_i = \sum_{j=1}^{n} A_{ji}\alpha_j, \qquad 1 \leq i \leq n.$$

These equations may be written in the equivalent form

$$\sum_{j=1}^{n} (\delta_{ij}T - A_{ji}I)\alpha_j = 0, \qquad 1 \leq i \leq n.$$

Let B denote the element of $K^{n\times n}$ with entries

$$B_{ij} = \delta_{ij}T - A_{ji}I.$$

When $n = 2$

$$B = \begin{bmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{bmatrix}$$

and

$$\begin{aligned} \det B &= (T - A_{11}I)(T - A_{22}I) - A_{12}A_{21}I \\ &= T^2 - (A_{11} + A_{22})T + (A_{11}A_{22} - A_{12}A_{21})I \\ &= f(T) \end{aligned}$$

where f is the characteristic polynomial:

$$f = x^2 - (\text{trace } A)x + \det A.$$

For the case $n > 2$, it is also clear that

$$\det B = f(T)$$

since f is the determinant of the matrix $xI - A$ whose entries are the polynomials

$$(xI - A)_{ij} = \delta_{ij}x - A_{ji}.$$

We wish to show that $f(T) = 0$. In order that $f(T)$ be the zero operator, it is necessary and sufficient that $(\det B)\alpha_k = 0$ for $k = 1, \ldots, n$. By the definition of B, the vectors $\alpha_1, \ldots, \alpha_n$ satisfy the equations

$$\sum_{j=1}^{n} B_{ij}\alpha_j = 0, \qquad 1 \leq i \leq n. \tag{6-6}$$

When $n = 2$, it is suggestive to write (6-6) in the form

$$\begin{bmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In this case, the classical adjoint, adj B is the matrix

$$\tilde{B} = \begin{bmatrix} T - A_{22}I & A_{21}I \\ A_{12}I & T - A_{11}I \end{bmatrix}$$

and

$$\tilde{B}B = \begin{bmatrix} \det B & 0 \\ 0 & \det B \end{bmatrix}.$$

Hence, we have

$$\begin{aligned} (\det B) \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} &= (\tilde{B}B) \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \\ &= \tilde{B} \left(B \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \right) \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

In the general case, let $\tilde{B} = \text{adj } B$. Then by (6-6)

$$\sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j = 0$$

for each pair k, i, and summing on i, we have

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \tilde{B}_{ki}B_{ij} \right) \alpha_j. \end{aligned}$$

Now $\tilde{B}B = (\det B)I$, so that

$$\sum_{i=1}^{n} \tilde{B}_{ki}B_{ij} = \delta_{kj} \det B.$$

Therefore

$$0 = \sum_{j=1}^{n} \delta_{kj}(\det B)\alpha_j$$
$$= (\det B)\alpha_k, \qquad 1 \leq k \leq n. \quad \blacksquare$$

The Cayley-Hamilton theorem is useful to us at this point primarily because it narrows down the search for the minimal polynomials of various operators. If we know the matrix A which represents T in some ordered basis, then we can compute the characteristic polynomial f. We know that the minimal polynomial p divides f and that the two polynomials have the same roots. There is no method for computing precisely the roots of a polynomial (unless its degree is small); however, if f factors

(6-7) $$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}, \qquad c_1, \ldots, c_k \text{ distinct}, d_i \geq 1$$

then

(6-8) $$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}, \qquad 1 \leq r_j \leq d_j.$$

That is all we can say in general. If f is the polynomial (6-7) and has degree n, then for every polynomial p as in (6-8) we can find an $n \times n$ matrix which has f as its characteristic polynomial and p as its minimal

Invariant Subspaces

Definition. *Let* V *be a vector space and* T *a linear operator on* V. *If* W *is a subspace of* V, *we say that* W *is* **invariant under** T *if for each vector* α *in* W *the vector* Tα *is in* W, *i.e., if* T(W) *is contained in* W.

EXAMPLE 6. If T is any linear operator on V, then V is invariant under T, as is the zero subspace. The range of T and the null space of T are also invariant under T.

EXAMPLE 7. Let F be a field and let D be the differentiation operator on the space $F[x]$ of polynomials over F. Let n be a positive integer and let W be the subspace of polynomials of degree not greater than n. Then W is invariant under D. This is just another way of saying that D is 'degree decreasing.'

space V. Let W be the subspace spanned by *all* of the characteristic vectors

EXAMPLE Let T be any linear operator on a finite-dimensional

space V. Let W be the subspace spanned by *all* of the characteristic vectors of T. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T. For each i, let W_i be the space of characteristic vectors associated with the characteristic value c_i, and let $\mathcal{B}_i$ be an ordered basis for W_i. The lemma before Theorem 2 tells us that $\mathcal{B}' = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is an ordered basis for W. In particular,

$$\dim W = \dim W_1 + \cdots + \dim W_k.$$

Let $\mathcal{B}' = \{\alpha_1, \ldots, \alpha_r\}$ so that the first few α's form the basis $\mathcal{B}_1$, the next few $\mathcal{B}_2$, and so on. Then

$$T\alpha_i = t_i\alpha_i, \qquad i = 1, \ldots, r$$

where $(t_1, \ldots, t_r) = (c_1, c_1, \ldots, c_1, \ldots, c_k, c_k, \ldots, c_k)$ with c_i repeated $\dim W_i$ times.

Now W is invariant under T, since for each α in W we have

$$\alpha = x_1\alpha_1 + \cdots + x_r\alpha_r$$
$$T\alpha = t_1x_1\alpha_1 + \cdots + t_rx_r\alpha_r.$$

Choose any other vectors $\alpha_{r+1}, \ldots, \alpha_n$ in V such that $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. The matrix of T relative to $\mathcal{B}$ has the block form (6-10), and the matrix of the restriction operator T_W relative to the basis $\mathcal{B}'$ is

$$B = \begin{bmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & t_r \end{bmatrix}.$$

The characteristic polynomial of B (i.e., of T_W) is

$$g = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$$

where $e_i = \dim W_i$. Furthermore, g divides f, the characteristic polynomial for T. Therefore, the multiplicity of c_i as a root of f is at least $\dim W_i$.

All of this should make Theorem 2 transparent. It merely says that T is diagonalizable if and only if $r = n$, if and only if $e_1 + \cdots + e_k = n$. It does not help us too much with the non-diagonalizable case, since we don't know the matrices C and D of (6-10).

Lemma. *Let* V *be a finite-dimensional vector space over the field* F. *Let* T *be a linear operator on* V *such that the minimal polynomial for* T *is a product of linear factors*

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}, \qquad c_i \text{ in } F.$$

Let W *be a proper* ($W \neq V$) *subspace of* V *which is invariant under* T. *There exists a vector* α *in* V *such that*

(a) α *is not in* W;
(b) $(T - cI)\alpha$ *is in* W, *for some characteristic value* c *of the operator* T.

Proof. What (a) and (b) say is that the T-conductor of α into W is a linear polynomial. Let β be any vector in V which is not in W. Let g be the T-conductor of β into W. Then g divides p, the minimal polynomial for T. Since β is not in W, the polynomial g is not constant. Therefore,

$$g = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$$

where at least one of the integers e_i is positive. Choose j so that $e_j > 0$. Then $(x - c_j)$ divides g:

$$g = (x - c_j)h.$$

By the definition of g, the vector $\alpha = h(T)\beta$ cannot be in W. But

$$\begin{aligned}(T - c_jI)\alpha &= (T - c_jI)h(T)\beta \\ &= g(T)\beta\end{aligned}$$

is in W. ∎

Theorem *Let* V *be a finite-dimensional vector space over the field* F

and let T *be a linear operator on* V. *Then* T *is triangulable if and only if the minimal polynomial for* T *is a product of linear polynomials over* F.

Proof. Suppose that the minimal polynomial factors

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}.$$

By repeated application of the lemma above, we shall arrive at an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ in which the matrix representing T is upper-triangular:

$$[T]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}. \tag{6-11}$$

Now (6-11) merely says that

$$T\alpha_j = a_{1j}\alpha_1 + \cdots + a_{jj}\alpha_j, \qquad 1 \leq j \leq n \tag{6-12}$$

that is, $T\alpha_j$ is in the subspace spanned by $\alpha_1, \ldots, \alpha_j$. To find $\alpha_1, \ldots, \alpha_n$, we start by applying the lemma to the subspace $W = \{0\}$, to obtain the vector α_1. Then apply the lemma to W_1, the space spanned by α_1, and we obtain α_2. Next apply the lemma to W_2, the space spanned by α_1 and α_2. Continue in that way. One point deserves comment. After $\alpha_1, \ldots, \alpha_i$ have been found, it is the triangular-type relations (6-12) for $j = 1, \ldots, i$ which ensure that the subspace spanned by $\alpha_1, \ldots, \alpha_i$ is invariant under T.

If T is triangulable, it is evident that the characteristic polynomial for T has the form

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}, \qquad c_i \text{ in } F.$$

Just look at the triangular matrix (6-11). The diagonal entries $a_{11}, \ldots, a_{1n}$ are the characteristic values, with c_i repeated d_i times. But, if f can be so factored, so can the minimal polynomial p, because it divides f. ∎

Corollary. *Let* F *be an algebraically closed field, e.g., the complex number field. Every* n × n *matrix over* F *is similar over* F *to a triangular matrix.*

Theorem *Let* V *be a finite-dimensional vector space over the field* F

and let T *be a linear operator on* V. *Then* T *is diagonalizable if and only if the minimal polynomial for* T *has the form*

$$p = (x - c_1) \cdots (x - c_k)$$

where $c_1, \ldots, c_k$ *are distinct elements of* F.

Proof. We have noted earlier that, if T is diagonalizable, its minimal polynomial is a product of distinct linear factors (see the discussion prior to Example 4). To prove the converse, let W be the subspace spanned by all of the characteristic vectors of T, and suppose $W \neq V$. By the lemma used in the proof of Theorem 5, there is a vector α not in W and a characteristic value c_j of T such that the vector

$$\beta = (T - c_j I)\alpha$$

lies in W. Since β is in W,

$$\beta = \beta_1 + \cdots + \beta_k$$

where $T\beta_i = c_i\beta_i$, $1 \leq i \leq k$, and therefore the vector

$$h(T)\beta = h(c_1)\beta_1 + \cdots + h(c_k)\beta_k$$

is in W, for every polynomial h.

Now $p = (x - c_j)q$, for some polynomial q. Also

$$q - q(c_j) = (x - c_j)h.$$

We have

$$q(T)\alpha - q(c_j)\alpha = h(T)(T - c_j I)\alpha = h(T)\beta$$

But $h(T)\beta$ is in W and, since

$$0 = p(T)\alpha = (T - c_j I)q(T)\alpha$$

the vector $q(T)\alpha$ is in W. Therefore, $q(c_j)\alpha$ is in W. Since α is not in W, we have $q(c_j) = 0$. That contradicts the fact that p has distinct roots. ∎

POSSIBLE QUESTIONS

2 Mark Questions:

1. Define dual basis.
2. What is minimal polynomial?
3. State Cayley Hamiltion theorem.
4. What is annihilator?
5. Defie dual space.

6 Mark Questions:

6. Let V be a finite dimensional vector space over the field F, and let W be a subspace of V, then prove that dim W + dim W^0 = dim V.
7. Let V be a finite dimensional vector space over the field F. Each basis for V^* is a dual of some basis for V.
8. Let $f_1, f_2, ..., f_r$ be a linear functional on a vector space V with respective null spaces $N_1, N_2, ..., N_r$ then prove that g is a linear combination of $f_1, f_2, ..., f_r$ if and only if N contains the intersection $N_1, N_2, ..., N_r$.
9. Let T be a linear operator on the finite dimensional space V.let $c_1, c_2, \ldots, c_k$ be the distinct characteristic value of T and let W_i be the space of characteristic vectors associate with characteristic values c_i. Prove that if $W = W_1 + W_2 + \ldots + W_k$ then dim W = dim W_1 + dimW_2 + ….+ dimW_k infact if B_1 is an order basis for W_1 then $B = (B_1, B_2, \ldots, B_k)$ is an ordered basis for W.
10. Let V and W be finite dimensional vector space over the field F. Let B be the ordered basis for V with dual basis B^* and let B' be an ordered basis for W with dual basis B'^*. Let T be the linear transformation form V into W. Let A be the matrix of T relative to B, B' and let B be the matrix of T^k relative to B'*, B* then prove that $B_{ij} = A_{ij}$.
11. L:et T be a linear operator on an n-dimensional vector space V. prove that the characteristic and minimal polynomial for T have the same roots, except the multiplicities.

12. Let T be a linear operator on a finite dimensional space V. Let $c_1, c_2, \ldots, c_k$ be the distinct characteristic values of T and let W_i be the null space of $(T-c_i)$. then prove the following are equivalent.

    i. T is diagonalizable.

    ii. The characteristic polynomial for T is $f = (x-c_1)^{d1} \ldots\ldots (x-c_k)^{dk}$

    and dim $W_i = d_i$, $i = 1, 2, \ldots,. k$

    iii. Dim W_1 + dim W_2 + …..+ dim W_k = dim V

13. Let W be an invariant subspace for T. the characteristic polynomialfor the restriction operator T_w divides the characteristic polynomial for T.Then prove that the minimal polynomial of T_w divides the minimal polynomial for T.

14. Diagonalize the matrix $\begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$

15. State and prove Cayley Hamilton theorem.

UNIT-III

SYLLABUS

Dual spaces, dual basis, double dual, transpose of a linear transformation and its matrix in the dual basis, annihilators, Eigen spaces of a linear operator, diagonalizability, invariant subspaces and Cayley-Hamilton theorem, the minimal polynomial for a linear operator.

KAHE

Linear Functionals

If V is a vector space over the field F, a linear transformation f from V into the scalar field F is also called a **linear functional** on V. If we start from scratch, this means that f is a function from V into F such that

$$f(c\alpha + \beta) = cf(\alpha) + f(\beta)$$

for all vectors α and β in V and all scalars c in F. The concept of linear functional is important in the study of finite-dimensional spaces because it helps to organize and clarify the discussion of subspaces, linear equations, and coordinates.

EXAMPLE

Let F be a field and let $a_1, \ldots, a_n$ be scalars in F. Define a function f on F^n by

$$f(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n.$$

Then f is a linear functional on F^n. It is the linear functional which is represented by the matrix $[a_1 \cdots a_n]$ relative to the standard ordered basis for F^n and the basis $\{1\}$ for F:

$$a_j = f(\epsilon_j), \qquad j = 1, \ldots, n.$$

Every linear functional on F^n is of this form, for some scalars $a_1, \ldots, a_n$. That is immediate from the definition of linear functional because we define $a_j = f(\epsilon_j)$ and use the linearity

$$\begin{aligned} f(x_1, \ldots, x_n) &= f\left(\sum_j x_j\epsilon_j\right) \\ &= \sum_j x_j f(\epsilon_j) \\ &= \sum_j a_j x_j. \end{aligned}$$

EXAMPLE

Here is an important example of a linear functional.

Let n be a positive integer and F a field. If A is an $n \times n$ matrix with entries in F, the **trace** of A is the scalar

$$\operatorname{tr} A = A_{11} + A_{22} + \cdots + A_{nn}.$$

The trace function is a linear functional on the matrix space $F^{n\times n}$ because

$$\operatorname{tr}(cA + B) = \sum_{i=1}^{n} (cA_{ii} + B_{ii})$$

$$= c \sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii}$$

$$= c \operatorname{tr} A + \operatorname{tr} B.$$

Theorem

Let V *be a finite-dimensional vector space over the field* F,

and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *be a basis for* V. *Then there is a unique dual basis* $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ *for* V^* *such that* $f_i(\alpha_j) = \delta_{ij}$. *For each linear functional* f *on* V *we have*

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i \tag{3-13}$$

and for each vector α *in* V *we have*

$$\alpha = \sum_{i=1}^{n} f_i(\alpha) \alpha_i. \tag{3-14}$$

Proof. We have shown above that there is a unique basis which is 'dual' to $\mathcal{B}$. If f is a linear functional on V, then f is some linear combination (3-12) of the f_i, and as we observed after (3-12) the scalars c_j must be given by $c_j = f(\alpha_j)$. Similarly, if

$$\alpha = \sum_{i=1}^{n} x_i \alpha_i$$

is a vector in V, then

$$f_j(\alpha) = \sum^{n} x_i f_j(\alpha_i)$$

$$= \sum_{i=1}^{n} x_i \delta_{ij}$$

$$= x_j$$

so that the unique expression for α as a linear combination of the α_i is

$$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i. \quad \blacksquare$$

Equation (3-14) provides us with a nice way of describing what the dual basis is. It says, if $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V and $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ is the dual basis, then f_i is precisely the function which assigns to each vector α in V the ith coordinate of α relative to the ordered basis $\mathcal{B}$. Thus we may also call the f_i the coordinate functions for $\mathcal{B}$. The formula (3-13), when combined with (3-14) tells us the following: If f is in V^*, and we let $f(\alpha_i) = a_i$, then when

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$$

we have

$$f(\alpha) = a_1x_1 + \cdots + a_nx_n. \tag{3-15}$$

In other words, if we choose an ordered basis $\mathcal{B}$ for V and describe each vector in V by its n-tuple of coordinates $(x_1, \ldots, x_n)$ relative to $\mathcal{B}$, then every linear functional on V has the form (3-15). This is the natural generalization of Example 18, which is the special case $V = F^n$ and $\mathcal{B} = \{\epsilon_1, \ldots, \epsilon_n\}$.

EXAMPLE 22. Let V be the vector space of all polynomial functions from R into R which have degree less than or equal to 2. Let t_1, t_2, and t_3 be any three *distinct* real numbers, and let

$$L_i(p) = p(t_i).$$

Then L_1, L_2, and L_3 are linear functionals on V. These functionals are linearly independent; for, suppose

$$L = c_1L_1 + c_2L_2 + c_3L_3.$$

If $L = 0$, i.e., if $L(p) = 0$ for each p in V, then applying L to the particular polynomial 'functions' 1, x, x^2, we obtain

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ t_1c_1 + t_2c_2 + t_3c_3 &= 0 \\ t_1^2c_1 + t_2^2c_2 + t_3^2c_3 &= 0 \end{aligned}$$

From this it follows that $c_1 = c_2 = c_3 = 0$, because (as a short computation shows) the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix}$$

is invertible when t_1, t_2, and t_3 are distinct. Now the L_i are independent, and since V has dimension 3, these functionals form a basis for V^*. What is the basis for V, of which this is the dual? Such a basis $\{p_1, p_2, p_3\}$ for V must satisfy

$$L_i(p_j) = \delta_{ij}$$

or

$$p_j(t_i) = \delta_{ij}.$$

These polynomial functions are rather easily seen to be

$$\begin{aligned} p_1(x) &= \frac{(x - t_2)(x - t_3)}{(t_1 - t_2)(t_1 - t_3)} \\ p_2(x) &= \frac{(x - t_1)(x - t_3)}{(t_2 - t_1)(t_2 - t_3)} \\ p_3(x) &= \frac{(x - t_1)(x - t_2)}{(t_3 - t_1)(t_3 - t_2)}. \end{aligned}$$

The basis $\{p_1, p_2, p_3\}$ for V is interesting, because according to (3-14) we have for each p in V

$$p = p(t_1)p_1 + p(t_2)p_2 + p(t_3)p_3.$$

Thus, if c_1, c_2, and c_3 are any real numbers, there is exactly one polynomial function p over R which has degree at most 2 and satisfies $p(t_j) = c_j$, $j = 1, 2, 3$. This polynomial function is $p = c_1p_1 + c_2p_2 + c_3p_3$.

Now let us discuss the relationship between linear functionals and subspaces. If f is a non-zero linear functional, then the rank of f is 1 because the range of f is a non-zero subspace of the scalar field and must (therefore) be the scalar field. If the underlying space V is finite-dimensional, the rank plus nullity theorem (Theorem 2) tells us that the null space N_f has dimension

$$\dim N_f = \dim V - 1.$$

In a vector space of dimension n, a subspace of dimension $n - 1$ is called a **hyperspace.** Such spaces are sometimes called hyperplanes or subspaces of codimension 1. Is every hyperspace the null space of a linear functional? The answer is easily seen to be yes. It is not much more difficult to show that each d-dimensional subspace of an n-dimensional space is the intersection of the null spaces of $(n - d)$ linear functionals (Theorem 16 below).

Definition. *If* V *is a vector space over the field* F *and* S *is a subset of* V, *the* **annihilator** *of* S *is the set* S^0 *of linear functionals* f *on* V *such that* $f(\alpha) = 0$ *for every* α *in* S.

It should be clear to the reader that S^0 is a subspace of V^*, whether S is a subspace of V or not. If S is the set consisting of the zero vector alone, then $S^0 = V^*$. If $S = V$, then S^0 is the zero subspace of V^*. (This is easy to see when V is finite-dimensional.)

Theorem 16. *Let* V *be a finite-dimensional vector space over the field* F, *and let* W *be a subspace of* V. *Then*

$$dim\ W + dim\ W^0 = dim\ V.$$

Proof. Let k be the dimension of W and $\{\alpha_1, \ldots, \alpha_k\}$ a basis for W. Choose vectors $\alpha_{k+1}, \ldots, \alpha_n$ in V such that $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. Let $\{f_1, \ldots, f_n\}$ be the basis for V^* which is dual to this basis for V.

The claim is that $\{f_{k+1}, \ldots, f_n\}$ is a basis for the annihilator W^0. Certainly f_i belongs to W^0 for $i \geq k+1$, because

$$f_i(\alpha_j) = \delta_{ij}$$

and $\delta_{ij} = 0$ if $i \geq k+1$ and $j \leq k$; from this it follows that, for $i \geq k+1$, $f_i(\alpha) = 0$ whenever α is a linear combination of $\alpha_1, \ldots, \alpha_k$. The functionals $f_{k+1}, \ldots, f_n$ are independent, so all we must show is that they span W^0. Suppose f is in V^*. Now

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i$$

so that if f is in W^0 we have $f(\alpha_i) = 0$ for $i \leq k$ and

$$f = \sum_{i=k+1}^{n} f(\alpha_i) f_i.$$

We have shown that if dim $W = k$ and dim $V = n$ then dim $W^0 = n - k$. ∎

Corollary. *If* W *is a* k*-dimensional subspace of an* n*-dimensional vector space* V*, then* W *is the intersection of* (n − k) *hyperspaces in* V.

Proof. This is a corollary of the proof of Theorem 16 rather than its statement. In the notation of the proof, W is exactly the set of vectors α such that $f_i(\alpha) = 0$, $i = k+1, \ldots, n$. In case $k = n-1$, W is the null space of f_n. ∎

$$\vdots \qquad\qquad \vdots$$
$$A_{m1}x_1 + \cdots + A_{mn}x_n = 0$$

for which we wish to find the solutions. If we let f_i, $i = 1, \ldots, m$, be the linear functional on F^n defined by

$$f_i(x_1, \ldots, x_n) = A_{i1}x_1 + \cdots + A_{in}x_n$$

then we are seeking the subspace of F^n of all α such that

$$f_i(\alpha) = 0, \qquad i = 1, \ldots, m.$$

In other words, we are seeking the subspace annihilated by $f_1, \ldots, f_m$. Row-reduction of the coefficient matrix provides us with a systematic

EXAMPLE Here are three linear functionals on R^4:

$$f_1(x_1, x_2, x_3, x_4) = x_1 + 2x_2 + 2x_3 + x_4$$
$$f_2(x_1, x_2, x_3, x_4) = 2x_2 + x_4$$
$$f_3(x_1, x_2, x_3, x_4) = -2x_1 - 4x_3 + 3x_4.$$

The subspace which they annihilate may be found explicitly by finding the row-reduced echelon form of the matrix

$$A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 2 & 0 & 1 \\ -2 & 0 & -4 & 3 \end{bmatrix}.$$

A short calculation, or a peek at Example 21 of Chapter 2, shows that

$$R = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Therefore, the linear functionals

$$g_1(x_1, x_2, x_3, x_4) = x_1 + 2x_3$$
$$g_2(x_1, x_2, x_3, x_4) = x_2$$
$$g_3(x_1, x_2, x_3, x_4) = x_4$$

span the same subspace of $(R^4)^*$ and annihilate the same subspace of R^4 as do f_1, f_2, f_3. The subspace annihilated consists of the vectors with

$$x_1 = -2x_3$$
$$x_2 = x_4 = 0.$$

EXAMPLE Let W be the subspace of R^5 which is spanned by the

vectors

$$\alpha_1 = (2, -2, 3, 4, -1), \qquad \alpha_3 = (0, 0, -1, -2, 3)$$
$$\alpha_2 = (-1, 1, 2, 5, 2), \qquad \alpha_4 = (1, -1, 2, 3, 0).$$

How does one describe W^0, the annihilator of W? Let us form the 4×5 matrix A with row vectors $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, and find the row-reduced echelon matrix R which is row-equivalent to A:

$$A = \begin{bmatrix} 2 & -2 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 2 \\ 0 & 0 & -1 & -2 & 3 \\ 1 & -1 & 2 & 3 & 0 \end{bmatrix} \longrightarrow R = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

If f is a linear functional on R^5:

$$f(x_1, \ldots, x_5) = \sum_{j=1}^{5} c_j x_j$$

then f is in W^0 if and only if $f(\alpha_i) = 0$, $i = 1, 2, 3, 4$, i.e., if and only if

$$\sum_{j=1}^{5} A_{ij} c_j = 0, \qquad 1 \leq i \leq 4.$$

This is equivalent to

$$\sum_{j=1}^{5} R_{ij} c_j = 0, \qquad 1 \leq i \leq 3$$

or

$$c_1 - c_2 - c_4 = 0$$
$$c_3 + 2c_4 = 0$$
$$c_5 = 0.$$

We obtain all such linear functionals f by assigning arbitrary values to c_2 and c_4, say $c_2 = a$ and $c_4 = b$, and then finding the corresponding $c_1 = a + b$, $c_3 = -2b$, $c_5 = 0$. So W^0 consists of all linear functionals f of the form

$$f(x_1, x_2, x_3, x_4, x_5) = (a + b)x_1 + ax_2 - 2bx_3 + bx_4.$$

The dimension of W^0 is 2 and a basis $\{f_1, f_2\}$ for W^0 can be found by first taking $a = 1$, $b = 0$ and then $a = 0$, $b = 1$:

$$f_1(x_1, \ldots, x_5) = x_1 + x_2$$
$$f_2(x_1, \ldots, x_5) = x_1 - 2x_3 + x_4.$$

The above general f in W^0 is $f = af_1 + bf_2$.

The Double Dual

One question about dual bases which we did not answer in the last section was whether every basis for V^* is the dual of some basis for V. One way to answer that question is to consider V^{**}, the dual space of V^*.

If α is a vector in V, then α induces a linear functional L_α on V^* defined by

$$L_\alpha(f) = f(\alpha), \qquad f \text{ in } V^*.$$

The fact that L_α is linear is just a reformulation of the definition of linear operations in V^*:

$$\begin{aligned} L_\alpha(cf + g) &= (cf + g)(\alpha) \\ &= (cf)(\alpha) + g(\alpha) \\ &= cf(\alpha) + g(\alpha) \\ &= cL_\alpha(f) + L_\alpha(g). \end{aligned}$$

If V is finite-dimensional and $\alpha \neq 0$, then $L_\alpha \neq 0$; in other words, there exists a linear functional f such that $f(\alpha) \neq 0$. The proof is very simple and was given in Section 3.5: Choose an ordered basis $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ for V such that $\alpha_1 = \alpha$ and let f be the linear functional which assigns to each vector in V its first coordinate in the ordered basis $\mathfrak{B}$.

Theorem

Let V *be a finite-dimensional vector space over the field* F.

For each vector α *in* V *define*

$$L_\alpha(f) = f(\alpha), \qquad f \text{ in } V^*.$$

The mapping $\alpha \to L_\alpha$ *is then an isomorphism of* V *onto* V^{**}.

Proof. We showed that for each α the function L_α is linear. Suppose α and β are in V and c is in F, and let $\gamma = c\alpha + \beta$. Then for each f in V^*

$$\begin{aligned} L_\gamma(f) &= f(\gamma) \\ &= f(c\alpha + \beta) \\ &= cf(\alpha) + f(\beta) \\ &= cL_\alpha(f) + L_\beta(f) \end{aligned}$$

and so

$$L_\gamma = cL_\alpha + L_\beta.$$

This shows that the mapping $\alpha \to L_\alpha$ is a linear transformation from V into V^{**}. This transformation is non-singular; for, according to the remarks above $L_\alpha = 0$ if and only if $\alpha = 0$. Now $\alpha \to L_\alpha$ is a non-singular linear transformation from V into V^{**}, and since

$$\dim V^{**} = \dim V^* = \dim V$$

Theorem 9 tells us that this transformation is invertible, and is therefore an isomorphism of V onto V^{**}. ▌

Corollary. *Let* V *be a finite-dimensional vector space over the field* F. *If* L *is a linear functional on the dual space* V* *of* V, *then there is a unique vector* α *in* V *such that*

$$L(f) = f(\alpha)$$

for every f *in* V*.

Corollary. *Let* V *be a finite-dimensional vector space over the field* F. *Each basis for* V* *is the dual of some basis for* V.

Proof. Let $\mathfrak{B}^* = \{f_1, \ldots, f_n\}$ be a basis for V^*. By Theorem 15, there is a basis $\{L_1, \ldots, L_n\}$ for V^{**} such that

$$L_i(f_j) = \delta_{ij}.$$

Using the corollary above, for each i there is a vector α_i in V such that

$$L_i(f) = f(\alpha_i)$$

for every f in V^*, i.e., such that $L_i = L_{\alpha_i}$. It follows immediately that $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V and that $\mathfrak{B}^*$ is the dual of this basis. ▌

Theorem

If S *is any subset of a finite-dimensional vector space* V, *then* $(S^0)^0$ *is the subspace spanned by* S.

Proof. Let W be the subspace spanned by S. Clearly $W^0 = S^0$. Therefore, what we are to prove is that $W = W^{00}$. We have given one proof. Here is another. By Theorem 16

$$\dim W + \dim W^0 = \dim V$$
$$\dim W^0 + \dim W^{00} = \dim V^*$$

and since $\dim V = \dim V^*$ we have

$$\dim W = \dim W^{00}.$$

Since W is a subspace of W^{00}, we see that $W = W^{00}$. ∎

The results of this section hold for arbitrary vector spaces; however, the proofs require the use of the so-called Axiom of Choice. We want to avoid becoming embroiled in a lengthy discussion of that axiom, so we shall not tackle annihilators for general vector spaces. But, there are two results about linear functionals on arbitrary vector spaces which are so fundamental that we should include them.

Definition. *If* V *is a vector space, a* **hyperspace** *in* V *is a maximal proper subspace of* V.

Theorem

If f *is a non-zero linear functional on the vector space* V, *then the null space of* f *is a hyperspace in* V. *Conversely, every hyperspace in* V *is the null space of a (not unique) non-zero linear functional on* V.

Proof. Let f be a non-zero linear functional on V and N_f its null space. Let α be a vector in V which is not in N_f, i.e., a vector such that $f(\alpha) \neq 0$. We shall show that every vector in V is in the subspace spanned by N_f and α. That subspace consists of all vectors

$$\gamma + c\alpha, \qquad \gamma \text{ in } N_f, \; c \text{ in } F.$$

Let β be in V. Define

$$c = \frac{f(\beta)}{f(\alpha)}$$

which makes sense because $f(\alpha) \neq 0$. Then the vector $\gamma = \beta - c\alpha$ is in N_f since

$$\begin{aligned} f(\gamma) &= f(\beta - c\alpha) \\ &= f(\beta) - cf(\alpha) \\ &= 0. \end{aligned}$$

So β is in the subspace spanned by N_f and α.

Now let N be a hyperspace in V. Fix some vector α which is not in N. Since N is a maximal proper subspace, the subspace spanned by N and α is the entire space V. Therefore each vector β in V has the form

$$\beta = \gamma + c\alpha, \qquad \gamma \text{ in } N, c \text{ in } F.$$

The vector γ and the scalar c are uniquely determined by β. If we have also

$$\beta = \gamma' + c'\alpha, \qquad \gamma' \text{ in } N, c' \text{ in } F.$$

then

$$(c' - c)\alpha = \gamma - \gamma'.$$

If $c' - c \neq 0$, then α would be in N; hence, $c' = c$ and $\gamma' = \gamma$. Another way to phrase our conclusion is this: If β is in V, there is a unique scalar c such that $\beta - c\alpha$ is in N. Call that scalar $g(\beta)$. It is easy to see that g is a linear functional on V and that N is the null space of g. ∎

Lemma. *If* f *and* g *are linear functionals on a vector space* V, *then* g *is a scalar multiple of* f *if and only if the null space of* g *contains the null space of* f, *that is, if and only if* $f(\alpha) = 0$ *implies* $g(\alpha) = 0$.

Proof. If $f = 0$ then $g = 0$ as well and g is trivially a scalar multiple of f. Suppose $f \neq 0$ so that the null space N_f is a hyperspace in V. Choose some vector α in V with $f(\alpha) \neq 0$ and let

$$c = \frac{g(\alpha)}{f(\alpha)}.$$

The linear functional $h = g - cf$ is 0 on N_f, since both f and g are 0 there, and $h(\alpha) = g(\alpha) - cf(\alpha) = 0$. Thus h is 0 on the subspace spanned by N_f and α—and that subspace is V. We conclude that $h = 0$, i.e., that $g = cf$. ∎

Theorem *Let* $g, f_1, \ldots, f_r$ *be linear functionals on a vector space* V *with respective null spaces* $N, N_1, \ldots, N_r$. *Then* g *is a linear combination of* $f_1, \ldots, f_r$ *if and only if* N *contains the intersection* $N_1 \cap \cdots \cap N_r$.

Proof. If $g = c_1f_1 + \cdots + c_rf_r$ and $f_i(\alpha) = 0$ for each i, then clearly $g(\alpha) = 0$. Therefore, N contains $N_1 \cap \cdots \cap N_r$.

We shall prove the converse (the 'if' half of the theorem) by induction on the number r. The preceding lemma handles the case $r = 1$. Suppose we know the result for $r = k - 1$, and let $f_1, \ldots, f_k$ be linear functionals with null spaces $N_1, \ldots, N_k$ such that $N_1 \cap \cdots \cap N_k$ is contained in N, the null space of g. Let $g', f_1', \ldots, f_{k-1}'$ be the restrictions of $g, f_1, \ldots, f_{k-1}$ to the subspace N_k. Then $g', f_1', \ldots, f_{k-1}'$ are linear functionals on the vector space N_k. Furthermore, if α is a vector in N_k and $f_i'(\alpha) = 0$, $i = 1, \ldots, k - 1$, then α is in $N_1 \cap \cdots \cap N_k$ and so $g'(\alpha) = 0$. By the induction hypothesis (the case $r = k - 1$), there are scalars c_i such that

$$g' = c_1f_1' + \cdots + c_{k-1}f_{k-1}'.$$

Now let

$$h = g - \sum_{i=1}^{k-1} c_if_i. \tag{3-16}$$

Then h is a linear functional on V and (3-16) tells us that $h(\alpha) = 0$ for every α in N_k. By the preceding lemma, h is a scalar multiple of f_k. If $h = c_kf_k$, then

$$g = \sum_{i=1}^{k} c_if_i. \quad \blacksquare$$

The Transpose of a Linear Transformation

Suppose that we have two vector spaces over the field F, V, and W, and a linear transformation T from V into W. Then T induces a linear

be a linear transformation from V *into* W. *The null space of* T^t *is the annihilator of the range of* T. *If* V *and* W *are finite-dimensional, then*

(i) *rank* (T^t) = *rank* (T)

(ii) *the range of* T^t *is the annihilator of the null space of* T.

Proof. If g is in W^*, then by definition

$$(T^t g)(\alpha) = g(T\alpha)$$

for each α in V. The statement that g is in the null space of T^t means that $g(T\alpha) = 0$ for every α in V. Thus the null space of T^t is precisely the annihilator of the range of T.

Suppose that V and W are finite-dimensional, say dim $V = n$ and dim $W = m$. For (i): Let r be the rank of T, i.e., the dimension of the range of T. By Theorem 16, the annihilator of the range of T then has dimension $(m - r)$. By the first statement of this theorem, the nullity of T^t must be $(m - r)$. But then since T^t is a linear transformation on an m-dimensional space, the rank of T^t is $m - (m - r) = r$, and so T and T^t have the same rank. For (ii): Let N be the null space of T. Every functional in the range

of T^t is in the annihilator of N; for, suppose $f = T^t g$ for some g in W^*; then, if α is in N

$$f(\alpha) = (T^t g)(\alpha) = g(T\alpha) = g(0) = 0.$$

Now the range of T^t is a subspace of the space N^0, and

$$\dim N^0 = n - \dim N = \text{rank } (T) = \text{rank } (T^t)$$

so that the range of T^t must be exactly N^0. ∎

Theorem

Let V *and* W *be finite-dimensional vector spaces over the*

field F. *Let* $\mathfrak{B}$ *be an ordered basis for* V *with dual basis* $\mathfrak{B}^*$, *and let* $\mathfrak{B}'$ *be an ordered basis for* W *with dual basis* $\mathfrak{B}'^*$. *Let* T *be a linear transformation from* V *into* W; *let* A *be the matrix of* T *relative to* $\mathfrak{B}$, $\mathfrak{B}'$ *and let* B *be the matrix of* T^t *relative to* $\mathfrak{B}'^*$, $\mathfrak{B}^*$. *Then* $B_{ij} = A_{ji}$.

Proof. Let

$$\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}, \qquad \mathfrak{B}' = \{\beta_1, \ldots, \beta_m\},$$
$$\mathfrak{B}^* = \{f_1, \ldots, f_n\}, \qquad \mathfrak{B}'^* = \{g_1, \ldots, g_m\}.$$

By definition,

$$T\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i, \qquad j = 1, \ldots, n$$

$$T^t g_j = \sum_{i=1}^{n} B_{ij} f_i, \qquad j = 1, \ldots, m.$$

On the other hand,

$$\begin{aligned}(T^t g_j)(\alpha_i) &= g_j(T\alpha_i) \\ &= g_j\left(\sum_{k=1}^{m} A_{ki}\beta_k\right) \\ &= \sum_{k=1}^{m} A_{ki} g_j(\beta_k) \\ &= \sum_{k=1}^{m} A_{ki}\delta_{jk} \\ &= A_{ji}.\end{aligned}$$

For any linear functional f on V

$$f = \sum_{i=1}^{m} f(\alpha_i) f_i.$$

If we apply this formula to the functional $f = T^t g_j$ and use the fact that $(T^t g_j)(\alpha_i) = A_{ji}$, we have

$$T^t g_j = \sum_{i=1}^{n} A_{ji} f_i$$

from which it immediately follows that $B_{ij} = A_{ji}$. ∎

Definition. *If* A *is an* m × n *matrix over the field* F, *the* **transpose** *of* A *is the* n × m *matrix* A^t *defined by* $A^t_{ij} = A_{ji}$.

Theorem 23 thus states that if T is a linear transformation from V into W, the matrix of which in some pair of bases is A, then the transpose transformation T^t is represented in the dual pair of bases by the transpose matrix A^t.

Theorem 24. *Let* A *be any* m × n *matrix over the field* F. *Then the row rank of* A *is equal to the column rank of* A.

Proof. Let $\mathfrak{B}$ be the standard ordered basis for F^n and $\mathfrak{B}'$ the standard ordered basis for F^m. Let T be the linear transformation from F^n into F^m such that the matrix of T relative to the pair $\mathfrak{B}$, $\mathfrak{B}'$ is A, i.e.,

$$T(x_1, \ldots, x_n) = (y_1, \ldots, y_m)$$

where

$$y_i = \sum_{j=1}^{n} A_{ij}x_j.$$

The column rank of A is the rank of the transformation T, because the range of T consists of all m-tuples which are linear combinations of the column vectors of A.

Relative to the dual bases $\mathfrak{B}'^*$ and $\mathfrak{B}^*$, the transpose mapping T^t is represented by the matrix A^t. Since the columns of A^t are the rows of A, we see by the same reasoning that the row rank of A (the column rank of A^t) is equal to the rank of T^t. By Theorem 22, T and T^t have the same rank, and hence the row rank of A is equal to the column rank of A. ∎

Now we see that if A is an $m \times n$ matrix over F and T is the linear transformation from F^n into F^m defined above, then

$$\text{rank } (T) = \text{row rank } (A) = \text{column rank } (A)$$

and we shall call this number simply the **rank** of A.

EXAMPLE Let V be an n-dimensional vector space over the field F, and let T be a linear operator on V. Suppose $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V. The matrix of T in the ordered basis $\mathfrak{B}$ is defined to be the $n \times n$ matrix A such that

$$T\alpha_j = \sum_{j=1}^{n} A_{ij}\alpha_i$$

in other words, A_{ij} is the ith coordinate of the vector $T\alpha_j$ in the ordered basis $\mathfrak{B}$. If $\{f_1, \ldots, f_n\}$ is the dual basis of $\mathfrak{B}$, this can be stated simply

$$A_{ij} = f_i(T\alpha_j).$$

Let us see what happens when we change basis. Suppose

$$\mathfrak{B}' = \{\alpha_1', \ldots, \alpha_n'\}$$

is another ordered basis for V, with dual basis $\{f_1', \ldots, f_n'\}$. If B is the matrix of T in the ordered basis $\mathfrak{B}'$, then

$$B_{ij} = f_i'(T\alpha_j').$$

Let U be the invertible linear operator such that $U\alpha_j = \alpha_j'$. Then the transpose of U is given by $U^t f_i' = f_i$. It is easy to verify that since U is invertible, so is U^t and $(U^t)^{-1} = (U^{-1})^t$. Thus $f_i' = (U^{-1})^t f_i$, $i = 1, \ldots, n$. Therefore,

$$\begin{aligned} B_{ij} &= [(U^{-1})^t f_i](T\alpha_j') \\ &= f_i(U^{-1}T\alpha_j') \\ &= f_i(U^{-1}TU\alpha_j). \end{aligned}$$

Now what does this say? Well, $f_i(U^{-1}TU\alpha_j)$ is the i, j entry of the matrix of $U^{-1}TU$ in the ordered basis $\mathfrak{B}$. Our computation above shows that this scalar is also the i, j entry of the matrix of T in the ordered basis $\mathfrak{B}'$. In other words

$$\begin{aligned} [T]_{\mathfrak{B}'} &= [U^{-1}TU]_{\mathfrak{B}} \\ &= [U^{-1}]_{\mathfrak{B}}[T]_{\mathfrak{B}}[U]_{\mathfrak{B}} \\ &= [U]_{\mathfrak{B}}^{-1}[T]_{\mathfrak{B}}[U]_{\mathfrak{B}} \end{aligned}$$

and this is precisely the change-of-basis formula which we derived earlier.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari (Po),
Coimbatore –641 021

Subject: RING THEORY AND LINEAR ALGEBRA-II **Subject Code: 16MMU403**

Class : II B.Sc Mathematics **Semester : IV**

UNIT -III

PART A (20x1=20 Marks)

(Question Nos. 1 to 20 Online Examinations)

Possible Questions

Question	Choice 1	Choice 2	Choice 3	Choice 4	Answer
Every …….. transformation is a linear transformation.	matrix	row	column	unit	matrix
Every matrix transformation is a…….transformation.	linear	non linear	homogeneous	non homogeneous	linear
…….. transformation preserve the operations of vector addition and scalar multiplication.	linear	non linear	matrix	row	linear
Linear transformation preserve the …….. of vector addition and scalar multiplication.	addition	functions	operations	values	operations
Linear transformation preserve the operations of ……. and scalar multiplication.	vector addition	vector subtraction	vector multiplication	vector division	vector addition
Linear transformation preserve the operations of vector addition and ……..	vector multiplication	scalar multiplication	matrix multiplication	vector division	scalar multiplication
If T is a linear transformation , then T(0)=……	0	1	2	3	0
T(cu+dv)=……..	T(cu)+T(dv)	cT(u)-dT(v)	T(u)+T(v)	cT(u)+dT(v)	cT(u)+dT(v)
Let T be a linear transformation then there exists a unique matrix A such that T(x)=……. for all x in R	0	Ax	x	1	Ax
Let T be a linear transformation then there exists a …….. matrix A such that T(x)=Ax for all x in R	zero	unique	identity	diagonal	unique
An nxn matrix B such that BA=I is called a ……of A	zero	left inverse	right inverse	identity	left inverse
An ……matrix B such that BA=I is called a left inverse of A	mxm	nxn	mxn	nxm	nxn
An nxn matrix B such that AB=I is called a ……of A	zero	left inverse	right inverse	identity	right inverse
An ……matrix B such that AB=I is called a right inverse of A	mxm	nxn	mxn	nxm	nxn
If AB=BA=I then B is called a ……..inverse of A.	two sided	left inverse	right inverse	identity	two sided
If AB=BA=……. then B is called a two sided inverse of A.	0	1	I	-1	I
A two sided inverse of Aand Ais said to be …….	invertible	inverse	identity	vertible	invertible
If A is invertible,so is A^{-1} and $(A^{-1})^{-1}$=…….	A^{-1}	A	0	c	A
If A is ……,so is A^{-1} and $(A^{-1})^{-1}$=A	invertible	inverse	identity	vertible	invertible
If both A and B are invertible ,so is AB,and $(AB)^{-1}$=……	B^{-1}	A^{-1}	BA	$B^{-1}A^{-1}$	$B^{-1}A^{-1}$
If both A and B are …….. ,so is AB,and $(AB)^{-1}=B^{-1}A^{-1}$	invertible	inverse	identity	vertible	invertible
A …….. of invertible matrices is invertible	addition	subtraction	product	division	product
A product of invertible …… is invertible	matrices	functions	vectors	equations	matrices
A product of invertible matrices is ……	invertible	inverse	identity	vertible	invertible
An ……matrix is invertible.	null	identity	elementary	singular	elementary
An elementary matrix is…….	invertible	inverse	identity	vertible	invertible
A …….. of V is a subset W of V which is itself a vectorspace over F with the operations of vector addition and scalar multiplication on V.	subspace	space	vector	function	subspace
A subspace of V is a subset W of V which is itself a vectorspace over F with the …….. of vector addition and scalar multiplication on V.	functions	operations	scalar	vector	operations
A subspace of V is a subset W of V which is itself a vectorspace over F with the operations of ……and scalar multiplication on V.	vector addition	vector subtraction	vector multiplication	vector division	vector addition
A subspace of V is a subset W of V which is itself a vectorspace over F with the operations of vector addition and …….. on V.	vector multiplication	scalar multiplication	matrix multiplication	vector division	scalar multiplication
The ……consisting of the zero vector alone is a subspace of V, called zero subspace of V.	subset	set	space	subspace	subset
The subset consisting of the……. vector alone is a subspace of V, called zero subspace of V.	zero	unit	finite	infinite	zero
The subset consisting of the zero vector alone is a subspace of V, called …….. of V.	zero subspace	zero space	zero subset	zero set	zero subspace
An …….. matrix A over the field F is symmetric if Aij=Aji for each i and j.	mxm	nxn	mxn	nxm	nxn
An nxn matrix A over the …… F is symmetric if Aij=Aji for each i and j.	field	scalar	vector	matrix	field
An nxn matrix A over the field F is ……..if Aij=Aji for each i and j.	symmetric	non symmetric	singular	non singular	symmetric
An nxn matrix A over the field F is symmetric if …….. for each i and j.	Aij<Aji	Aij>Aji	Aij=Aji	Aij≠Aji	Aij=Aji
Any set which contains a lineary dependent set is ……...	linearly dependent	linearly independent	linear	non linear	linearly dependent
Any subset of a lineary independent set is ……..	linearly dependent	linearly independent	linear	non linear	linearly independent
Any set which contains the ……..vector is linearly dependent.	0	unit	inverse	complex	0
Any set which contains the 0 vector is……..	linearly dependent	linearly independent	linear	non linear	linearly dependent
A set S of vectors is …… iff each finite subset of S is linearly independent.	linearly dependent	linearly independent	linear	non linear	linearly independent
A set S of vectors is linearly independent iff each ……. subset of S is linearly independent.	one	finite	infinite	null	finite

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari (Po),
Coimbatore –641 021

Subject: RING THEORY AND LINEAR ALGEBRA-II **Subject Code: 16MMU403**
Class : II B.Sc Mathematics **Semester : IV**

UNIT -IV

PART A (20x1=20 Marks)

(Question Nos. 1 to 20 Online Examinations)

Possible Questions

Question	Choice 1	Choice 2	Choice 3	Choice 4	Answer
Hom(V,V) is the set of all vector space-homomorphisms of V into itself	V into V	V into W	W into W	W into V	V into V
A linear transformation on V,over F is an element of ----------	$A_F(W)$	$B_F(V)$	$A_F(V)$	$W_F(V)$	$A_F(V)$
A linear transformation on ---------- is an element of $A_F(V)$	W over F	V over V	F over F	V over F	V over F
An element $T \in A(V)$ is called ---------- if there exists an $S \in A(V)$ such that TS =1	both invertible	right-invertible	left-invertible	invertible	right-invertible
An element ---------- is called right-invertible if there exists an $S \in A(V)$ such that TS =1	$V \in A(V)$	$T \in A(T)$	$T \in A(V)$	$T \in A(T)$	$T \in A(V)$
An element $T \in A(V)$ is called right-invertible if there exists an $S \in A(V)$ such that ----------	TS =1	TS=0	ST =1	TS =2	TS =1
The ----------$T \in A(V)$ is said to be unitary if $(uT,vT)=(u,v)$ for all $u,v \in V$	normal transformation	linear transformation	unitary	Nilpotent transformation	linear transformation
The linear transformation $T \in A(V)$ is said to be---------- if $(uT,vT)=(u,v)$ for all $u,v \in V$	normal transformation	linear transformation	unitary	Nilpotent transformation	unitary
The linear transformation $T \in A(V)$ is said to be unitary if $(uT,vT)=$ for all $u,v \in V$	(u,v)	uv	uT	vT	(u,v)
The ---------- Ton V is unitary if and only if it takes an orthonormal basis of V into an orthonormal basis of V	normal transformation	linear transformation	unitary	Nilpotent transformation	linear transformation
The linear transformation Ton V is unitary if and only if it takes an---------- of V into an orthonormal basis of V	basis	orthogonal basis	orthonormal basis	normal basis	orthonormal basis
The linear transformation Ton V is unitary if and only if it takes an orthonormal basis of V into an -------- of V	basis	orthogonal basis	orthonormal basis	normal basis	orthonormal basis
If V is finite dimensional and ---------- there is an $S \in A(V)$ such that $E=TS \neq 0$ is an idempotent	T>0	T=0	$T \neq 0$	T<0	
If V is finite dimensional and $T \neq 0$ there is an $S \in A(V)$ such that E=-------- is an idempotent	T>0	$TS \neq 0$	$T \neq 0$	TS=0	
If V is finite dimensional and $T \neq 0$ there is an $S \in A(V)$ such that $E=TS \neq 0$ is an ----------	regular	idempotent	grounded	nilpotent	
The ---------- W of V is invariant under $T \in A(V)$ if $WT \subset W$	subspace	space	field	sub field	
The subspace W of V is invariant under $T \in A(V)$ if----------	W over F	$W \subset TV$	$WT \subset T$	W= TV	
The element $\lambda \in F$ is a characteristic root of $T \in A(V)$ if and only if for some ------- in V, $vT=\lambda v$	$\lambda=0$	$\lambda \neq 0$	v=0	$v \neq 0$	
The element $\lambda \in F$ is a characteristic root of $T \in A(V)$ if and only if for some $v \neq 0$ in V,----------	vT=T	$vT=\lambda v$	vT=v	Tv=T	
If $T \in A(V)$ is nilpotent then k is called the ----------of $Tk=0$ but $T^{k-1} \neq 0$	index of nilpotence	nilpotence	linear transformation	idempotent	

UNIT-IV

SYLLABUS

Inner product spaces and norms, Gram-Schmidt orthogonalisation process, orthogonal complements, Bessel's inequality, the adjoint of a linear operator.

KAHE

Inner Product Spaces

Definition. *Let* F *be the field of real numbers or the field of complex numbers, and* V *a vector space over* F. *An* **inner product** *on* V *is a function which assigns to each ordered pair of vectors* α, β *in* V *a scalar* $(\alpha|\beta)$ *in* F *in such a way that for all* α, β, γ *in* V *and all scalars* c

(a) $(\alpha + \beta|\gamma) = (\alpha|\gamma) + (\beta|\gamma)$;

(b) $(c\alpha|\beta) = c(\alpha|\beta)$;

(c) $(\beta|\alpha) = \overline{(\alpha|\beta)}$, *the bar denoting complex conjugation*;

(d) $(\alpha|\alpha) > 0$ *if* $\alpha \neq 0$.

It should be observed that conditions (a), (b), and (c) imply that

(e) $$(\alpha|c\beta + \gamma) = \bar{c}(\alpha|\beta) + (\alpha|\gamma).$$

One other point should be made. When F is the field R of real numbers, the complex conjugates appearing in (c) and (e) are superfluous; however, in the complex case they are necessary for the consistency of the conditions. Without these complex conjugates, we would have the contradiction:

$$(\alpha|\alpha) > 0 \quad \text{and} \quad (i\alpha|i\alpha) = -1(\alpha|\alpha) > 0.$$

In the examples that follow and throughout the chapter, F is either the field of real numbers or the field of complex numbers.

EXAMPLE On F^n there is an inner product which we call the **standard inner product.** It is defined on $\alpha = (x_1, \ldots, x_n)$ and $\beta = (y_1, \ldots, y_n)$ by

(8-1) $$(\alpha|\beta) = \sum_j x_j \bar{y}_j.$$

When $F = R$, this may also be written

$$(\alpha|\beta) = \sum_j x_j y_j.$$

In the real case, the standard inner product is often called the dot or scalar product and denoted by $\alpha \cdot \beta$.

Inner Product Spaces

Definition. *An* **inner product space** *is a real or complex vector space, together with a specified inner product on that space.*

A finite-dimensional real inner product space is often called a **Euclidean space.** A complex inner product space is often referred to as a **unitary space.**

Theorem

If V *is an inner product space, then for any vectors* α, β *in* V *and any scalar* c

(i) $||c\alpha|| = |c|\,||\alpha||$;

(ii) $||\alpha|| > 0$ *for* $\alpha \neq 0$;
(iii) $|(\alpha|\beta)| \leq ||\alpha||\,||\beta||$;
(iv) $||\alpha + \beta|| \leq ||\alpha|| + ||\beta||$.

Proof. Statements (i) and (ii) follow almost immediately from the various definitions involved. The inequality in (iii) is clearly valid when $\alpha = 0$. If $\alpha \neq 0$, put

$$\gamma = \beta - \frac{(\beta|\alpha)}{||\alpha||^2}\alpha.$$

Then $(\gamma|\alpha) = 0$ and

$$\begin{aligned} 0 \leq ||\gamma||^2 &= \left(\beta - \frac{(\beta|\alpha)}{||\alpha||^2}\alpha \,\middle|\, \beta - \frac{(\beta|\alpha)}{||\alpha||^2}\alpha\right) \\ &= (\beta|\beta) - \frac{(\beta|\alpha)(\alpha|\beta)}{||\alpha||^2} \\ &= ||\beta||^2 - \frac{|(\alpha|\beta)|^2}{||\alpha||^2}. \end{aligned}$$

Hence $|(\alpha|\beta)|^2 \leq ||\alpha||^2||\beta||^2$. Now using (c) we find that

$$\begin{aligned} ||\alpha + \beta||^2 &= ||\alpha||^2 + (\alpha|\beta) + (\beta|\alpha) + ||\beta||^2 \\ &= ||\alpha||^2 + 2\,\mathrm{Re}\,(\alpha|\beta) + ||\beta||^2 \\ &\leq ||\alpha||^2 + 2\,||\alpha||\,||\beta|| + ||\beta||^2 \\ &= (||\alpha|| + ||\beta||)^2. \end{aligned}$$

Thus, $||\alpha + \beta|| \leq ||\alpha|| + ||\beta||$. ∎

The inequality in (iii) is called the **Cauchy-Schwarz inequality.** It has a wide variety of applications. The proof shows that if (for example) α is non-zero, then $|(\alpha|\beta)| < ||\alpha||\,||\beta||$ unless

$$\beta = \frac{(\beta|\alpha)}{||\alpha||^2}\,\alpha.$$

Thus, equality occurs in (iii) if and only if α and β are linearly dependent.

Definitions. *Let α and β be vectors in an inner product space* V. *Then α is* **orthogonal** *to β if $(\alpha|\beta) = 0$; since this implies β is orthogonal to α, we often simply say that α and β are orthogonal. If* S *is a set of vectors in* V, S *is called an* **orthogonal set** *provided all pairs of distinct vectors in* S *are orthogonal. An* **orthonormal set** *is an orthogonal set* S *with the additional property that $||\alpha|| = 1$ for every α in* S.

EXAMPLE

Let V be the space of continuous complex-valued (or real-valued) functions on the interval $0 \leq x \leq 1$ with the inner product

$$(f|g) = \int_0^1 f(x)\overline{g(x)}\,dx.$$

Suppose $f_n(x) = \sqrt{2}\,\cos 2\pi nx$ and that $g_n(x) = \sqrt{2}\,\sin 2\pi nx$. Then $\{1, f_1, g_1, f_2, g_2, \ldots\}$ is an infinite orthonormal set. In the complex case, we may also form the linear combinations

$$\frac{1}{\sqrt{2}}\,(f_n + ig_n), \qquad n = 1, 2, \ldots.$$

In this way we get a new orthonormal set S which consists of all functions of the form

$$h_n(x) = e^{2\pi inx}, \qquad n = \pm 1, \pm 2, \ldots.$$

The set S' obtained from S by adjoining the constant function 1 is also orthonormal. We assume here that the reader is familiar with the calculation of the integrals in question.

Theorem

An orthogonal set of non-zero vectors is linearly inde-

pendent.

Proof. Let S be a finite or infinite orthogonal set of non-zero vectors in a given inner product space. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_m$ are distinct vectors in S and that

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m.$$

Then

$$\begin{aligned}(\beta|\alpha_k) &= (\sum_j c_j\alpha_j|\alpha_k) \\ &= \sum_j c_j(\alpha_j|\alpha_k) \\ &= c_k(\alpha_k|\alpha_k).\end{aligned}$$

Since $(\alpha_k|\alpha_k) \neq 0$, it follows that

$$c_k = \frac{(\beta|\alpha_k)}{||\alpha_k||^2}, \qquad 1 \leq k \leq m.$$

Thus when $\beta = 0$, each $c_k = 0$; so S is an independent set. ∎

Corollary. *If a vector β is a linear combination of an orthogonal sequence of non-zero vectors $\alpha_1, \ldots, \alpha_m$, then β is the particular linear combination*

$$\beta = \sum_{k=1}^{m} \frac{(\beta|\alpha_k)}{||\alpha_k||^2}\alpha_k. \tag{8-8}$$

This corollary follows from the proof of the theorem. There is another corollary which although obvious, should be mentioned. If $\{\alpha_1, \ldots, \alpha_m\}$ is an orthogonal set of non-zero vectors in a finite-dimensional inner product space V, then $m \leq \dim V$. This says that the number of mutually orthogonal directions in V cannot exceed the algebraically defined dimension of V. The maximum number of mutually orthogonal directions in V is what one would intuitively regard as the geometric dimension of V, and we have just seen that this is not greater than the algebraic dimension. The fact that these two dimensions are equal is a particular corollary of the next result.

Theorem

Let V be an inner product space and let $\beta_1, \ldots, \beta_n$ be

any independent vectors in V. *Then one may construct orthogonal vectors* $\alpha_1, \ldots, \alpha_n$ *in* V *such that for each* $k = 1, 2, \ldots, n$ *the set*

$$\{\alpha_1, \ldots, \alpha_k\}$$

is a basis for the subspace spanned by $\beta_1, \ldots, \beta_k$.

Proof. The vectors $\alpha_1, \ldots, \alpha_n$ will be obtained by means of a construction known as the **Gram-Schmidt orthogonalization process.** First let $\alpha_1 = \beta_1$. The other vectors are then given inductively as follows: Suppose $\alpha_1, \ldots, \alpha_m$ $(1 \leq m < n)$ have been chosen so that for every k

$$\{\alpha_1, \ldots, \alpha_k\}, \qquad 1 \leq k \leq m$$

is an orthogonal basis for the subspace of V that is spanned by $\beta_1, \ldots, \beta_k$. To construct the next vector α_{m+1}, let

$$\alpha_{m+1} = \beta_{m+1} - \sum_{k=1}^{m} \frac{(\beta_{m+1}|\alpha_k)}{||\alpha_k||^2}\,\alpha_k. \tag{8-9}$$

Then $\alpha_{m+1} \neq 0$. For otherwise β_{m+1} is a linear combination of $\alpha_1, \ldots, \alpha_m$ and hence a linear combination of $\beta_1, \ldots, \beta_m$. Furthermore, if $1 \leq j \leq m$, then

$$\begin{aligned}(\alpha_{m+1}|\alpha_j) &= (\beta_{m+1}|\alpha_j) - \sum_{k=1}^{m} \frac{(\beta_{m+1}|\alpha_k)}{||\alpha_k||^2}\,(\alpha_k|\alpha_j)\\ &= (\beta_{m+1}|\alpha_j) - (\beta_{m+1}|\alpha_j)\\ &= 0.\end{aligned}$$

Therefore $\{\alpha_1, \ldots, \alpha_{m+1}\}$ is an orthogonal set consisting of $m + 1$ non-zero vectors in the subspace spanned by $\beta_1, \ldots, \beta_{m+1}$. By Theorem 2, it is a basis for this subspace. Thus the vectors $\alpha_1, \ldots, \alpha_n$ may be constructed one after the other in accordance with (8-9). In particular, when $n = 4$, we have

$$\begin{aligned}\alpha_1 &= \beta_1\\ \alpha_2 &= \beta_2 - \frac{(\beta_2|\alpha_1)}{||\alpha_1||^2}\,\alpha_1\\ \alpha_3 &= \beta_3 - \frac{(\beta_3|\alpha_1)}{||\alpha_1||^2}\,\alpha_1 - \frac{(\beta_3|\alpha_2)}{||\alpha_2||^2}\,\alpha_2\\ \alpha_4 &= \beta_4 - \frac{(\beta_4|\alpha_1)}{||\alpha_1||^2}\,\alpha_1 - \frac{(\beta_4|\alpha_2)}{||\alpha_2||^2}\,\alpha_2 - \frac{(\beta_4|\alpha_3)}{||\alpha_3||^2}\,\alpha_3. \quad \blacksquare\end{aligned} \tag{8-10}$$

Corollary. *Every finite-dimensional inner product space has an orthonormal basis.*

Proof. Let V be a finite-dimensional inner product space and $\{\beta_1, \ldots, \beta_n\}$ a basis for V. Apply the Gram-Schmidt process to construct an orthogonal basis $\{\alpha_1, \ldots, \alpha_n\}$. Then to obtain an orthonormal basis, simply replace each vector α_k by $\alpha_k/||\alpha_k||$. ∎

EXAMPLE

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where a, b, c, and d are complex numbers. Set $\beta_1 = (a, b)$, $\beta_2 = (c, d)$, and suppose that $\beta_1 \neq 0$. If we apply the orthogonalization process to β_1, β_2, using the standard inner product in C^2, we obtain the following vectors:

$$\alpha_1 = (a, b)$$

$$\alpha_2 = (c, d) - \frac{((c, d)|(a, b))}{|a|^2 + |b|^2}(a, b)$$

$$= (c, d) - \frac{(c\bar{a} + d\bar{b})}{|a|^2 + |b|^2}(a, b)$$

$$= \left(\frac{cb\bar{b} - d\bar{b}a}{|a|^2 + |b|^2}, \frac{d\bar{a}a - c\bar{a}b}{|a|^2 + |b|^2}\right)$$

$$= \frac{\det A}{|a|^2 + |b|^2}(-\bar{b}, \bar{a}).$$

Now the general theory tells us that $\alpha_2 \neq 0$ if and only if β_1, β_2 are linearly independent. On the other hand, the formula for α_2 shows that this is the case if and only if $\det A \neq 0$.

Theorem

Let W *be a subspace of an inner product space* V *and*

let β be a vector in V.

(i) *The vector α in* W *is a best approximation to β by vectors in* W *if and only if $\beta - \alpha$ is orthogonal to every vector in* W.

(ii) *If a best approximation to β by vectors in* W *exists, it is unique.*

(iii) *If* W *is finite-dimensional and $\{\alpha_1, \ldots, \alpha_n\}$ is any orthonormal basis for* W, *then the vector*

$$\alpha = \sum_k \frac{(\beta|\alpha_k)}{||\alpha_k||^2} \alpha_k$$

is the (unique) best approximation to β by vectors in W.

Proof. First note that if γ is any vector in V, then $\beta - \gamma = (\beta - \alpha) + (\alpha - \gamma)$, and

$$||\beta - \gamma||^2 = ||\beta - \alpha||^2 + 2 \operatorname{Re} (\beta - \alpha|\alpha - \gamma) + ||\alpha - \gamma||^2.$$

Now suppose $\beta - \alpha$ is orthogonal to every vector in W, that γ is in W and that $\gamma \neq \alpha$. Then, since $\alpha - \gamma$ is in W, it follows that

$$\begin{aligned} ||\beta - \gamma||^2 &= ||\beta - \alpha||^2 + ||\alpha - \gamma||^2 \\ &> ||\beta - \alpha||^2. \end{aligned}$$

Conversely, suppose that $||\beta - \gamma|| \geq ||\beta - \alpha||$ for every γ in W. Then from the first equation above it follows that

$$2 \operatorname{Re} (\beta - \alpha|\alpha - \gamma) + ||\alpha - \gamma||^2 \geq 0$$

for all γ in W. Since every vector in W may be expressed in the form $\alpha - \gamma$ with γ in W, we see that

$$2 \operatorname{Re} (\beta - \alpha|\tau) + ||\tau||^2 \geq 0$$

for every τ in W. In particular, if γ is in W and $\gamma \neq \alpha$, we may take

$$\tau = -\frac{(\beta - \alpha|\alpha - \gamma)}{||\alpha - \gamma||^2} (\alpha - \gamma).$$

Then the inequality reduces to the statement

$$-2 \frac{|(\beta - \alpha|\alpha - \gamma)|^2}{||\alpha - \gamma||^2} + \frac{|(\beta - \alpha|\alpha - \gamma)|^2}{||\alpha - \gamma||^2} \geq 0.$$

This holds if and only if $(\beta - \alpha|\alpha - \gamma) = 0$. Therefore, $\beta - \alpha$ is orthogonal to every vector in W. This completes the proof of the equivalence of the two conditions on α given in (i). The orthogonality condition is evidently satisfied by at most one vector in W, which proves (ii).

Now suppose that W is a finite-dimensional subspace of V. Then we know, as a corollary of Theorem 3, that W has an orthogonal basis. Let $\{\alpha_1, \ldots, \alpha_n\}$ be any orthogonal basis for W and define α by (8-11). Then, by the computation in the proof of Theorem 3, $\beta - \alpha$ is orthogonal to each of the vectors α_k ($\beta - \alpha$ is the vector obtained at the last stage when the orthogonalization process is applied to $\alpha_1, \ldots, \alpha_n, \beta$). Thus $\beta - \alpha$ is orthogonal to every linear combination of $\alpha_1, \ldots, \alpha_n$, i.e., to every vector in W. If γ is in W and $\gamma \neq \alpha$, it follows that $||\beta - \gamma|| > ||\beta - \alpha||$. Therefore, α is the best approximation to β that lies in W. ∎

Definition. *Let* V *be an inner product space and* S *any set of vectors in* V. *The* **orthogonal complement** *of* S *is the set* $S^\perp$ *of all vectors in* V *which are orthogonal to every vector in* S.

The orthogonal complement of V is the zero subspace, and conversely $\{0\}^\perp = V$. If S is any subset of V, its orthogonal complement $S^\perp$ (S perp) is always a subspace of V. For S is non-empty, since it contains 0; and whenever α and β are in $S^\perp$ and c is any scalar,

$$\begin{aligned}(c\alpha + \beta|\gamma) &= c(\alpha|\gamma) + (\beta|\gamma)\\ &= c0 + 0\\ &= 0\end{aligned}$$

for every γ in S, thus $c\alpha + \beta$ also lies in S. In Theorem 4 the characteristic property of the vector α is that it is the only vector in W such that $\beta - \alpha$ belongs to $W^\perp$.

$$||\beta - \gamma||^2 = ||E\beta||^2 + ||\beta - E\beta - \gamma||^2$$
$$\geq ||\beta - (\beta - E\beta)||^2$$

with strict inequality when $\gamma \neq \beta - E\beta$. Therefore, $\beta - E\beta$ is the best approximation to β by vectors in $W^\perp$. ▌

Definition. *Whenever the vector α in Theorem 4 exists it is called the* **orthogonal projection of β on W.** *If every vector in V has an orthogonal projection on W, the mapping that assigns to each vector in V its orthogonal projection on W is called the* **orthogonal projection of V on W.**

By Theorem 4, the orthogonal projection of an inner product space on a finite-dimensional subspace always exists. But Theorem 4 also implies the following result.

Corollary. *Let V be an inner product space, W a finite-dimensional subspace, and E the orthogonal projection of V on W. Then the mapping*

$$\beta \rightarrow \beta - E\beta$$

is the orthogonal projection of V on $W^\perp$.

Proof. Let β be an arbitrary vector in V. Then $\beta - E\beta$ is in $W^\perp$, and for any γ in $W^\perp$, $\beta - \gamma = E\beta + (\beta - E\beta - \gamma)$. Since $E\beta$ is in W and $\beta - E\beta - \gamma$ is in $W^\perp$, it follows that

Theorem
Let W be a finite-dimensional subspace of an inner product space V and let E be the orthogonal projection of V on W. Then E is an idempotent linear transformation of V onto W, $W^\perp$ is the null space of E, and

$$V = W \oplus W^\perp.$$

Proof. Let β be an arbitrary vector in V. Then $E\beta$ is the best approximation to β that lies in W. In particular, $E\beta = \beta$ when β is in W. Therefore, $E(E\beta) = E\beta$ for every β in V; that is, E is idempotent: $E^2 = E$. To prove that E is a linear transformation, let α and β be any vectors in

V and c an arbitrary scalar. Then, by Theorem 4, $\alpha - E\alpha$ and $\beta - E\beta$ are each orthogonal to every vector in W. Hence the vector

$$c(\alpha - E\alpha) + (\beta - E\beta) = (c\alpha + \beta) - (cE\alpha + E\beta)$$

also belongs to $W^\perp$. Since $cE\alpha + E\beta$ is a vector in W, it follows from Theorem 4 that

$$E(c\alpha + \beta) = cE\alpha + E\beta.$$

Of course, one may also prove the linearity of E by using (8-11). Again let β be any vector in V. Then $E\beta$ is the unique vector in W such that $\beta - E\beta$ is in $W^\perp$. Thus $E\beta = 0$ when β is in $W^\perp$. Conversely, β is in $W^\perp$ when $E\beta = 0$. Thus $W^\perp$ is the null space of E. The equation

$$\beta = E\beta + \beta - E\beta$$

shows that $V = W + W^\perp$; moreover, $W \cap W^\perp = \{0\}$. For if α is a vector in $W \cap W^\perp$, then $(\alpha|\alpha) = 0$. Therefore, $\alpha = 0$, and V is the direct sum of W and $W^\perp$. ▌

Corollary. *Under the conditions of the theorem,* $\mathrm{I} - \mathrm{E}$ *is the orthogonal projection of* V *on* $\mathrm{W}^\perp$. *It is an idempotent linear transformation of* V *onto* $\mathrm{W}^\perp$ *with null space* W.

Proof. We have already seen that the mapping $\beta \to \beta - E\beta$ is the orthogonal projection of V on $W^\perp$. Since E is a linear transformation, this projection on $W^\perp$ is the linear transformation $I - E$. From its geometric properties one sees that $I - E$ is an idempotent transformation of V onto W. This also follows from the computation

$$\begin{aligned}(I - E)(I - E) &= I - E - E + E^2 \\ &= I - E.\end{aligned}$$

Moreover, $(I - E)\beta = 0$ if and only if $\beta = E\beta$, and this is the case if and only if β is in W. Therefore W is the null space of $I - E$. ▌

The Gram-Schmidt process may now be described geometrically in the following way. Given an inner product space V and vectors $\beta_1, \ldots, \beta_n$ in V, let P_k $(k > 1)$ be the orthogonal projection of V on the orthogonal complement of the subspace spanned by $\beta_1, \ldots, \beta_{k-1}$, and set $P_1 = I$.

Then the vectors one obtains by applying the orthogonalization process to $\beta_1, \ldots, \beta_n$ are defined by the equations

$$\alpha_k = P_k\beta_k, \qquad 1 \leq k \leq n. \tag{8-12}$$

Theorem 5 implies another result known as **Bessel's inequality**.

Corollary. *Let* $\{\alpha_1, \ldots, \alpha_n\}$ *be an orthogonal set of non-zero vectors in an inner product space* V. *If* β *is any vector in* V, *then*

$$\sum_k \frac{|(\beta|\alpha_k)|^2}{||\alpha_k||^2} \leq ||\beta||^2$$

and equality holds if and only if

$$\beta = \sum_k \frac{(\beta|\alpha_k)}{||\alpha_k||^2}\,\alpha_k.$$

Proof. Let $\gamma = \sum_k [(\beta|\alpha_k)/||\alpha_k||^2]\,\alpha_k$. Then $\beta = \gamma + \delta$ where $(\gamma|\delta) = 0$. Hence

$$||\beta||^2 = ||\gamma||^2 + ||\delta||^2.$$

It now suffices to prove that

$$||\gamma||^2 = \sum_k \frac{|(\beta|\alpha_k)|^2}{||\alpha_k||^2}.$$

This is straightforward computation in which one uses the fact that $(\alpha_j|\alpha_k) = 0$ for $j \neq k$. ∎

In the special case in which $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal set, Bessel's inequality says that

$$\sum_k |(\beta|\alpha_k)|^2 \leq ||\beta||^2.$$

The corollary also tells us in this case that β is in the subspace spanned by $\alpha_1, \ldots, \alpha_n$ if and only if

$$\beta = \sum_k (\beta|\alpha_k)\,\alpha_k$$

or if and only if Bessel's inequality is actually an equality. Of course, in the event that V is finite dimensional and $\{\alpha_1, \ldots, \alpha_n\}$ is an orthogonal basis for V, the above formula holds for every vector β in V. In other words, if $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V, the kth coordinate of β in the ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ is $(\beta|\alpha_k)$.

Linear Functionals and Adjoints Theorem

Let V *be a finite-dimensional inner product space, and* f *a*

linear functional on V. *Then there exists a unique vector* β *in* V *such that* $f(\alpha) = (\alpha|\beta)$ *for all* α *in* V.

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthonormal basis for V. Put

$$\beta = \sum_{j=1}^{n} \overline{f(\alpha_j)}\alpha_j \tag{8-13}$$

and let f_β be the linear functional defined by

$$f_\beta(\alpha) = (\alpha|\beta).$$

Then

$$f_\beta(\alpha_k) = (\alpha_k| \sum_j \overline{f(\alpha_j)}\alpha_j) = f(\alpha_k).$$

Since this is true for each α_k, it follows that $f = f_\beta$. Now suppose γ is a vector in V such that $(\alpha|\beta) = (\alpha|\gamma)$ for all α. Then $(\beta - \gamma|\beta - \gamma) = 0$ and $\beta = \gamma$. Thus there is exactly one vector β determining the linear functional f in the stated manner. ∎

The proof of this theorem can be reworded slightly, in terms of the representation of linear functionals in a basis. If we choose an orthonormal basis $\{\alpha_1, \ldots, \alpha_n\}$ for V, the inner product of $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ and $\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$ will be

$$(\alpha|\beta) = x_1\bar{y}_1 + \cdots + x_n\bar{y}_n.$$

If f is any linear functional on V, then f has the form

$$f(\alpha) = c_1x_1 + \cdots + c_nx_n$$

for some fixed scalars $c_1, \ldots, c_n$ determined by the basis. Of course $c_j = f(\alpha_j)$. If we wish to find a vector β in V such that $(\alpha|\beta) = f(\alpha)$ for all α, then clearly the coordinates y_j of β must satisfy $\bar{y}_j = c_j$ or $y_j = \overline{f(\alpha_j)}$. Accordingly,

$$\beta = \overline{f(\alpha_1)}\alpha_1 + \cdots + \overline{f(\alpha_n)}\alpha_n$$

is the desired vector.

Some further comments are in order. The proof of Theorem 6 that we have given is admirably brief, but it fails to emphasize the essential geometric fact that β lies in the orthogonal complement of the null space of f. Let W be the null space of f. Then $V = W + W^\perp$, and f is completely determined by its values on $W^\perp$. In fact, if P is the orthogonal projection of V on $W^\perp$, then

$$f(\alpha) = f(P\alpha)$$

for all α in V. Suppose $f \neq 0$. Then f is of rank 1 and dim $(W^\perp) = 1$. If γ is any non-zero vector in $W^\perp$, it follows that

$$P\alpha = \frac{(\alpha|\gamma)}{||\gamma||^2}\gamma$$

for all α in V. Thus

$$f(\alpha) = (\alpha|\gamma) \cdot \frac{f(\gamma)}{||\gamma||^2}$$

for all α, and $\beta = [\overline{f(\gamma)}/||\gamma||^2]\,\gamma$.

Theorem

For any linear operator T on a finite-dimensional inner product space V, there exists a unique linear operator T on V such that*

(8-14) $$(T\alpha|\beta) = (\alpha|T^*\beta)$$

for all α, β in V.

Proof. Let β be any vector in V. Then $\alpha \to (T\alpha|\beta)$ is a linear functional on V. By Theorem 6 there is a unique vector β' in V such that $(T\alpha|\beta) = (\alpha|\beta')$ for every α in V. Let T^* denote the mapping $\beta \to \beta'$:

$$\beta' = T^*\beta.$$

We have (8-14), but we must verify that T^* is a linear operator. Let β, γ be in V and let c be a scalar. Then for any α,

$$\begin{aligned}(\alpha|T^*(c\beta + \gamma)) &= (T\alpha|c\beta + \gamma)\\ &= (T\alpha|c\beta) + (T\alpha|\gamma)\\ &= \bar{c}(T\alpha|\beta) + (T\alpha|\gamma)\end{aligned}$$

$$= \bar{c}(\alpha|T^*\beta) + (\alpha|T^*\gamma)$$
$$= (\alpha|cT^*\beta) + (\alpha|T^*\gamma)$$
$$= (\alpha|cT^*\beta + T^*\gamma).$$

Thus $T^*(c\beta + \gamma) = cT^*\beta + T^*\gamma$ and T^* is linear.

The uniqueness of T^* is clear. For any β in V, the vector $T^*\beta$ is uniquely determined as the vector β' such that $(T\alpha|\beta) = (\alpha|\beta')$ for every α. ∎

Theorem

Let V *be a finite-dimensional inner product space and let* $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ *be an (ordered) orthonormal basis for* V. *Let* T *be a linear operator on* V *and let* A *be the matrix of* T *in the ordered basis* $\mathfrak{B}$. *Then* $A_{kj} = (T\alpha_j|\alpha_k)$.

Proof. Since $\mathfrak{B}$ is an orthonormal basis, we have

$$\alpha = \sum_{k=1}^{n} (\alpha|\alpha_k)\alpha_k.$$

The matrix A is defined by

$$T\alpha_j = \sum_{k=1}^{n} A_{kj}\alpha_k$$

and since

$$T\alpha_j = \sum_{k=1}^{n} (T\alpha_j|\alpha_k)\alpha_k$$

we have $A_{kj} = (T\alpha_j|\alpha_k)$. ∎

POSSIBLE QUESTIONS

2 Mark Questions:

1. Define norm of a vector.
2. State Bessel's inequality.
3. Define orthogonal complement.
4. State Gram schmidt orthogonalization process.
5. Define inner product space.

6 Mark Questions:

6. If V is an inner product space, then for any vectors α, β in V and any scalar c
   i. $\|c\alpha\| = |c| \ \|\alpha\|$.
   ii. $\|\alpha\| > 0$ for $\alpha \neq 0$.
   iii. $|(\alpha/\beta)| \leq \|\alpha\| \ \|\beta\|$
   iv. $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$
7. Let V be a finite dimensional inner product space. If T and U are linear operator on V and c is a scalar, the prove
   i. $(T+U)^* = T^* + U^*$.
   ii. $(cT^*) = \bar{c}T^*$.
   iii. $(TU)^* = U^*T^*$.
   iv. $(T^*)^* = T$.
8. Prove that an orthogal set of non-zero vectors is linearly independent.

9. State and prove Besel's inequality.

10. State and prove Gram schmidt orthogonalization process.

11. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthogonal set of non zero vectors in an inner poduct space V. if β is any vector in V, then prove that

$$\sum_k \frac{|(\beta/\alpha_k)|^2}{\|\alpha_k\|^2} \leq \|\beta\|$$

And equality holds if and only if

$$\beta = \sum_k \frac{(\beta/\alpha_k)}{\|\alpha_k\|^2}\alpha_k$$

12. Find the orthonomar basis using Gram schmidt orthogonalization process for the vectors $\beta_1 = (3, 0, 4)$, $\beta_2 = (-1, 0, 7)$ and $\beta_3 = (2, 9, 11)$ in R^3.

13. Let V be a finite dimensional inner product space, and f a linear functional on V. then prove there exists a unique vector β in V such that $f(\alpha) = (\alpha|\beta)$ for all α in V.

14. Let W be a subspace of an inner product space V and let β be a vector in V then prove the following

    i. The vector α in W is a best approximation to β by vectors in W if and only if $\beta - \alpha$ is orthogonal to every vector in W.

    ii. If a best approximation to β by vectors in W exists, it is unique.

    iii. If W is finite dimensional and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is any orthonormal basis for W, then the vector

$$\alpha = \sum_k \frac{(\beta/\alpha_k)}{\|\alpha_k\|^2}\alpha_k$$

    is the (unique) best approximation to β by vectors in W.

15. For any linear operator T on a finite dimensional inner product space V, then prove there exists a unique linear operator T* on V such that

$$(T_\alpha|\beta) = (\alpha|T^*|\beta) \text{ for all } \alpha, \beta \text{ in V.}$$

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari (Po),
Coimbatore –641 021

Subject: RING THEORY AND LINEAR ALGEBRA-II **Subject Code: 16MMU403**
Class : II B.Sc Mathematics **Semester : IV**

UNIT -V

PART A (20x1=20 Marks)

(Question Nos. 1 to 20 Online Examinations)

Possible Questions

Question	Choice 1	Choice 2	Choice 3	Choice 4	Answer
The -------- of a matrix A is the sum of the elements on the main diagonal of A	transpose	inverse	trase	conjucate	trase
The trace of a matrix A is the ---------- of the elements on the main diagonal of A	sum	inverse	product	subtract	sum
The trace of a matrix A is the sum of the elements on the --------- of A	diagonal	main diagonal	elements	all elements	main diagonal
The matrix A is said tobe a---------- if A'=A	symmetric matrix	singular matrix	nonsingular matrix	skew- symmetric matrix	symmetric matrix
The matrix A is said to be a symmetric matrix if A'=A	A≠A'	A<A'	A>A'	A'=A	A'=A
The matrix A is said to be a---------- if A'=-A	symmetric matrix	singular matrix	nonsingular matrix	skew- symmetric matrix	skew- symmetric matrix
The matrix A is said to be a skew- symmetric matrix if --------	A≠A'	A<A'	A'=- A	A'=A	A'=- A
A and B are symmetric matrices,AB is ---------- iff AB=BA	symmetric matrix	singular matrix	nonsingular matrix	skew- symmetric matrix	symmetric matrix
A and B are symmetric matrices,AB is symmetric iff ----------	A=B	AB=BA	A≠B	AB≠BA	AB=BA
The determinant of a triangular matrix is the ---------- of its entries on the main diogonal	sum	inverse	product	subtract	product
The determinant of a triangular matrix is the product of its entries on the----------	diagonal	main diagonal	elements	all elements	main diagonal
The of a triangular matrix is the product of its entries on the main diogonal	transpose	inverse	trase	determinant	determinant
Interchanging two rows of A changing the sign of its ----------	transpose	inverse	trase	determinant	determinant
Interchanging two rows of A changing the---------- of its determinant	value	sign	sign and value	transpose	sign
Interchanging two columns of A changing the sign of its ----------	transpose	inverse	trase	determinant	determinant
Interchanging two columns of A changing the ------- of its determinant	value	sign	sign and value	transpose	sign
The characteristic roots of A are the roots with the correct multiplicity of the secular equation , ------- of A	det (x-A)	det(x-a)	dm (x-a)	dim(x+A)	det (x-A)
The ---------- of A are the roots with the correct multiplicity of the secular equation ,det (x-A) of A	root	multiple root	characteristic roots	product roots	characteristic roots
The characteristic roots of A are the roots with the correct multiplicity of the---------- , det(x-A) of A	linear equation	secular equation	non linear equation	non-secular equation	secular equation
A polynomial with coefficients which are ----------has all its roots in the complex field	real number	complex numbers	rational number	irrational number	complex numbers
A polynomial with coefficients which are complex numbers has all its------ in the complex field	root	multiple root	characteristic roots	product roots	root
A polynomial with coefficients which are complex numbers has all its roots in the ---	real field	rational field	complex field	irrational field	complex field
T∈A(V) is unitary if and only if----------	TT*=1	TT*>1	TT*<1	TT*≤1	TT*=1
T∈A(V) is --------- if and only if TT* =1	transformation	transformation	unitary	transformation	unitary
T∈A(V) is called harmitian if T*=T	transformation	harmition	unitary	transformation	harmition
T∈A(V) is called harmitian if ----------	T=T*	T=1	T≠T*	TT*=1	T≠T*
If T ∈A(V) isHermitian then all its --------- are real	root	multiple root	characteristic roots	product roots	characteristic roots
If T ∈A(V) isHermitian then all its characteristic roots are ----------	real	complex	rational	irrational	real
If T ∈A(V) is --------- then all its characteristic roots are real	transformation	harmition	unitary	transformation	harmition
If T ∈A(V) is --------- if TT*=T*T	normal	harmition	unitary	Nilpotent	normal
If T ∈A(V) is normal if ----------	T=T*	T=1	T=T*	TT*=1	T=T*
The Hermitian ---------- T is non negative if and only if its characterstic roots are non negative	normal transformation	linear transformation	unitary	Nilpotent transformation	linear transformation
The Hermitian linear transformation T is non negative if and only if its--------- are non negative	root	multiple root	characteristic roots	product roots	characteristic roots
The Hermitian linear transformation T is non negative if and only if its characterstic roots are ----------	non negative	negative	rational	irrational	non negative
The Hermitian linear transformation T is ----------if and only if its characterstic roots are non-negative	non negative	negative	rational	irrational	non negative
If N is normal and if vN^k=0then----------	vN=0	vN=1	k=1	vk=0	vN=0