(Springer Undergraduate Mathematics Series) Thomas

Further Linear Algebra-Springer (2002)

T.S. BIyth and E.F. Robertson Further Linear Algebra SPRINGER □ UNDERGRADUATE □ MATHEMATICS □ SERIES & Springer

Springer Undergraduate Mathematics Series

Other books in this series A First Course in Discrete Mathematics I. Anderson Analytic Methods for Partial Differential Equations G. Evans, J. Blackledge, P. Yardley Applied Geometry for Computer Graphics and CAD, Second Edition D. Marsh Basic Linear Algebra, Second Edition T.S. Blyth andE.F. Robertson Basic Stochastic Processes Z. Brzezniak and T. Zastawniak Calculus of One Variable K.E. Hirst Complex Analysis J.M. Howie Elementary Differential Geometry A. Pressley Elementary Number Theory G.A. Jones andJM. Jones Elements of Abstract Analysis M. OSearcdid Elements of Logic via Numbers and Sets D.L. Johnson Essential Mathematical Biology N.F. Britton Essential Topology M.D. Crossley Fields and Galois Theory J.M. Howie Fields, Flows and Waves: An Introduction to Continuum Models D.F. Parker Further Linear Algebra T.S. Blyth andEF. Robertson Geometry JR. Fenn Groups, Rings and Fields D.A.R. Wallace Hyperbolic Geometry, Second Edition J. W. Anderson Information and Coding Theory G.A Jones and J.M. Jones Introduction to Laplace Transforms and Fourier Series P.P.G. Dyke Introduction to Ring Theory P.M. Cohn Introductory Mathematics: Algebra and Analysis G. Smith Linear Functional Analysis B.P. RynneandM.A. Youngson Mathematics for Finance: An Introduction to Financial Engineering M. Capinksi and T. Zastawniak Matrix Groups: An Introduction to Lie Group Theory A. Baker Measure, Integral and Probability, Second Edition M. Capinksi and E. Kopp Multivariate Calculus and Geometry, Second Edition S. Dineen Numerical Methods for Partial Differential Equations G. Evans, J. Blackledge, P. Yardley Probability Models J.Haigh Real Analysis J.M. Howie Sets, Logic and Categories P. Cameron Special Relativity N.M.J. Woodhouse Symmetries D.L. Johnson Topics in Group Theory G. Smith and 0. Tabachnikova Vector Calculus P.C Matthews

T.S. Blyth and E.F. Robertson Further Linear Algebra Springer

Professor T.S.Blyth Professor E.F. Robertson School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews, KY169SS, UK Cevtr illustration fitments rtpnduad by Kmdpetmisam of Apcech Systems, Inc. Publisher of the GAUSS Mathematical and Statistical System. 23804 Si Kent-Kangky Road. Maple VaUey. WA 98038. USA. TeL (206) 432 - 78SS Fax (206) 432 • 7832 email: info9aptcch.com URL www^ptechxom AroericanStotisticalAssocutiori:C"naiiceVol8Nol.i995artidety Springa-VerUg: Muhematka in Education and Research Vol 4 Issue 3 1995 article by Roman E Maeder, Beatrice Amrhein and Oliver Gloor 'Illustrated Mathematics: Visualization of Mathematical Objects' page 9 fig 11, originally published as a CD ROM 'Illustrated Mathematics' by TEL0& ISBN 0-387-14222-3. German edition by Birkhausen ISBN 3-7643-51O0-*. Mathematica in Education and Research Vol 4 Issue 3 1995 artide by Richard J Gaylord and Kazume Nishidate Traffic Engineenng with Cellular Automata' page 35 fig 2. Mathematica in Education and Research Vol S Issue 2 1996 artide by Michad Trott The Implicitization of a Trefoil Knot'page 14. Mathematica in Education and Research Vol S Issue 2 1996 artide by Lee de Cola 'Coins, Trees, Ban and Bells: Simulation of the Binomial Process' page 19 fig 3. Mathematica in Education and Research Vol S Issue 2 1996 article by Richard Gaylord and Kazume Nishidate 'Contagious Spreading* page 33 fig I. Mathematica in Education and Research Vol S Issue 2 1996 artide by Joe Buhler and Stan Wagon 'Secrets of the Middling Constant* page 50 fig I. British Library Cataloguing in Publication Data A catalogue record for this book is available from the British Library Library of Congress Cataloging-in-Publication Data Blyth, T.S. (Thomas Scon) Further linear algebra / T.S. Blyth and E.F. Robertson p. cm. - (Springer undergraduate mathematics series, ISSN 1615-2085) Includes index. ISBN 1-85233-425-8 (acid-free paper) LAlgebras, Linear. I. Robertson, E.F. II. Title. III. Series. QA184.2.B59 2002 512\55—dc21 2001042987 Springer Undergraduate Mathematics Series ISSN 1615-2085 ISBN 1-85233-425-8 Springer Science* Business Media springer.com C

Springer-Verlag London Limited 2002 Printed in Great Britain 2nd printing 2006 The use of registered names, trademarks etc in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use. The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made. Typesetting: Camera ready by T.S.B. Printed and bound at the Athenaeum Press Ltd., Gateshead, Tyne & Wear 12/3830-54321 Printed on acid-free paper SPIN 11576679

Preface Most of the introductory courses on linear algebra develop the basic theory of finite- dimensional vector spaces, and in so doing relate the notion of a linear mapping to that of a matrix. Generally speaking, such courses culminate in the diagonalisation of certain matrices and the application of this process to various situations. Such is the case, for example, in our previous SUMS volume Basic Linear Algebra. The present text is a continuation of that volume, and has the objective of introducing the reader to more advanced properties of vector spaces and linear mappings, and consequently of matrices. For readers who are not familiar with the contents of Basic Linear Algebra we provide an introductory chapter that consists of a compact summary of the prerequisites for the present volume. In order to consolidate the student's understanding we have included a large number of illustrative and worked examples, as well as many exercises that are strategically placed throughout the text. Solutions to the exercises are also provided. Many applications of linear algebra require careful, and at times rather tedious, calculations by hand. Very often these are subject to error, so the assistance of a computer is welcome. As far as computation in algebra is concerned, there are several packages available. Here we include, in the spirit of a tutorial, a chapter that gives a brief introduction to the use of MAPLE1 in dealing with numerical and algebraic problems in linear algebra. Finally, for the student who is keen on the history of mathematics, we include a chapter that contains brief biographies of those mathematicians mentioned in our two volumes on linear algebra. The biographies emphasise their contributions to linear algebra. T.S.B., E.F.R. St Andrews 'MAPLE™ is a registered trademark of Waterloo Maple Inc., 57 Erb Street West, Waterloo, Ontario, Canada, N2L6C2. www.maplesoft.coin

Contents

The story so far.... Our title Further Linear Algebra suggests already that the reader will be familiar with the basics of this discipline. Since for each reader these can be different, depending on the content of the courses taken previously, we devote this introductory chapter to a compact summary of the prerequisites, thereby offering the reader a smooth entry to the material that follows. Readers already familiar with our previous SUMS volume Basic Linear Algebra may comfortably proceed immediately to Chapter 1.

 In essence, linear algebra is the study of vector spaces. The notion of a vector space holds a ubiquitous place in mathematics and plays an important role, along with the attendant notion of a matrix, in many diverse applications. If F is a given field then by a vector space over F, or an F-vector space, we mean an additive abelian group V together with an 'action* of F on V, described by (X, x) *-* Xx, such that the following axioms hold: (1) {VXeF)(Vx,yeV) X{x + y) = Xx + Xy; (2) (VA,/i € F){Vx e V) (X + v)x = Xx + iix\ (3) (VA./zeFHvxev) {Xn)x = x{^x)\ (4) (VxeV) \Fx = x. Additional properties that follow immediately from these axioms are: (5) (VA6F) M),, = 0,,; (6) (VxGV) Ofx = (V; (7) if Xx = 0V then either X = 0F or x = <fy; (8) (VA 6 F)(Vx € V) {-X)x = -{Xx) = X(-x). We often refer to the field F as the 'ground field* of the vector space V and the elements of F as 'scalars*. When F is the field IR of real numbers we say that V is a real vector space; and when F is the field C of complex numbers we say that V is a complex vector space. As with every algebraic structure, there are two important items to consider in relation to vector spaces, namely the substructures (i.e. the subsets that are also vector spaces) and the morphisms (i.e. the structure-preserving mappings between vector spaces). As for the substructures, we say that a non-empty subset W of an F-vector space V is a subspace if it is closed under the operations of V, in the sense that (1) ifx.yeWthenx + yeW; (2) ifxeWandXeFlhenXxeW. These two conditions may be expressed as the single condition (V^€F)(Vx,yeW) Xx + ny<EW. Every subspace of a vector space V

contains 0V, so {0V} is the smallest subspace. The set-theoretic intersection of any collection of subspaces of a vector space V is also a subspace of V. In particular, if S is any non-empty subset of V then there is a smallest subspace of V that contains 5, namely the intersection of all the subspaces that contain S. This can be usefully identified as follows. We say that v £ V is a linear combination of elements of 5 if there exist xY,..., xn 6 S and X,,..., Xn 6 F such that n v= £>«*« = Mi + •■■ + *,, *,,• i=l

 Then the set of linear combinations of elements of S forms a subspace of V and is the smallest subspace of V to contain 5. We call this the subspace spanned by S and denote it by Span S. We say that 5 is a spanning set of V if Span S = V. A non-empty subset S of a vector space V is said to be linearly independent if the only way of writing 0V as a linear combination of elements of 5 is the trivial way, in which all the scalars are 0F. No linearly independent subset can contain 0V. A subset that is not linearly independent is said to be linearly dependent. A subset S is linearly dependent if and only if at least one element of S can be expressed as a linear combination of the other elements of S. A basis of a vector space V is a linearly independent subset that spans V. A nonempty subset 5 is a basis of V if and only if every element of V can be expressed in a unique way as a linear combination of elements of 5. If V is spanned by a finite subset S = {v,,..., v,,} and if / = {w,,..., wm] is a linearly independent subset of V then necessarily m^.n. From this there follows the important result that if V has a finite basis B then every basis of V is finite and has the same number of elements as B. A vector space V is said to be finite-dimensional if it has a finite basis. The number of elements in any basis of V is called the dimension of V and is written dim V. As a courtesy, we regard the empty set 0 as a basis of the zero subspace {0V} which we can then say has dimension 0. Every linearly independent subset of a finite-dimensional vector space V can be extended to form a basis of V. More precisely, if V is of dimension n and if is a linearly independent subset of V then there exist vm+,,..., v,, £ V such that {vi,...,vrilvlfI+ll...lv;,} is a basis of V. If V is of finite dimension then so is every subspace W, and dim W ¢. dim V with equality if and only if W = V. As for the morphisms, if V and W are vector spaces over the same field F then by a linear mapping from V to W we mean a mapping/: V -+ W such that (1) (V*,y€V) f(x + y)=f(x)+f(y); (2) (VA € F)(Vx € V) f{Xx) = Xf{x). These two conditions may be expressed as the single condition (W./iGFKVx.yGV) f{\x + ny)=\f(x) + nf{y). Additional properties that follow immediately from these axioms are: (3)/(0,) = 0^, (4) (V*€V) f(-x) = -f(x). If V and W are vector spaces over the same field F then the set Lin (V, W) of linear mappings from V to W can be made into a vector space over F in a natural way by

4 Further Linear Algebra defining/ + g and \f by the prescriptions (Vjc € V) (f + g)(x) = f(x) + g(x), (*/)(*) = \f(x). If / : V -+ W is linear and X is a subspace of V then the image of X under/, namely the set rW ={/(*) ;*€*}, is a subspace of W; and if Y is a subspace of W then the pre-image of Y under/, namely the set r(«={*eV;/Wer}. is a subspace of V. In particular, the subspace/""(V) = {f(x) ; jc e V} of W is called the image of/ and is written Im/; and the subspace f~{0w} of V is called the kernel of/ and is written Ker/. A linear mapping/: V —» W is injective if and onlyifKer/= {0V}. The main connection between Im/ and Ker/ is summarised in the Dimension Theorem which states that if V and IV are finite-dimensional vector spaces over the same field and if/: V —»W is linear then dim V = dim Im/ + dim Ker/. By the rank of a linear mapping/ we mean dim Im/; and by the nullity of/ we mean dim Ker/. A bijective linear mapping is called an isomorphism. If V and W are vector spaces each of dimension n over a field F and if/: V —» W is linear then the properties of being injective, surjective, bijective are equivalent, and/ is an isomorphism if and only if/ carries a basis to a basis. If V is a vector space of dimension n over a field F then V is isomorphic to the vector space Fn over F that consists of n-tuples of elements of F. As a consequence, if V and W are vector spaces of the same dimension n over a field F then V and W are isomorphic. Every linear mapping is completely and uniquely determined by its action on a basis. Thus, two linear mappings /, g : V —► W are equal if and only if they agree on any basis of V. Linear mappings from one finite-dimensional vector space to another can be represented by matrices. If F is a field and m, n are positive integers then by an m x n matrix over F we mean a rectangular array that consists of mn elements Xjj [i = 1,..., m; j = 1,..., n) of F in a boxed display of m rows and n columns: *1I *12 *I3 • *21 *22 *23 ■ *31 *32 *33 • *m\ xml xm3 • • *1,, ■• *2n ■■ *3,, ■ -*iiwi

5 We often find it convenient to abbreviate the above display to simply [x;j]mxn. The transpose of A = [aij)mxn is the matrix A1 = [a^. Matrices A = [ai}]mxn and B = [bi}]pxq are said to be equal if m = p, n = q and a{J = btj for all ij. With this definition of equality for matrices, the algebra of matrices may

be summarised as follows. Given m x n matrices A = [fly],,,,,,, and B ~ [bij]mxn over F, define A + B = [aiy + fc^,,; and for every \ € F define XA = [\aij]mxn. Then under these operations the set Matwxn Fofmxn matrices over F forms a vector space over F of dimension mn. The connection with linear mappings is described as follows. First an ordered basis of a vector space V is a finite sequence (vj,^, of elements of V such that (vi • • • ■. vn} 's a basis of V. For convenience, we use the abbreviated notation (vf)n. Suppose now that V and W are vector spaces of dimensions m and n respectively over a field F. Let (v,)^, (w.)n be given ordered bases of V, W. If / : V -► W is linear then/ is completely and uniquely determined by its action on the basis (vf)wl. This action is described by expressing each/(v;) as a linear combination of the basis elements wx,..., wn of W: /(v,) = xnwi + xnw2 + .. + xlnwn; /(V2) = ^21^1 + *22w2 + • • • + *2n*V> /(VJ = *«lWl + ^2^2 + *' * + Xmnwn- The action of/ on (v,)m is therefore determined by the mn scalars xi} appearing in the above equations. Put another way, the action of/ is completely determined by a knowledge of the m x n matrix X = [xiy]. For technical reasons, the transpose of this matrix X is called the matrix of/ relative to the fixed ordered bases (v(-)ffll (w,)n. When it is clear what these fixed ordered bases are, we denote the matrix in question by simply Mat /. If /•£ : V -» W are linear then relative to fixed ordered bases of each we have Mat if + g) = Mat / + Mat g, and Mat \f=\ Mat / for every scalar \. Hence the assignment/ •-» Mat/ gives an isomorphism from the vector space Lin(V, W) to the vector space Matwxn F. In the particular case where W = V the identity mapping idv corresponds under this isomorphism to the identity matrix /n = [6jj]nxn where 10 otherwise. Consider now the following situation: in which the notation U\ («,),, for example denotes a vector space U with a fixed ordered basis («.),, and/;/! denotes a linear mapping/ represented, relative to the

6 Further Linear Algebra ordered bases («,),,, (v,)n , by the matrix A. The situation concerning the composite linear mapping go/: U —» W is i/;W.—!l— wi(wt)r What is the matrix of this composite linear mapping? It is natural to expect that this will depend on the matrices A and B. In fact, with A = [ajj]nxm and B = [6y]/;XM (note the sizes!), we define the product BA to be the p x m matrix whose (ij')-th element is n [5A],7 = bnaij + ^a2y + ^/3^3/ + • • • + *>, ,,<*,,; = E bikakj- k=\ Then we have that Mat(g of) = BA = Mat g • Mat/. In other words, the matrix that represents g of is the product of the matrix that represents g with the matrix that represents/. When the relevant sums and

products are defined, we have the identities A(BC) = (Afl)C; A{B + C) = AB + AC\ (B + C)A = BA + CA. In particular, under the operations of addition and multiplication, the set Mat„xn F forms a ring with identity element /„. Also, multiplication by scalars is such that, when the relevant products are defined, \{AB) = {\A)B = A{\B). In particular, Matnxn F becomes what is called an algebra. Given a matrix A = [fl/yl^m over the field F, consider the subspace of Matnx, F that is spanned by the columns of A. The dimension of this subspace is called the column rank of A. Likewise, the dimension of the subspace of Mat, xm F spanned by the rows of A is called the row rank of A. Somewhat surprisingly, the column rank and the row rank of a matrix over a field are the same. (This is not true for matrices over a ring, for example.) So we can talk simply of the rank of a matrix over a field. Annxm matrix X is a left inverse of an m x n matrix M \iXM - /„; and a right inverse if MX = lm. An m x n matrix A/ has a left inverse if and only if rank M = n, and has a right inverse if and only if rank M - m. A square matrix M = \m,j\nxn has a left inverse X if and only if it has a right inverse Y; moreover in this case X = Y and so M has a unique inverse which we denote by M~l. We say that a (necessarily square) matrix is invertible if it has an inverse. An elementary matrix of size n x n is a matrix that is obtained from /„ by applying a single permutation to the rows or columns of /„. If E is an elementary n x n matrix then for every rxr matrix A the matrices EA and AE have the same rank as A. A square matrix is invertible if and only if it is a product of elementary matrices.

The story so far.... 7 If A and B are invertible then so is AB with [AB)~l = B~]A~l. If A is invertible then so is its transpose, with (A1)-1 = [A~1)'. A matrix is orthogonal if A1 exists and is A'. If V is a vector space over F of dimension n then a linear mapping/: V —» V is an isomorphism if and only if it is represented, relative to some ordered basis, by an invertible matrix. A very important question concerns what happens to the matrix of a linear mapping when we change reference from one ordered basis to another. Consider first the particular case of the identity mapping on an m-dimensional vector space V over F and the situation V;(v,). ^ .V;M). T T old basis new basis The matrix A that represents \6V is called the transition matrix from the basis (v,)„(to the basis (v,)m. Clearly, transition matrices are invertible. If now W is an //-dimensional vector space over F and if a linear mapping/ : V —» W is represented relative to ordered bases (v,)ml (vv()„ by the n x m matrix A, what is the matrix that represents/ relative to new ordered bases (vj)m, (h>J)„? To solve this problem, which is the substance of the Change of Basis Theorem, we consider the situation V;(v,)TO W.P V;«)TO w;K)n \dw.Q w;(na /* in which we have to determine the matrix X. Since, from the diagram, fo\6v=f- id,y o/ we deduce, via the isomorphism Lin(y, W) ~ MatnxmF, that AP = QX. Then, since the transition matrix Q is invertible, we have the solution X = Q~XAP. A direct consequence of this is that the rank of a linear mapping is the same as the rank of any matrix that represents it. In fact, let V, W be of dimensions m, n and let/: V -> W be of rank p. Then the Dimension Theorem gives dim Ker/ = m-p. Let {v,,..., vm_p] be a basis of Ker / and extend this to a basis B={u ,«p,v,,...,vm^} of V. Then {f(ux),... tf{up)} is linearly independent and so is a basis for Im/. Now extend this to a basis C = {/(«,),. •. ,f(up)t w,,..., wn_p] of W. Then the matrix of/ relative to the ordered bases B, C is 'p ° 0 0

Further Linear Algebra Suppose now that A is an n x m matrix that represents/ relative to fixed ordered bases BVt Bw. If Q and P are the transition matrices from the bases B, C to the bases BVt Bw then we have Q~lAP = *, ° 0 0 the matrix on the right being of rank p. But since transition matrices are invertible they are products of elementary matrices. Consequently the rank of A is the same as the rank of Q~lAPt namely p = rank /. If Ay B are n x n matrices over F then B is said to be similar to A if there is an invertible matrix P such that B = P~lAP. The relation of being similar is an equivalence relation on the set Matnxn F. Concerning similarity an important problem, from both the theoretical and practical points of view, is the determination of particularly simple representatives (or canonical forms) in each similarity class. The starting point of this investigation involves the notion of a diagonal matrix, this being a square matrix A = [fl/y]nxn for which a^ - 0 whenever i j^j. The problem of deciding when a given square matrix is similar to a diagonal matrix is equivalent to that of deciding when a linear mapping can be represented by a diagonal matrix. To tackle this problem, we need the following machinery which involves the theory of determinants. We shall not summarise this theory here, but refer the reader to Chapter 8 of our Basic Linear Algebra for a succinct treatment. If A is an n x n matrix then an eigenvalue of A is a scalar X for which there exists a non-zero n x 1 matrix x such that Ax = Xx. Such a (column) matrix x is called an eigenvector associated with X. A scalar X is an eigenvalue of A if and only if det (A - XI,,) = 0. Here det (A - Xln) is a polynomial of degree n in X, called the characteristic polynomial, and det (A - X/n) = 0 is the characteristic equation. The algebraic multiplicity of an eigenvalue X is the greatest integer k such that (X - X)k is a factor of the characteristic polynomial. The notion of characteristic polynomial can be defined for a linear mapping. Given a vector space V of dimension n over F and a linear mapping f : V —* V, let A be the matrix of/ relative to some fixed ordered basis of V. Then the matrix of / relative to any other ordered basis is of the form P~l AP where P is the transition matrix from the new basis to the old basis. Now the characteristic polynomial of P~lAP is tet{p-lAP -XIn) = det [P~l{A -XIn)P] = det/,-1det(A-X/n)det/> = det(A-X/n), i.e. we have Cp-iAP[X) - cA(X). It follows that the characteristic polynomial is independent of the choice of basis, so we can define the characteristic polynomial of / to be the characteristic polynomial of any matrix that represents/.

story so far.... 9 If X is an eigenvalue of the n x n matrix A then the set £x = {x €MatnxlF; Ax=Xx} i.e. the set of eigenvectors associated with the eigenvalue X together with the zero column 0, is a subspace of the vector space Matnx, F. This subspace is called the eigenspace associated with the eigenvalue X. The dimension of the eigenspace Ex is called the geometric multiplicity of the eigenvalue X. If/: V -» W is a linear mapping then a scalar X is said to be an eigenvalue of/ if there is a non-zero x^V such that f(x) = Xx, such an element x being called an eigenvector associated with X. The connection with matrices is as follows. Given annxn matrix A, consider the linear mapping /*:MatnxlF->MatnxlF given byfA(x) = Ax. Relative to the natural ordered basis of Matnx, F (given by the columns of /n), we have Mat fA - A. Clearly, the matrix A and the linear mapping fA have the same eigenvalues. Eigenvectors that correspond to distinct eigenvalues are linearly independent. A square matrix is said to be diagonalisable if it is similar to a diagonal matrix. Annxn matrix is diagonalisable if and only if it admits n linearly independent eigenvectors. A linear mapping/ : V —► V is said to be diagonalisable if there is an ordered basis (v,),, of V with repect to which the matrix of/ is a diagonal matrix; equivalently, if and only if V has a basis consisting of eigenvectors of /. A square matrix A [resp. a linear mapping f\ is diagonalisable if and only if, for every eigenvalue X, the geometric multiplicity of X coincides with its algebraic multiplicity. If a given matrix A is similar to a diagonal matrix D then there is an invertible matrix P such that P~XAP = D where the diagonal entries of D are the eigenvalues of A. The practical problem of determining such a matrix P is dealt with as follows. First we observe that the equation P~XAP - D can be written AP = PD. Let the columns of P be p,,..., pn and let where \,..., Xn are the eigenvalues of A. Comparing the i-th columns of each side of the equation AP = PD, we obtain (/=1,...,/1) APi^Xfo. In other words, the i-th column of P is an eigenvector of A corresponding to the eigenvalue \t. So, when a given nx n matrix A is diagonalisable, to determine an invertible matrix P that will reduce A to diagonal form (more precisely, such that

10 Further Linear Algebra P~XAP = diag{A,,..., \n}) we simply determine an eigenvector corresponding to each of the n eigenvalues then paste these eigenvectors together as the columns of the matrix P. Consider again the vector space Matnxn F which is of dimension n2. Every set of n2 + 1 elements of Matnxn F must be linearly dependent. In particular, given any A 6 Matnxn Ft the n2 + 1 powers A = /,,, At A , A , ...,j4 are linearly dependent and so there is a non-zero polynomial p(X) = a0 + axX + a2X2 + • ■ • + anlXni with coefficients in F such that p(A) = 0. The same is of course true for any f e Lin (y, y) where V is of dimension n, for we have Lin (y, V) ~ Mat nxn F. But a significantly better result holds: there is in fact a polynomial p(X) which is of degree at most n and such that p{A) = 0. This is the celebrated Cayley-Hamilton Theorem, the polynomial in question being the characteristic polynomial cA(X) = det{A-XIn)ofA. It follows from this that there is a unique monic polynomial mA (X) of least degree such that mA(A) - 0. This is called the minimum polynomial of A Both mA(X) and cA(X) have the same zeros, so mA[X) divides cA{X). In a similar way, we can define the notion of the minimum polynomial of a linear mapping, namely as the minimum polynomial of any matrix that represents the mapping.

1 Inner Product Spaces In some aspects of vector spaces the ground field F of scalars can be arbitrary. In this chapter, however, we shall restrict F to be R or C, for the results that we obtain depend heavily on the properties of these fields. Definition Let V be a vector space over C. By an inner product on V we shall mean a mapping / : V x V -» C, described by (x,y) •-> (x\y), such that for all jc, x',y 6 V and all a 6 C, the following identities hold: (1) (x + x'\y) = (x\y) + (x'\y)- (2) (ax\y)«a(x\y); (3) (x | y) - (y | jc) so that in particular (jc | jc) e IR; (4) (jc | jc) ^ 0, with equality if and only if jc = 0V. By a complex inner product space we mean a vector space V over C together with an inner product on V. By a real inner product space we mean a vector space V over R together with an inner product on V (this being defined as in the above but with the bar denoting complex conjugate omitted). By an inner product space we shall mean

either a complex inner product space or a real inner product space. There are other useful identities that follow immediately from (1) to (4) above, namely: (5) (*b + /> = (*W + My). In fact, by (1) and (3) we have (jc|>+y) = (y+y|x) = (7R+(7w = (x|>> + (x|y). (6) (x\ay) = a(x\y). This follows from (3) and (4) since (x | ay) = {ay \ x) = a(y \ x) = a(y \ x) = a(x \ y). (7) (jr|0,,) = 0=(0,,|jc). This is immediate from (1), (2), (5) on taking jc' = -x, / = -y, and a - -1.

12 Further Linear Algebra Example 1.1 On the vector space IR" of n-tuples of real numbers let n ({xu...fxn)\{ylt...tyn))^,£xiyr i=i Then it is readily verified that this defines an inner product on IR", called the standard inner product on IR". In the cases where n - 2,3 this inner product is often called the dot product or scalar product. This terminology is popular when dealing with the geometric applications of vectors. Indeed, several of the results that we shall establish will generalise familiar results in euclidean geometry of two and three dimensions. Example 1.2 On the vector space C of n-tuples of complex numbers let n (tei.--,^)IK,...,»0) = £v»£ i=i Then it is readily verified that this defines an inner product on C, called the standard inner product on C. Example 1.3 Let a, b € IR with a < b and let V be the real vector space of continuous functions /: [a, b] -* R Define a mapping from V x V to IR by V,8)»V\8) = Jfg. Then, by elementary properties of integrals, this defines an inner product on V. EXERCISES 1.1 Let R,,[X] be the real vector space of polynomials at degree most n. Prove that JO Jo defines an inner product on IR,,[X]. n 1.2 For a square matrix A = [a,y]nxn define its trace by tr A = £ an- Prove i=i that the vector space Matnxn IR can be made into a real inner product space by defining (A|fl) = trfl'A. Likewise, prove that Mat nxn C can be made into a complex inner product space by defining (A|fl) = trfl*A where B" - B' is the complex conjugate of the transpose of B.

1. Inner Product Spaces 13 Definition Let V be an inner product space. For every x 6 V we define the norm of x to be the non-negative real number For jc, y 6 V we define the distance between x and y to be <*(*.>)= I*-y|. Example 1.4 In the real inner product space IR2 under the standard inner product, if x = (x,, x2) then \\x\\2 - x] + jr|. so ||x|| is the distance from x to the origin. Likewise, if y = (>! ,y2) then we have lk->ll2 = (^,-yi)2 + (^->2)2, which is simply the theorem of Pythagoras. Recall from (4) above that we have W=o <=» * = <>,,. Theorem 1.1 Let V be an inner product space. Then, for all x, y 6 V and every scalar \, (1) |M|=N||*||; (2) [Cauchy-Schwarz inequality] \(x \ y)\ «£ \\x\\ \\y\\; (3)

[Triangle inequality] \\x + y|| ^ ||*|| + ||y||. Proof (1) ||Xx||2 = (Xx|U) = U(x|x) = |X|2|W|2. (2) The result is trivial if x = 0V. Suppose then that x f 0V, so that ||x|| f 0. Utz = y-^fx.Then IWI2 (z\x) = {y\x)-{j$(**) = <> and therefore 0ZUl2 = (z\i) = (z\y) from which (2) follows.

14 Further Linear Algebra (3) We have ||* + y||2 = (* + y|* + y) = (*|*) + (*|y) + (y|*) + (y|y) = |W|2 + (x|y) + (x|y)+|H|2 = |W|2 + 2Re(x|y)+|M|2 <IWI2 + 2|(x|y)|+|M|2 MIWI + IMI)2, from which the triangle equality follows immediately. D Example 1.5 In the inner product space C" under the standard inner product, the Cauchy-Schwarz inequality becomes n r rn E«*5<JEKI2nfelAI2- k=\ *=i V*=i Example 1.6 Let V be the set of sequences (a,), j> j of real numbers that are square-summable in the sense that £ a2 exists. Define an addition and a multiplication by real scalars in the component-wise manner, namely let (fl|).->i + (*.■).> l = («. + *,•).•> l, M<0,>i = (Hbi- Then it is readily verified that under these operations V becomes a real vector space. Now by the Cauchy-Schwarz inequality applied to the inner product space IR" with the standard inner product, we have the inequality n i=l n n e«a<,/i:«?i:*?. i=i i=i n n If now (pn)n>,, (qn)n> i 6 V are given by pn = £ a.fc, and ?n = . /£ a2 £ *>? then i=i y i=i i=i we have \pn | < q,, where {q,,)n> i converges to /£ a) E *?•Il foIlows that (#.),, * i yi>i i>i is absolutely convergent and hence is convergent. Thus £ 0,6, exists and we can define ((a,),^ | (*,),>i) = E «A- In this way, V becomes a real inner product space that is often called l2-space or Hilbert space.

1. Inner Product Spaces 15 EXERCISES 1.3 If V is a real inner product space and xt y 6 V, prove that ll^ + >l|2=IW|2+ll>|2 + 2(x|y). Interpret this result in IR2. 1.4 If W is a complex inner product space and jc, y 6 W, prove that ||* + y||2-/|K* + y||2= W|2+ ||y||2-r(W2+ ||y||2) + 2<*|y). 1.5 If V is an inner product space establish the parallelogram identity (V*,y€V) l^ + >l|2+ lk-y||2 = 2|W|2 + 2||y||2. Interpret this result in IR2. 1.6 If V is a real inner product space and x,y 6 V are such that \\x\\ = ||y||, show that {x + y \ x -y) - 0. Interpret this result in IR2. 1.7 Interpret the Cauchy-Schwarz inequality in the inner product space of Example 1.3. 1.8 Interpret the Cauchy-Schwarz inequality in the complex inner product space MatnxnC of Exercise 1.2. If, in this inner product space, Em denotes the matrix whose (p, ?)-th entry is 1 and all other entries are 0, determine (£w | £,,). 1.9 Show that equality holds in the Cauchy-Schwarz inequality if and only if {xty} is linearly dependent. Definition If V is an inner

product space then xt y 6 V are said to be orthogonal if (x | y) = 0. A non-empty subset S of V is said to be an orthogonal subset if every pair of distinct elements of 5 is orthogonal. An orthonormal subset of V is an orthogonal subset S such that I x || = 1 for every x £S. Example 1.7 Relative to the standard inner products, the standard bases of IR" and of C form orthonormal subsets. Example 1.8 The matrices Epq of Exercise 1.8 form an orthonormal subset of the complex inner product space Matnxn C.

16 Further Linear Algebra Example 1.9 In IR2 the elements x = (*,,x2) and y = (y,,y2) are orthogonal if and only if *,>', + x2y2 = 0. Geometrically, this is equivalent to saying that the lines joining x and y to the origin are mutually perpendicular. EXERCISES 1.10 In the inner product space R2[X] of Exercise 1.1» determine for which a 6 IR the subset {X, X2 - a] is orthogonal. 1.11 Let V be the inner product space of real continuous functions on the interval [-7r, n] with inner product vi «>=£/«• Prove that the subset S - {x t-* 1, x *-> sin kxt x i-» cos kx ; k - 1,2,3,...} is orthogonal. 1.12 In the inner product space Mat 2x2 C with (A\B) = tr B*A, is the subset ■{ i 0 0 0 1 0 i 0 0 1 0 0 i 0 1 0 0 0 i } orthonormal? Clearly, an orthonormal subset of V can always be constructed from an orthogonal subset S consisting of non-zero elements of V by the process of normalising x each element of 5, i.e. by replacing x 6 S by x* = -r —-. IMI An important property of orthonormal sets is the following. Theorem 1.2 In an inner product space orthogonal subsets of non-zero elements are linearly independent. Proof Let 5 be an orthogonal subset of V\ {0V} and let *,,..., xn 6 S. Suppose that n 52 *i = 0v- Then we have K 1*12 = M*.-1 *.■) = t **(** I *,) = (t V* I *.-) = (<V I *.-) = o, whence each X, = 0 since jc, ^ 0V. Hence S is linearly independent. □ We now describe the subspace that is spanned by an orthonormal subset.

1. Inner Product Spaces 17 Theorem 1.3 Let {C|,..., en] be an orthonormal subset of the inner product space V. Then [Bessel's inequality] (Vx 6 V) £|(jc |ek)\2 < ||.x||2. Moreover, if W is the subspace spanned by {eu..., en) then the following statements are equivalent: (1) xew- (2)E|(*M2=NI2; k=l n (3) * = EM **)**; k=l (4)(Vy€V) (*|y> = EW «*><«* I >)• Proof n Let z = x - Yl(x | £*)**. Then a simple calculation gives <*U> = <*1*)-E<*k>W«*>= M2-E| (*k*)f- (2) =>> (3) is now immediate from this since if (2) holds then z = 0V. (3) => (4): If x = £(x I **)** *en for all > 6 V we have <* I y) = (1> I «*>** I >) = £(* I **><«* I >) ■ (4) ^- (2) follows by taking y - x in (4). (3) =>{\) is

clear. n (1) => (3): If x = Y, ^kek men for7 = 1, • • •," we have h = E^l'i) = (E V* U) = (*|*,) whence (3) holds. D Definition By an orthonormal basis of an inner product space we mean an orthonormal subset that is a basis. Example 1.10 The standard bases of IR" and of C" are orthonormal bases.

18 Further Linear Algebra Example 1.11 In Matnxn C with (A \ B) = tr B*A an orthonormal basis is {Epq ; pt q = 1,..., n) where Epq has a 1 in the (p, q)-th position and 0 elsewhere. We shall now show that every finite-dimensional inner product space contains an orthonormal basis. In so doing, we give a practical method of constructing such a basis. Theorem 1.4 [Gram-Schmidt orthonormal isation process] Let V be an inner product space and for every non-zero x^V let x*- jc/||jc||. !f[xlt...txk} is a linearly independent subset of V, define recursively >i = *V> Then {>,,..., yk] is orthonormal and spans the same subspace as {*,,..., xk}. Proof Since xx f 0V it is clear that y{ f 0V. Then x2 - (^2 l>'i)>'i iS a non-trivial linear combination of x2,xl and hence y2 f 0V. In general, >', is a non-trivial linear combination of jc, ,..., jc, and consequently yt f 0V. It is clear from the definition of >, that x, is a linear combination of >,,...,yt. It follows therefore that {xlt...txk} and {>'!,••.,>'*} span the same subspace of V. It now suffices to prove that {>!,...,>*} »s an orthogonal subset; and this we do inductively. For k - 1 the result is trivial. Suppose then that {>,,... ,>r_i} is orthogonal where r > 2. Then, writing r-l = <*/. |*r-E(*r I >.-)>.■ II ,=1 we see from the definition of yr that r-l ar)7=*r-E(*r|y.-}y. 1=1

and so, fory < r, *rWyi) = My,)-Y;Myi)to\yj) i=i = Myj)-Myj) = o. Since or, f 0 we deduce that (yr\yj) = 0 fory < r. Consequently, {>,,... ,yr} is orthogonal. D Corollary If V is a finite-dimensional inner product space then V has an orthonormal basis. Proof Apply the Gram-Schmidt process to a basis of V. D Example 1.12 Consider the basis {jc1 , jc2, x3} of IR3 where *, = (0,1,1), x2 = (1,0,1), x3 = (1,1,0). In order to apply the Gram-Schmidt process to this basis using the standard inner product, we first let y, = *i/IWI = ^(0, i,i). Then ^2-(^1^)^ = 0.0.1)-^((1.0,1)1(0,1,1))^(0,1,1) = (1,0,1)-1(0,1,1) = 5(2.-1,1) so, normalising this, we take y2 = -^(2, -1,1). [Note that, by Theorem 1.1, we have in general This helps avoid unnecessary arithmetic in the calculations.] Similarly, *a - (*3 Myi - (*i \y\)y\ = |(i. i.-0 and so, normalising this, we takey3 = ^(1,1, -1). Thus an orthonormal basis constructed from the given basis is {#0.1.1). £(2.-1.0. #1.1.-0)-

20 Further Linear Algebra Theorem 1.5 If {ex,..., en} is an orthonormal basis of an inner product space V then *=i (2)(v*ev) W'-tK'k^r-. Jt=l (3)(to,yeV) (x\y) = t(*\'k)My)- *=i Proof This is immediate from Theorem 1.3. D Definition The identity (1) in Theorem 1.5 is often called the Fourier expansion of jc relative to the orthonormal basis {e,,..., £,,}, the scalars (x \ ek) being called the Fourier coefficients of x. The identity (3) of Theorem 1.5 is known as Parseval's identity. Example 1.13 The Fourier coefficients of (1,1,1) relative to the orthonormal basis {y,, v2, y 3} constructed in the previous example are ((1,1,1)1^(0,1,1)) = ^; ((1,1,1)1^(2,-1,1)) = $; ((1,1,1)1^(1,1,-1)) = ^. Just as every linearly independent subset can be extended to a basis, so can every orthonormal subset be extended to an orthonormal basis. Theorem 1.6 Let V be an inner product space of dimension n. If {xlt...txk} is an orthonormal subset of V then there exist xk+l,..., xn 6 V such that {*,,..., x,,} is an orthonormal basis of V. Proof Let W be the subspace spanned by {jc, ,..., xk}. By Theorem 1.2, this set is a basis for W and so can be extended to a basis {x^,..., xkl xk+],..., x,,] of V. Applying the Gram-Schmidt process to this basis, we obtain an orthonormal basis for V. The result now follows on noting that the first k elements of this orthonormal basis are precisely jclt..., xk\ for, by the formulae in Theorem 1.4 and the orthogonality of jc,,...,^, we see that y{, = jc*= jc, foreachi<*. □

Product Spaces 21 1.13 Use the Gram-Schmidt process to construct in IR3 an orthonormal basis from the basis {(1,1,1,),(0,1,1),(0,0,1)}. 1.14 Determine an orthonormal basis for the subspace of IR4 spanned by {(1,1,0,1),(1,-2,0,0), (1,0,-1,2)}. 1.15 Determine an orthonormal basis for the subspace of C3 spanned by {(1,/,0),(1,2,1-0}. 1.16 Consider the real inner product space of real continuous functions on the interval [0,1] under the inner product •l /«• (f\ 8)=/. Jo Find an orthonormal basis for the subspace spanned by {/i,/2} where /1W=land/2(x) = x. 1.17 In the inner product space IR[X] with inner product (P*)-Jf determine an orthonormal basis for the subspace spanned by (1) {x,*2}; (2) {1.2X-1.12X2}. 1.18 Let V be a real inner product space. Given xty 6 V define the angle between x and y to be the real number t? 6 [0, -n] such that COStf= <*!>> WW Prove that, in the real inner product space IR2, ParsevaPs identity gives cos(tf, -1?2) = cos t?! cos t?2 + sin ^l sin t?2- 1.19 Referring to Exercise 1.11, consider the orthogonal subset 5n= {jch-> 1, xt- *s\nkxt x i-> coskx ; /:= 1,...,n}. If Tn is the orthonormal subset obtained by normalising every element of 5n, interpret Bessel's inequality relative to Tn.

Determine, relative to r,,, the Fourier expansion of (1) x~x2; (2) x i-» sin x + cos x.

22 Further Linear Algebra 1.20 Show that *-«A.o.^). (-^.0. Jj).(o.'.«} is an orthonormal basis of R3. Express the vector (2, -3,1) as a linear combination of elements of B. 1.21 Let V be a real inner product space and let/ : V -* V be linear. Prove that if B = {e,,..., e,,] is an ordered orthonormal basis of V and if M is the matrix of/ relative to B then 1.22 Let V be a real inner product space and let A = {a,,..., a,} be a subset of V. Define CA to be the r x r matrix whose (i j)-th element is Prove that a,,..., flr are linearly dependent if and only if det GA = 0. An isomorphism from one finite-dimensional vector space to another carries bases to bases. We now investigate the corresponding situation for inner product spaces. Definition If V and W are inner product spaces over the same field then /: V -» W is an inner product isomorphism if it is a vector space isomorphism that preserves inner products, in the sense that (Vx.yeV) <f{x)\f{y)) = (x\y). Theorem 1.7 Let V and W be finite-dimensional inner product spaces over the same field. Let {e{,..., en) be an orthonormal basis of V. Then f : V -+ W is an inner product isomorphism if and only if {/(e,),... J{en)} is an orthonormal basis of W. Proof =>: If/: V -* W is an inner product isomorphism then clearly {/"(e,),... ,/(e#l)} is a basis of W. It is also orthonormal since fl iff =J; W«i)l/l«y))«tel«yH l0 if/ +j. <^ : Suppose now that {/(e,),...,/(e,,)} is an orthonormal basis of W. Then/ carries a basis of V to a basis of IV and so/ is a vector space isomorphism. Now for

1. Inner Product Spaces 23 all x G V we have, using the Fourier expansion of x relative to the orthonormal basis (fM\f{ej))^(f(t^\ei)e)\f(ej)) = (EW«iW«j|/fy)) = t(*\'iM'i)\A'j)) i=i and similarly (/(ey) |/(x)) = (ey | x). It now follows by ParsevaPs identity applied to both V and W that OWI/M) = £</W \AejM'j) l/W) = EM';><*; W and consequently / is an inner product isomorphism. D EXERCISES 1.23 Let V and W be finite-dimensional inner product spaces over the same field with dim V ~ dim W. If/ : V -* W is linear, prove that the following statements are equivalent: (1) / is an inner product isomorphism; (2) / preserves norms, i.e. (Vx,G V) \\f{x)\\ = ||x||; (3) / preserves distances, i.e. (Vjc, y G V) ^(/"(x),/^)) = </(*, >). 1.24 Let V be the infinite-dimensional vector space of real continuous functions on the interval [0,1]. Let Vl consist of V with the inner product <fl*)= ff(x)g(x)x2dxt Jo and let V2 consist of V with the inner product <f|*)= ff(x)g(x)dx. Jo Prove that the mapping p : Vx -> V2

given by/ *-> <p{f) where (VxGV) fo>(/)lM = jc/W is linear and preserves inner products, but is not an inner product isomorphism.

2 Direct Sums of Subspaces are as If A and B are non-empty subsets of a vector space V over a Meld F then the subspace spanned byAuB, i.e. the smallest subspace of V that contains both A and B, is the set of linear combinations of elements ofAUB. In other words, it is the set of elements of the form m n EMi + E^A i=l /=1 where a, E A, bj E #. and \,/iy- E /7. In the particular case where A and B i m n subspaces of V we have £ X,a, E A and J) /iyfy E 5, so this set can be described i=l /=1 {a+fc; flE^.fcE^}. We call this the sum of the subspaces A, B and denote it by A + B. This operation of addition on subspaces is clearly commutative. More generally, if Ax,..., An are n subspaces of V then we define their sum to be the subspace that is spanned by (J A,. i=i n We denote this subspace by j^ + • • • + An or E^i» noting that the operation of i=l addition on subspaces is also associative. Clearly, we have n E^i = {ai+ - + 0«; o/eaj. i=i Example 2.1 Let X, K, D be the subspaces of IR2 given by *={(*,()); XEIR}, K={(0,y);yElR}, D= {(x,x); xElR}. Since every (x,y) E IR2 can be written in each of the three ways (x,0) + (0,y), (x-y,0) + (y,y), (0,y-x) + (x,x), wesee that IR2 = X + y = X+ D= y+ D.

2. Direct Sums of Subspaces 25 Definition « n A sum J]) Aj of subspaces Ax,..., AH is said to be direct if every x6^A can be i=l ,=i written in a unique way as x - ax + • • • + an with at 6 A§ for each i. n n When the sum £ ^ is direct we shall write it as 0 A,or Ax © • • • © An and call i=i i=i this the direct sum of the subspaces Alt...,An. Example 2.2 In Example 2.1 we have B2 = X@Y = X®D=Y®D. Example 2.3 In IR3 consider the subspaces A={(x,y,z)'> Jt + y + Z = 0}, fl={(x,jt,z); jc.zgIR}. Then IR3 = A + B since, for example, we can write each (x, y, z) as Q(* ~y)>-J(* ->),0) + (|(* + y), |(x + y),z) which is of the form a + b where a 6 ^ and b 6 #• The sum A + 5 is not direct, however, since such an expression is not unique. To see this, observe that (x.y.z) can also be expressed as (i(x-y+ 1),4(^-^-1),-1) + (^ + ^-1),^ +).-1),2 + 1) which is of the form a' + b' with a' e A, b' e B and a' + a,b' + b. EXERCISES 2.1 Let V be the subspace of IR3 given by V= {(*,*,0); xelR}. Find distinct subspaces I/,, U2 of IR3 such that IR3 = V © Ux = V © t/2. 2.2 Let f J, f2, r3, f 4 : IR3 -»IR3 be the linear mappings given by *i(*,y,z)= {x + y,y + ztz + x)\ h(x>y>z)= {x-yty-z,o)\ ^.^.2)= (-y.*i*); ^,y,z) = (*,y,y)- Prove that, for« = 1,2,3,4, IR3 = Imf,© Kerf,-.

26 Further Linear Algebra A useful criterion for a sum to be direct is provided by the following result. Theorem 2.1 If Au...,Anare subspaces of a vector space V then the following statements are equivalent: n (1) the sum £ A, is direct; n (2) if £ a, = 0 with a, 6 A, then each a, = 0; i=l (3) A^YsAj^Wv) far every i. if' Proof (1) =*• (2): By the definition of direct sum, if (1) holds then 0V can be written n in precisely one way as a sum £ a, with a, 6 Aj for each i. Condition (2) follows i=l immediately from this. (2) => (3): Suppose that (2) holds and let x e A,n£ Ay, say x - a{f = £ a;. We can write this as a{ - £ aj = 0 whence, by (2), each ai = 0 and therefore x - 0V. if n n (3) => (1): Suppose now that (3) holds and that £ a, = £ ^i wnere fl/i ^/ G A,- i=i i=i for each i. Then we have a*-*.• = £(*/-fly) where the left hand side belongs to A, and the right hand side belongs to]T /4,. By (3), we deduce that at - 6, = 0 and, since this holds for each /,(1) follows. D Corollary V= Al®A2ifandonlyifV = Al+A2and A, HA2 = {0V}. D Example 2.4 A mapping/: R -> R is said to be even if/(-x) = f[x) for every x 6 R; and odd if/(-x) = -f{x) for every x 6 IR- The set A of even functions is a subspace of the vector space V = Map(R, R), as is the set B of odd functions. Moreover, V - A© B. To see this, given any/: IR -» R let/*: R -»IR and/" : R -»IR be given by Then/* is even and/- is odd. Since cleaiiy/ = f* +f~ we have V = A + fl. Now AHB consists of only the zero mapping and so, by the above corollary, we have that V=A@B.

Sums of Subspaces 27 Example 2.5 Let V be the vector space Matnxn IR. If A, B are respectively the subspaces of V that consist of the symmetric and skew-symmetric matrices then V = A © B. For, every M G V can be written uniquely in the form X + Y where X is symmetric and Y is skew-symmetric; as is readily seen, X - \{M + M') and Y = \(M -M'). EXERCISES 2.3 If V is a finite-dimensional vector space and if A is a non-zero subspace of V prove that there exists a subspace B of V such that V = A © £. 2.4 If X = {(x,0) ; x 6 IR} find infinitely many subspaces Z, such that IR2 = X©Z,. 2.5 If V is a finite-dimensional vector space and A, B are subspaces of V such that V - A + B and dim y = dim A + dim B, prove that y = A 0 B. 2.6 Let V be a finite-dimensional vector space and let/: y —* V be a linear mapping. Prove that V = Im/ © Ker/ if and only if Im/ = Im/2. 2.7 If y is a finite-dimensional vector space and/: V —» y is a linear mapping, prove that there is a positive integer p such that Imf = Im//>+l and deduce that (VA:^ 1) Im/' = Im/***, Kerf = Ker/*"*. Show also that V = Im/' © Ker/*\ A significant property of direct sums is the following, which says informally that bases of the summands can be pasted together to obtain a basis for their sum. Theorem 2.2 Let Vl,...,Vn be non-zero subspaces of a finite-dimensional vector space V such that y = 0 V;. If B{ is a basis of V{ then (j Bk is a basis of V. Proof For each i, let the subspace Vt have dimensiondt and basis B, = {bilt...,bid). Since V = 0 V^? we have Vi,0 £ Vj = {0V} and therefore Vt, n V}, = {0V} for i fj. Consequently B, n By = 0 for / ^y. Now a typical element of the subspace spanned n by (J B, is of the form i=i 0) EVu + - + EV-.; i=i y=i

28 Further Linear Algebra i.e. of the form (2) ^ + --- + ¾ where ^ = 5^,,6,,,. n Since V = £ V] ar,d since B, is a basis of V, it is clear that every x G V can be ;=i n expressed in the form (1), and so V is spanned by |J fl,. If now in (2) we have ;=i jci+- -+xn - 0V then by Theorem 2.1 we deduce that each jc, = 0V and consequently n that each >iy = 0. Thus |J Bt is a basis of V. D i=l Corollary dim®V;. = £>mV;, D i=\ i=l EXERCISES 2.8 Let V be a real vector space of dimension 4 and let B - {&,, 62, 63,64) 4 be a basis of V. With each x£V expressed as x = £ jf ,-6,-, let i=i y2 = {x G V; *3 = -x2l x4 = -¾}. Show that (1) 1^ and V2 are subspaces of V\ (2) Bx - {6^64,62+63}isabasisof^,,andB2 - {62-63,6,-64} is a basis of V2; (3) v=v,©v2; (4) the transition matrix from the basis B to the basis fl, U fi2 's /> = rj 0 0 11 oiio 0 1 -1 0 li 0 0 4, (5) /»-' = 2/>. A real 4 x 4 matrix A/ is said to be centro-symmetric if m/y = m5_, 5_, for all ij. If M is centro-symmetric, prove that A/ is similar to a matrix of the form 'a)3 0 0" i no 0 0 £ < 0 0 7} t?_

2. Direct Sums of Subspaces 29 2.9 Let V be a vector space of dimension n over IR. If/ : and such that/2 = \6V prove that V=lm{\6v+f)@lm{\dv-f). Deduce that annxn matrix A over IR is such that A2 - A is similar to a matrix of the form '/, o " .° -'"-p.' Our purpose now is to determine precisely when a vector space is the direct sum of finitely many non-zero subspaces. As we shall see, this is closely related with the following types of linear mapping. Definition Let A and B be subspaces of a vector space V such that V - A © B. Then every x 6 V can be expressed uniquely in the form x = a + b where a 6 A and b £ B. By the projection on A parallel to B we shall mean the linear mapping pA : V —» V given byPaUH0- Example 2.6 We know that IR2 = X © D where X = {(x,0); x e H} and D = {(*, x); x e IR}- The projection on X parallel to D is given by p(x,y) = [x -yt 0). Thus the image of the point (xty) is the point of intersection with X of the line through [xt y) parallel to the line D. The terminology used is thus suggested by the geometry. EXERCISE 2.10 In

IR3 let A = Span{(l,0,1),(-1,1,2)}. Determine the projection of (1,2,1) onto A parallel to Span{(0,1,0)}. Definition A linear mapping/: V -► V is said to be a projection if there are subspaces A, B of V such that V = A © B and/ is the projection on A parallel to B. A linear mapping / : V -» V is said to be idempotent if/2 = /. These notions are related as follows. Theorem 2.3 lfV- A®B and if p is the projection on A parallel to B then (1) A = Imp={xey; x = p{x)}\ (2) B = Kerp; (3) p is idempotent. V —» V is linear I,, if and only if

30 Further Linear Algebra Proof (1) It is clear that A = Imp D {x 6 V; x = p(jf)}. If now a 6 A then its unique representation as a sum of an element of A and an element of B is clearly a = a + 0V. Consequently p(a) = a and the containment becomes equality. (2) Let x 6 V have the unique representation x - a + b where o £ A and b £ fl. Then since p(x) = a we have p(x) = (V <=» a = 0v <f=> x = beB. In other words, Kerp = B. (3) For every x 6 ^ we have p(x) 6 A and so, by (1), p(x) = p{p(x)]. Thus P2 = P. □ That projections and idempotents are essentially the same is the substance of the following result. Theorem 2.4 A linear mapping f :V —*Visa projection if and only if it is idempotent, in which case V = Im/ © Ker/ and f is the projection on Im/ parallel to Ker/. Proof Suppose that/ is a projection. Then there exist subspaces A and B with V= AQ)B, and/ is the projection on A parallel to B. By Theorem 2.3(3),/ is idempotent. Conversely, suppose that/ : V —» V is idempotent. If x £ Im/n Ker/ then we have x = /(>) for some y, and /(*) = 0V. Consequently, x = /(>) = f[f{y)] = /(x) = 0V and hence Im/ 0 Ker/ = {0y}. Now for every x £ V we observe that /I* -/(*)]=/(*) -/2M=/(*) -/(*) = ov and so x - /(x) £ Ker/. The identity x = f(x)+x -f[x) now shows that we also have V - Im/ + Ker/. It follows by the Corollary of Theorem 2.1 that V = Im/© Ker/. Suppose now that x = a + b where a £ Im/ and b £ Ker/. Then a = f(y) for some >, and/(6) = 0y. Consequently, /M = f(a + 6) = /(a) + (V = /[/(y)] = f(y) = a. Thus we see that / is the projection on Im/ parallel to Ker/. □ Corollary f// • V- *Visa projection then so is \6V -f, and Im/ = Ker(idy -/). Proof Since/2 = /we have (idy -/)2 = idy -f-f+f2 - idy -/. Moreover, by Theorem 2.3, we have xelmf <=* *=/(*) <=* (i<V-/)W = <V and so Im/ = Ker(idv -/). □

Sums of Subspaces 31 EXERCISES 2.11 If/ is the projection on A parallel to B, prove that idv -/ is the projection on B parallel to A. 2.12 Let V be a vector space over IR. If p, ,p2: V —» V are projections prove that the following are equivalent: 0) Pi + Pi is a projection; (2) Pl°P2 = P2°Pl=0- When Pi + p2 is a projection, prove that Im^ + p2) = Imp, © Imp2 and Ker(pj + p2) = Kerp, n Kerp2- 2.13 If V is a finite-dimensional vector space and if php2 : V —» V are projections, prove that the following statements are equivalent: (1) Imp, = Imp2; (2) Piop2 = p2andp2op1=p1. 2.14 Let V be a finite-dimensional vector space and let p,,..., pk: V —» V be projections such that Imp, = Imp2 = • • • = Imp*. Let \,,..., \k G F k be such that £ \ - 1 • Prove that p = \ xpx + • • + A^ is a projection i=i with Imp = Imp;. 2.15 Let y be a vector space over IR and let/: V —» y be idempotent. Prove that idy +/ is invertible and determine its inverse. We now show how the decomposition of a vector space into a direct sum of finitely many non-zero subspaces may be expressed in terms of projections. Theorem 2.5 If V is a vector space then there are non-zero subspaces Vu...tVnof V such that n V = 0 Vj if and only if there are non-zero linear mappings p,,..., pn : V -» V such f=i that (1) J>,-id,,; (2) (ifflPiopj-O. Moreover, each p,- is necessarily a projection and Vt = Imp,. Proof Suppose first thatl/=0V;-. Then for/= 1,...,n we have ^= V;©£V}. Letp^ be i=i it* the projection on 1^ parallel to J2 Vj. Using the notation pf (X) = {pM ; x 6 X} it'

 for every subspace X of V, we have that forj f /, PilPjM] € Pp(Impy) = pViVj) by Theorem 2.3 = pf (KerpJ by Theorem 2.4 = {0,} and so p,- o py = 0. Also, since every x 6 V can be written uniquely in the form n * = E * i where x, 6 ^- for each i, we have /i /i « * = E*» = Ea-W = lEft/M i=l i=l f=l n whence Y, Pi~ l^v i=l Conversely, suppose that p,,... ,pn satisfy (1) and (2). Then we note that n n A = ft o idy = P, o Y,Pj = E(ft °P;) = ft °A- and so each p, is idempotent and therefore, by Theorem 2.4, is a projection. Now for every x 6 V we have /i * /i x = idyM = (J»(x) = EftM € EIm ft i=l i=l i=l n which shows that V = £ Imp;. i=l If now x 6 Imp, D J^Impy then, by Theorem 2.3, we have x = p,(x) and W x-J^Xj where py(xy) = Xj for every j f i. Consequently, it* x = p,(x) = p,(E *y) = ft(EftW) = Eftlftt*)] = °v tf' ;*« w n and it follows that V7 = ® Imp,. D Examp/e2.7 Consider the direct sum decomposition IR2 = X © Y where X is the 'x-axis* and Y is the 'y-axis*. The projection on X parallel to Y is given by px(x, y) = (x, 0), and the projection on Y parallel to X is given by py[xty) = (0,y). We have px+Py = id, Px°Pr = Q = Pr°Px-

 We now pass to the consideration of direct sums of subspaces in inner product spaces. Although we shall take a closer look at this later, there is a particular direct sum of subspaces that we can deal with immediately. For this purpose, consider the following notion. Definition Let V be an inner product space. For every non-empty subset £ of V we define the orthogonal complement of £ to be the set EL of elements of V that are orthogonal to every element of £; in symbols, E^^iyeV; (V*€£) (*|y) = 0). It is readily verified that E1 is a subspace of V. Moreover, we have V1 = {0V} and {(V} 1 = V. The above terminology is suggested by the following result. Theorem 2.6 Let V be an inner product space and let W be a finite-dimensional subspace of V. Then Proof Since W is of finite dimension there exists, by the Corollary of Theorem 1.4, an n orthonormal basis of W, say {eit...ten}. Given x 6 V, let x' - £{x | efte-, 6 W and i=l consider the element x" = x - x'. Forj = 1,..., n we have (x» I e,) = (x |;) - (V |«,) = (x |,,) - £(* |«,)(«, | et) 1=1 ■(*l«i)-(*l«i) = 0. It follows that x" e W1 and hence that x - x' + x" € W + W1. Consequently we have V = W + W1. Now if x e W fl W1 we have (x \ x) = 0 whence x = 0^. Thus W n W1 = {0V} and we conclude that V = W © W1. □ The principal properties of orthogonal complements are contained in the following two results. Theorem 2.7 If V is a finite-dimensional inner product space and if W is a subspace of V then dimW-L = dimV-dimW. Moreover, (W1)1 = W.

 Proof It is immediate from Theorem 2.6 that from which the first statement follows. As for the second, it is clear from the definition of orthogonal complement that we have W C (W1)1. Also, by what we have just proved, dim(WJ-)-L = dim^-dimWJ-. Hence dim(WJ-)-L = dim W and the result follows. D Theorem 2.8 If V is a finite-dimensional inner product space and if A,B are subspaces of V then (1) ACB => B±CA±\ (2) (AnfiJ-^A-L + fl1; (3) {A + B^^A^nB1. Proof (1) If A C B then clearly every element of V that is orthogonal to B is also orthogonal to A, so B1 C A1. (2) Since A, B C A + B we have, by (1), (A + B)-LCA-Lnfi-L. Similarly, since A n B C A, B we have A1, B1 C (A 0 fi)1 and so A-L + B-LC(AnB)-L. Together with Theorem 2.7, these observations give AnB=(AnB)-L-LC(A-L + B-L)-LCAJ"LnBJ-L = AnB from which we deduce that AflB=(A-L + B-L)-L. Consequently (A n B)1 = (A1 + B1)1-1- = A1 + B1. (3) This follows from (2) on replacing A, B by A1, B1. D Example 2.8 Consider the subspace W of IR3 given by W = Span{(0,1,1),(1,0,1)}. In order to determine W1 we can proceed as follows. First extend {(0,1,1), (1,0,1)} to the basis {(0,1,1), (1,0,1), (1,1,0)} of IR3. Now apply the Gram-Schmidt process to obtain (as in Example 1.12) the orthonormal basis {^(o.'. i). :75(2.-1,1).^0,1,-1))-

2. Direct Sums of Subspaces 35 Recall from Theorem 1.4 that W = Span{^(0,1,1), ^(2, -1,1)}. It follows that IV1- Span {(1,1,-1)}. EXERCISE 2.16 Determine the orthogonal complement in IR3 of the subspace Span{(l,0,1),(1,2,-2)}. Suppose now that W is a finite-dimensional subspace of an inner product space V. By Theorem 2.6, given x £ V we can express x uniquely in the form x = a+b where a £ W and b £ W1. Then, by orthogonality, ||*||2 = (<. + fc|a+i)= Ml 2+||fc||2. It follows that, for every > 6 W, |x->||2=|<.-y + fc||2 = |a-,||2+||fc||2 + 2(a-,|fc) = 1"-y||2+W2 = 1"-y|2+11*-«112 > I*-all2- Thus we see that the element of W that is 'nearest' the element of x of V is the component a of x in W. Now let {ex,..., en) be an orthonormal basis of W. Express the element a £ W that is nearest a given x E V as the linear combination n /= i By Theorem 1.3, we have \ = (a \ e) and by orthogonality n Thus we see that the element of W that is nearest to x is £(x |«,.)«,-, the scalars being the Fourier coefficients. Example 2.9 Let us apply the above observations to the inner product space V of continuous functions/ : [0,2-n] —► IR under the inner product {f\8)= / /*• Jo An orthonormal subset of V is S=ix>-* vb« *•""> ^cos^jc, jci-» ^-sinfo; /:= 1,2,3,...}.

36 __ Further Linear Algebra Let Wn be the subspace, of dimension In + 1, with orthonormal basis B"~ {**-* ^7» Xi~* 7*cos**' Xi-* j;s'mkx* k~ li"-w}- Then the element/, of Wn that is nearest a givenf £ V (or, put another way, is the best approximation to/ in Wn) is n fnb) = c04- + J>* cos *x + dk sin *x) v *=i where c0 = (/(*) | ^) and (V^l) c,= (/(x)|^cos^), ^*{fljc)|j5-sinfac). This can be written in the standard form n /,,(*) = ^0 + E(fl* C0S kx + fc* Sin *X) where a0 = J / f(x)dx and ./o /»2w y»2ir (*^1) fl*=W /(x)cos*x<fjr; fcft=i/ /(jt)sin*jt</jt. ./o JO If/ is infinitely differentiable then it can be shown that the sequence (/",,),,>i is a Cauchy sequence having/ as its limit. Thus we can write / = ja0 + £ (aA cos fo + 6* sin fo) k>\ which is the Fourier series representation of/.

3 Primary Decomposition The characterisation of direct sums given in Theorem 2.5 opens the door to a deep study of linear mappings on finite-dimensional vector spaces and their representation by matrices. In order to embark on this, we require the following notion. Definition If V is a vector space over a field F and if / : V -» V is linear then a subspace W of V is said to be /-invariant (or/-stable) if it satisfies the property /-(W) C W, i.e. ifx £W=>f{x)£W. Example 3.1 If/: V -* V is linear then Im/ and Ker/ are /-

invariant subspaces of V. Example 3.2 Let D be the differentiation map on the vector space R[X] of all real polynomials. Then D is linear and the subspace f\n[X] consisting of the polynomials of degree at most n is D-invariant. Example 3.3 Iff : y _> v is linear and x £ V with x f Oy then the subspace spanned by {*} is /-invariant if and only if x is an eigenvector of/. In fact, the subspace spanned by {x\ is Fx = {\ x ; X € F) and this is/-invariant if and only if for every \ £ F there exists fi 6 ^ such that/(Xx) = fix. Taking \ = lf we see that x is an eigenvector of/. Conversely, if x is an eigenvector of/ then/(x) = fix for some /i G F and so, for every \ G F, we have/(Xx) = X/(*) = V* € ^- Example 3.4 Let/ : V -* V be linear. Then for every p G F[X] the subspace Kerp(/) of V is /-invariant.

3B Further Linear Algebra EXERCISES 3.1 Let/: Fl4 -> IR4 be given by /(a, 6, c, J) = (a + b + 2c -d, b + J, 6 + c, 26 - </)• Show that the subspace W = {*, 0, z, 0); jc, z 6 IR} is/-invariant. 3.2 Let V be the vector space of real continuous functions defined on the interval [0,1]. Let <p : V -» V be given by / •-» p(/) where (V*GV) MfflM-*/"/. Jo Let W be the subspace consisting of those/ £ y of the form/(x) = ax+6 for some a, 6 6 IR. Prove that W is ^-invariant. 3.3 Let V be a finite-dimensional vector space and let p be a projection. If f:V-tV\s linear, prove that (1) Imp is /-invariant if and only if p of o p = / o p; (2) Kerp is /-invariant if and only if p o/ o p- pof. Deduce that Imp and Kerp are /-invariant if and only if p and/ commute. 3.4 Let V be a finite-dimensional vector space and let/, g : V -► V be linear mappings such that/o g = idy. Prove that gof- \6V. Prove also that a subspace of V is /-invariant if and only if it is g-invariant. Does this hold if V is infinite-dimensional? 3.5 Let V be a finite-dimensional vector space and let/: V -► V be linear. n Suppose that y = 0 Vs where each V; is an/-invariant subspace of V. If/ : V; -► V is the restriction of/ to V;, prove that (fl)Im/=©Im/.; i=i (6)Ker/=©Ker/, i=i Theorem 3.1 Iff : y -► V is linear then for every polynomial p 6 F[X] the subspaces Im p{f) and Kerp(/) are f-invariant. Proof Observe first that for every polynomial p we have/ o p(/) = p(/) o/. It follows from this that if x - p{f){y) then /(jc) = p(/)[/"(y)], whence Imp(jQ is/-invariant. Likewise, if p(f){x) = 0y then p(/) [^(x)] = 0Vt whence Kerp(/) is /-invariant (c.f. Exercise 3.3). □

sition 39 In what follows we shall often have occasion to deal with expressions of the form p(f) where p is a polynomial and/ is a linear mapping, and in so doing we shall find it convenient to denote composites by simple juxtaposition. Thus, for example, we shall write/pC/) for/op(/). Suppose now that V is of finite dimension n, that/ :V -+V\s linear, and that the subspace W of Vis/-in variant. Then/induces a linear mapping/: W —*W, namely that given by the assignment w *-> f(w) - f(w) 6 W. Choose a basis {w,,..., wr} of IV and extend it to a basis of V. Consider the matrix of/ relative to this basis. Since W is/-invariant each /(w{) 6 W and so, for each it fiwg) = \nwi + •. • + \irwr + 0 Vl + • • • + 0 vn_,, whence it follows that this matrix is of the form A B 0 C where A is the r x r matrix that represents the mapping induced on W by/. Suppose now that V - Vi®V2 where both V{ and V2 are/-invariant subspaces of V. If Bx is a basis of V, and B2 is a basis of V2 then by Theorem 2.2 we have that B = B, U B2 is a basis of V, and it is readily seen that the matrix of/ relative to B is of the form A, 0" 0 A2 where A1(A2 represent the mappings induced on Vlt V2 by/. k More generally, if V - 0 V, where each V; is/-invariant and if B, is a basis of V, i=i it for each i then the matrix of/ relative to the basis B-\JB{ is of the block diagonal i=i form "A, in which A, is the matrix representing the mapping induced on V, by/, so that A, is of size dim V(x dim Vr Our objective now is to use the notions of direct sum and invariant subspace in order to find a basis of V such that the matrix of/ relative to this basis has a particularly useful form. The key to this study is the following fundamental result.

40 Further Linear Algebra Theorem 3.2 [Primary Decomposition Theorem] Let V bea non-zero finite-dimensional vector space over afield F and let f: V —» V be linear. Let the characteristic and minimum polynomials of f be respectively, where plt...tpkare distinct irreducible polynomials in F[X]. Then k each of the subspaces Vt = Kerp'*(/) is f-invariant and V-QVj. i=i Moreover, if f : Vt —> Vj is the linear mapping that is induced on V{ by f then the characteristic polynomial of £ is pt' and the minimum polynomial off isp*'. Proof If k = 1 then p\'{f) = mf{f) = 0 and trivially V = Ker p\x{f) = Vx. Suppose then thatk^2. Fori = 1,..., Jklet ft-«//#? "IK- it* Then there is no irreducible factor that is common to each of qx,..., qk and so there exist a,,..., ak € F[X] such that alql+a2q2 + --- + akqk=\. Writing qtat - r, for each i, and substituting/ in this polynomial identity, we obtain (1) f,(/) + . .. + f4tf)-id,,. Now by the definition of ¢, we have that if i f) then mf divides q{qr Consequently 1,{f)<ljif) - 0 for i £j% and then (2) (ifj) f.-tffyW-O. By (1), (2), and Theorem 2.5 we see that each f,-(/) is a projection and i=\ Moreover, by Theorem 3.1, each of the subspaces Im t,[f) is/-invariant. We now show that

Imf,tf)-Kcr/tftf). Observe first that, since pf'^, = mft we have p'^iflq, if) - mfif) - 0 and therefore P?lfkiif)aiif) - ° whence Imf,(/) C Kerp"(/). To obtain the reverse inclusion, observe that for every j we have

3. Primary Decomposition 41 from which we see that Kertf(/)cnKerf,(/) CKer £f,(/) W = Ker(idv-f,(/)) by(l) = Im f, (/) by the Corollary to Theorem 2.4. As for the induced mapping f. \ V. -* Vit let m, be its minimum polynomial. Since p'f{f) is the zero map on Vh so also is pf'(/[•)• Consequently, we have that mf.\fi. Thus mfj\mf and the mf. are relatively prime. Suppose now that g 6 F[X] is a multiple of ntf. for every i. Then clearly g{ft) is the zero map on Vt. For every k k x - Y, v,- 6 0 K = V we then have /=i i=i *w(jr)»x:«w{v,)-x:*«xv,)-ov i=i i=i and so g(f) = 0 and consequently mf\g. Thus we see that mf is the least common multiple of my,,..., m^. Since these polynomials are relatively prime, we then have k k mf ~ Ylmfr ^ut we ^now mat mf - TIP?* an(* mat m/JP/'- Since all of the i=l i=l polynomials in question are monic, it follows that m^. = pe/ fori- 1,..., k. Finally, using Theorem 2.2, we can paste together bases of the subspaces V, to form a basis of V with respect to which, as seen above, the matrix of/ is of the block diagonal form Since, from the theory of determinants, k det(X/-M) = ndet(X/-A,) i=i it we see that cf = J[cfr Now we know that mfl = pe{ and so, since mfl and cfi have i-i k k d the same zeros, we must have cfl = pT{ some r, ^ pr Thus {] P? = «"/ = JI ft" from i=i i=i which it follows that r, = J, for i = 1,..., A:. D Corollary 1 (;= 1 it) dimV^rf.degfl,

42 Further Linear Algebra Proof dim V, is the degree of cf. = pf. D Corollary 2 If all the eigenvalues of f lie in the ground field F, so that c/=(X->,)d"(X->2^... (X->^; m/ = (X->,)"(X->2r..(X->jl)", k then V; = Ker (J - >,idv)ri is f-invariant, of dimension dit and V=QVh D Example 3.5 Consider the linear mapping/: R3 -» IR3 given by /(*,>,z) = (-z, x + zt y + z). Relative to the standard ordered basis of IR3, the matrix of/ is A = "0 0 1 0 0 1 -1" 1 1 It is readily seen from this that cf -mf- (X+1)(X-1)2. By Corollary 2 of Theorem 3.2 we have IR3 = Ker(/ + id) © Ker {f - id)2 with Ker(/" + id) of dimension 1 and Ker(/"- id)2 of dimension 2. Now {f + id)(x, >, z) = (x - z, x + y + z, y + 2z) so a basis for Ker (f + id) is {(1, -2,1)}. Also, (/-id)2(x,y,z)=(*-y + z,-2x + 2y-2z,x-;y + z) so a basis for Ker(/" - id)2 is {(0,1,1), (1,1,0)}. Thus a basis for IR3 with respect to which the matrix of/ is in block diagonal form is 5= ((1,-2,1),(0,1,1),(1,1,0)}. The transition matrix from B to the standard basis is 10 1" P= -2 1 1 110 and the block diagonal matrix that represents/ relative to

the basis B is -1 P~lAP= 2 1 -1 0

3. Primary Decomposition 43 Example 3.6 Consider the differential equation {IT + a^D"" + • • • • + axD + aQ)f = 0 with constant (complex) coefficients. Let V be the solution space, i.e. the set of all infinitely differentiable functions satisfying the equation. From the theory of differential equations we have that V is finite-dimensional with dim V = n. Consider the polynomial m = Xn + an.xXn~x + • • + a,X + a0. Over <D, this polynomial factorises as m = {X - £*,)'»(* -a2)" ■ • • {X -otk)e'. Then D : V —► V is linear and its minimum polynomial is m. By Corollary 2 of Theorem 3.2, V is the direct sum of the solution spaces V, of the differential equations (D-a,id)"/ = 0. Now the solutions of (D - a id)r/ = 0 can be determined using the fact that, by a simple inductive argument, (D - a id)r/ =^//(^./). Thus / is a solution if and only if {/(e^'f) - 0, which is the case if and only if e~°'/ is a polynomial of degree at most r - 1. A basis for the solution space of (D - a id)r/ = 0 is then {«", fc0",..., f'e0"}. EXERCISES 3.6 Consider the linear mapping/: IR3 -► IR3 given by /(*,>, z) = (2* + y -z, -2x-y + 3z, z). Find the minimum polynomial of/ and deduce that IR3 = Ker/©Ker(/-id)2. Find a block diagonal matrix that represents/. 3.7 Consider the linear mapping/: IR3 -► IR3 given by /(*, y, z) = {x + y + z, x + y + z, x + y + z). Determine the characteristic and minimum polynomials of/. Show that the matrices 'i l r i i i i i i , N = '3 0 0" 0 0 0 0 0 0 are similar.

44 Further Linear Algebra It is natural to consider special cases of the Primary Decomposition Theorem. Here we shall look at the situation in which each of the irreducible factors p, of mf is linear and each et - 1, i.e. when m/=(X-X,)(X-X2)...(X->4). This gives the following important result, for which we recall that/ : V -» V is said to be diagonalisable if there is a basis of V consisting of eigenvectors of /; equivalently, if there is a basis of V with respect to which the matrix of/ is diagonal. Theorem 3.3 Let V be a non-zero finite-dimensional vector space and let f : V —► V be linear. Then the following statements are equivalent: (1) the minimum polynomial mjoffisa product of distinct linear factors', (2) f is diagonalisable. Proof (1) => (2): Suppose that (1) holds and that m, = (X-X,)(X-X2)...(X-X4) where X,,..., Xk are distinct elements of the ground Meld. By Theorem 3.2, V is the direct sum of the /-invariant subspaces Vj = Ker(f - X,idv). For every x G Vt r we have {J -X jidv)(jr) = 0V and so/(x) = X,-jc. Thus each X, is an eigenvalue of/, and

every non-zero element of Vj is an eigenvector of/ associated with X,-. By Theorem 2.2 we can paste together bases of V,,..., Vk to form a basis for V. Then V has a basis consisting of eigenvectors of/ and so/ is diagonalisable. (2) => (1): Suppose now that (2) holds and let X,,..., \k be distinct eigenvalues of/. Consider the polynomial p=(X-X,)(X-X2)...(X-X4). Clearly, p{f) maps every basis vector to 0V and consequently p(f) = 0. The minimum polynomial m^ therefore divides p, and must coincide with p since every eigenvalue of/ is a zero of mf. □ Example 3.7 Consider the linear mapping/: P3 -► IR3 given by /1*,y,z) = (7jc -y -2z, -x + ly + 2z, -2x + 2y + lOz). Relative to the standard ordered basis of IR3, the matrix of/ is " 7 -1 -2" A =

 The reader can verify thatc7= (X-6)2(X-12)andm/ = (X-6)(X-\2). It follows by Theorem 3.3 that/ is diagonalisable. EXERCISES 3.8 If/ : IR" -► IRn is linear and such that/3 = /, prove that/ is diagonalisable. 3.9 Let/: R3 -► R3 be the linear mapping given by the prescription f{x,y,z) = (-2x -y + z, 2x + y -3z, -z). Find the eigenvalues and the minimum polynomial of/ and show that/ is not diagonalisable. 3.10 Determine whether or not each of the following mappings/: IR3 -► IR3 is diagonalisable: (a) /(*,>,z) = (3x -y + z, -x + 5y -z, x -y + 3z); (*>) /(*, y, z) = (2x, x + 2y, -x + y + z); M /(*, y, z) = (jt - z, 2y, x + y + 3z). An interesting result concerning diagonalisable mappings that will be useful to us later is the following: Theorem 3.4 Let Vbea non-zero finite-dimensional vector space and let f,g: V —*Vbe diagonalisable linear mappings. Then f and g are simultaneously diagonalisable (in the sense that there is a basis of V that consists of eigenvectors of both f and g) if and onlyiffog = gof. Proof =^ : Suppose that there is a basis {v,,..., vn} of V such that each v, is an eigenvector of both/ and g. If /(v,) = X, v, and g(v{) = /i,v, then we have figM = X,-/i,-v,- = /i,V. = g\f[vi)]. Thus/o g and g of agree on a basis and so/ o g = g of. «<= : Suppose now that fog = gof. Since / is diagonalisable its minimum polynomial is of the form »y = (x->,X*->2)-"(*-**) where \,,..., X* are distinct eigenvalues of/. By Corollary 2 of Theorem 3.2, we k have V = 0 V; where V; = Ker (/"-A.idv). Now since, by the hypothesis,/og = gof /=1 we have, for v, £ Vh /fe(v.-)] = «Wv,)] = ^,.) = ^,.)

 and so g(v,) G Vs. Thus each Vt is g-invariant. Now let g, : V, -► Vy be the linear mapping thus induced by g. Since g is diagonalisable, so is each g,, for the minimum polynomial of g, divides that of

g. We can therefore find a basis £, of V, consisting of eigenvectors of g,. Since every eigenvector of g, is an eigenvector of g, and since k every element of 1^ is an eigenvector of /, it follows that |J £, is a basis of V that consists of eigenvectors of both / and g. D Corollary Let AtB be n x n matrices over a field F. If A and B are diagonalisable then they are simultaneously diagonalisable (i.e. there is an invertible matrix P such that P~x AP and P~x BP are diagonal) if and only if AB - BA. □

4 Reduction to Triangular Form The Primary Decomposition Theorem shows that for a linear mapping/ on a finite- dimensional vector space V there is a basis of V with respect to which / can be represented by a block diagonal matrix. As we have seen, in the special situation where the minimum polynomial of/ is a product of distinct linear factors, this matrix is diagonal. We now turn our attention to a slightly more general situation, namely that in which the minimum polynomial of/ factorises as a product of linear factors that are not necessarily distinct, i.e. is of the form t=i where each e,^\. This, of course, is always the case when the ground field is C, so the results we shall establish will be valid for all linear mappings on a finite- dimensional complex vector space. To let the cat out of the bag, our specific objective is to show that when the minimum polynomial of/ factorises completely there is a basis of V with respect to which the matrix of/ is triangular. We recall that a matrix A = [a,y]nxn is (upper) triangular if ati - 0 whenever i > j. In order to see how to proceed, we observe first that by Corollary 2 of Theorem 3.2 we can write V as a direct sum of the/-invariant subspaces V{ - Ker (/"-A.idv)". Let/-: V; -v Vlf be the linear mapping induced on the 'primary component' V; by/, and consider the mapping/ - ^idv. : Vs -► V*. We have that (/J - \idv.)" is the zero map on Vn and so/ - >,idVj is nilpotent, in the following sense. Definition A linear mapping/ : V -► V is said to be nilpotent if f" = 0 for some positive integer m. Likewise, a square matrix A is said to be nilpotent if there is a positive integer m such that Am = 0. Example 4.1 The linear mapping/: R3 -► R3 given by /(*, y, z) = (0, xt y) is nilpotent. In fact, f2(xty,z) =/(0,x,y) = (0,0,*) and then/3 = 0.

 Example 4.2 If/ : C -v C is such that all the eigenvalues of/ are 0 then we have cf-Xn. By the Cayley-Hamilton Theorem, /n = cf{f) - 0 so that/ is nilpotent. Example 4.3 The differentiation mapping D: Bn[X] -*■ Bn[X] is nilpotent. EXERCISES 4.1 Show that the linear mapping/:

R3 —► R3 defined by /(*, >, z) = (-* -y - z, 0, x + > + z) is nilpotent. 4.2 If/ : IR2[*] -+ IR2M 's me l'near mapping whose action on the basis {1,X,X2}oflR2[X]isgivenby /(1)= -5-8X-5X2 /(*) = 1 + X + X2 /(X2) = 4 + 7X + 4X2 show that/ is nilpotent. 4.3 If /: V -v V is nilpotent prove that the only eigenvalue of/ is 0. 4.4 If V = Matnxn IR and A e V, show that/A : V -* V given by fA(X) = AX-XA is linear. Prove that if A is nilpotent then so \sfA. As we shall see, the notion of a nilpotent mapping shares an important role with with that of a diagonalisable mapping. In order to prepare the way for this we now produce, relative to a nilpotent linear mapping, a particularly useful basis. Theorem 4.1 Let V be a non-zero finite-dimensional vector space and letf:V—*Vbea nilpotent linear mapping. Then there is an ordered basis { V|,..., vn} of V such that /(v.) = 0V; /(v2)eSpan{v,}; /(v3) e Span{v,,v2}; f(v,,) e Span{v,,...,v,,_,}.

 Proof Since/ is nilpotent there is a positive integer m such that/" = 0. If / = 0 then the situation is trivial, for every basis of V satisfies the stated conditions. So we can assume that/ f 0. Let k be the smallest positive integer such that/* = 0. Then f'lf 0 for 1 < i < k - 1. Since/*"• f 0 there exists v e V such that/*-,(v) f 0V. Let v, = /*"' (v) and observe that/(v,) = 0V. We now proceed recursively. Suppose that we have been able to find v,,..., vr satisfying the conditions and consider the subspace W = Span {v,,..., vr}. If W = V then there is nothing more to prove. If W f V there are two possibilities, depending on whether Im/ C W or Im/ £ W. In the former case, let vr+1 be any element of V\W. In the latter case, since we have the chain {0V} = Im/* C Im/*"1 C • • • C Im/2 C Im/, there is a positive integer j such that Im/' £ W and Im/'+l C W. In this case we choose vr+, G Im/7 with vr+, $ IV. Each of these choices is such that {vlt..., vr+,} is linearly independent, with /(vr+,) £ W. □ Corollary V f '• V —* V " nilpotent then there is an ordered basis of V with respect to which the matrix of f is upper triangular with all diagonal entries 0. Proof From the above, the action of/ on the basis {vY,..., vn} can be described by /(v,)= 0v!+ 0v2+ 0v3+-+0vn /(^)=^12^+ 0v2+ 0v3+-+0vn /(v3)=a,3v,+a23V2+ 0v3 + ... + 0v,, /(V,,) = 0|,,Vi + fl2nV2 + «3» V3 + • • + 0V,, where the ai} belong to the ground field. The matrix of/ relative to this basis is then the n x n matrix /i = "0 a,2 a,3 0 0 <z23 0 0 0 0 0 0 «2n 0 which is upper triangular with all diagonal entries 0. D

 Returning to our consideration of the primary components, we can apply the above results to the nilpotent linear mapping g, = /- X,idv. on the direct summand V,- of dimension dt. Since ft - 8i + \idVj, we deduce from the above that there is an ordered basis of the subspace V, with respect to which Mat f; = Mat g, + \ Mat \6Vi *i 012 013 0 \ a23 0 0 0 0 0 0 \- J Consequently, we have the following result. Theorem 4.2 [Triangular Form] Let V be a non-zero finite-dimensional vector space over afield F and let f: V —*V be a linear mapping whose characteristic and minimum polynomials are 9=n(x->^;. m^ntx-*.)'• for distinct \it...t\k£F and et < </,-. Then there is an ordered basis of V with respect to which the matrix of f is upper triangular, more specifically, is a block diagonal matrix *2 M = l*J in which A,- is a dt x dt upper triangular matrix \ ? ■•■ ?" 0 \t ••• ? 0 0 ••• \t in which the entries marked ? are elements of F. □ Corollary Every square matrix over the field C of complex numbers is similar to an upper triangular matrix. □

4. Reduction to Triangular Form 51 1 0 -1 0 2 0 f 1 3 Example 4.4 Consider the linear mapping /: R3 —> IR3 given by /(*, >, z) = (* + z, 2> + z, -jc + 3z). Relative to the standard ordered basis of IR3, the matrix of / is /i = -1 0 3 The reader will readily verify that cf = mf = [X - 2)3. It follows by Corollary 2 of Theorem 3.2 that R3=Ker(/-2id)\ We now find a basis for H3 in the style of Theorem 4.1. First, we note that (f - 2 id)(jc,>, z) = (-jc + z, z, -jc + z). We therefore choose, for example, vl = (0,l,0)GKer(/-2id)\{0}. As for v2, we require that v2 be independent of v(and such that (/-2id)(v2)eSpan{v,}; i.e. we have to choose v2 = (jc, yt z) independent of v, = (0,1,0) such that (-jc + z.z.-jc + z) = <*(<), 1,0). We may take, for example, a - 1 and choose v2 = (1,0,1). We now require v3 independent of {v,, v2} such that (/-2id) (v3)eSpan{vltv2}. We may choose, for example v3 = (0,0,1). Consider now the basis B= {v,,v2,v3}= {(0,1,0),(1,0,1),(0,0,1)}. The transition matrix from B to the standard basis of IR3 is "0 1 0" P = 1 0 0 0 1 1 and so the matrix of/ relative to B is the upper triangular matrix P~lAP = 2 1 1 0 2 1 0 0 2

52 Further Linear Algebra Example 4.5 Referring to Example 3.5, consider again the linear mapping /: IR3 -* IR3 given by /(*,y,z) = (-z, x + z,y + z). We have IR3 = Ker {f + id) © Ker {f - id)2 with Ker (^+ id) of dimension 1 and Ker(/"- id)2 of dimension 2. Since (J + id)(x,y, z) = (x - z, x + > + z, > + 2z), a basis for V, = Ker(/"+ id) is {(1,-2,1)}. Now consider finding a basis for V2 = Ker (f - id)2 in the style of Theorem 4.1. First note that (/"-id)(x,y,z) = (-* -z, x -y + z, >). Begin by choosing W| = H,0,l)eKertf-id)\{0}. We now require w2

independent of w, with (f-idK^GSpanfM;,}; i.e. we have to find w2 = (x, >, z) independent of w, = (-1,0,1) such that {-x -z, x -y + z, y) = a(-l,0,1). We may take, for example, a = 1 and choose w2 - (0,1,1). Now since (/•-id)(M/,)= (0,0,0) = 0m/, +0m/2 if-i6)(w2) =(-1,0,1)= 1m/,+0m/2 we see that the matrix of / - id relative to the basis {w,, w2] is 0 f 0 0 ' The matrix of the mapping f2 that is induced on V2 = Ker (f - id)2 by / is then 1 1 0 1 ' Consequently, the matrix of / relative to the basis 5=((1,-2,1),(-1,0,1),(0,1,1)} is the upper triangular matrix [-1 0 0" 0 1 1 . 0 0 1

to Triangular Form 53 EXERCISES 4.5 Let /: IR3 -► IR3 be the linear mapping given by f{x1y1z)={2x + y-z,-2x-y + 3z1z). Find the characteristic and minimum polynomials of / and show that / is not diagonalisable. Find an ordered basis of IR3 with respect to which the matrix of / is upper triangular. 4.6 Let /: IR3 —> IR3 be the linear mapping given by /(*, y, z) = (2x - 2y, x - y, -x + 3> + z). Find the characteristic and minimum polynomials of / and show that / is not diagonalisable. Find an ordered basis of IR3 with respect to which the matrix of/ is upper triangular. We have seen above that if / : V —► V is linear and every eigenvalue of / lies in the ground field of V then each induced mapping / on the /-invariant subspace V; = Ker (/"-X.idv)" can be written in the form / = g, + X,idv. where g, is nilpotent. Clearly, X,idVi is diagonalisable (its minimum polynomial being X-\t). Thus every induced mapping ft has a decomposition as the sum of a diagonalisable mapping and a nilpotent mapping that, moreover, commute. That this is true of / itself is the substance of the following important result. Theorem 4.3 [Jordan Decomposition] Let V be a non-zero finite-dimensional vector space over afield F and let f : V —» V be a linear mapping all of whose eigenvalues belong to F. Then there is a diagonalisable linear mapping 8 : V —* V and a nilpotent linear mapping r\ : V —► V such that f = 8 + rj and 8 o r\ = f\ o 8. Moreover, there are polynomials p,q€ F[X] such that 8 = p{f) and r\ - q{f). Furthermore, 8 and r\ are uniquely determined, in the sense that if 8' ,r\' : V —* V are respectively diagonalisable and nilpotent linear mappings such that f = 8'+rj'with 6'ot/ = t/o6' then 8 = 8' and i\ = 7/. Proof k The minimum polynomial of / is mj = f[(* - X,-)" where X,,..., \k G F are /=i k distinct. Moreover, V = 0 V: where V; = Ker{f- A.idy)". ;=i k Let 8 : V -> V be given by 8 = J] X jP, where /¾f: V -► V is the projection on V; t=i parallel to J] Vj. Then for every v, e V, we have fi(v,-) = (£ X7pj)(vi) = X,v, and consequently V has a basis consisting of eigenvectors of 6 and so 6 is diagonalisable.

 Now define r\ = / - 6. Then for every v, e Vt we have lM=f(vi)-tM=if-><iHvft>i) and consequently ijr,(v,-) = (/"-\idv)"(v,) = Ov. It follows that, for some r, Ker rf contains a basis of V, so rf = 0 and hence 77 is nilpotent. k Now since V = 0 V] every vG V can be written uniquely in the form v = v, + • • + vk with v, 6 V;. Since each Vt is /-invariant, we then have ftVMJ = ftlflvi) + • • ■ +/(^)1=/(0=/IaM1 and hence p; o/ = / o p, for each i. Consequently, 6o/=E\p,o/=E^(p.0/) = E^0Pi)=/0EV.=/o«- 1=1 1=1 1=1 1=1 It follows from this that 6 o 77 = 6(/-6) = 6/-62 =/6 -62 = (/-6)6 = 7706. We now show that there are polynomials p, q £ F\X] such that 6 = /?(/) and 77 = q(f). For this purpose, we observe first thatp, = /^(/) where f, is the polynomial described in the proof of Theorem 3.2. Then, by the definition of 6, we have 6 = p(f) k where p = £ Ajf,. Since tj = / - 6, there is then a polynomial g such that r\ = ¢(/). i'=i As for uniqueness, suppose that 6', t/ : V —► V are diagonalisable and nilpotent respectively, with / = 6' + r\' and S' or\' -r\' o 6'. Since, as we have just seen, there are polynomials p, q such that 6 = /7(/) and r\ = ¢(/) it follows that 6' 0 6 = 6 0 6' and that Tj'oj] = i]oj]'. Now since 6 + r\ = / = 6' + tj' we have 6 - 6' = 7j' - r\ and so, since 77,77' commute, we can use the binomial theorem to deduce from the fact that 77 and 77' are nilpotent that so also is 77' - 77, which can therefore be represented by a nilpotent matrix N. Also, since 6,6' commute it follows by Theorem 3.4 that there is a basis of V consisting of eigenvectors of both 6 and 6'. Each such eigenvector is then an eigenvector of 6 - 6', and consequently 6 - 6' is represented by a diagonal matrix D. Now N and D are similar, and the only possibility is N = D = 0. Consequently, we have 6-6'= 0 = 77'-77 whence 8' = S and 77' = 77 as required. □ There is, of course, a corresponding result in terms of square matrices, namely that if all the eigenvalues of A £ Mat nxn F lie in F then A can be expressed uniquely as the sum of a diagonalisable matrix D and a nilpotent matrix N with DN = ND. Example 4.6 In Example 4.5 we saw that if/: R3 -► IR3 is given by /(*,>,z)=(-z, * + z. > + z)

 then relative to the basis the matrix of / is ^=((1,-2,1),(-1,0,1),(0,1,1)} r = "-1 0 0 0 1 0 0" 1 1 '-1 0 0" 0 1 0 0 0 1 , tf= '0 0 0" 0 0 1 0 0 0 Clearly, we can write T = D + N where D = and this is the Jordan decomposition of T. To refer matters back to the standard ordered basis, we compute PDP~l and PNP~l where P is the transition matrix 1 -1 0" />= -2 0 1 1 1 1 It is readily seen that PDP~l = i i _n 2 2 2 1 0 1 , PNP~l = r i _i _n 2 2 2 0 0 0 i i i L 2 2 2J L 2 2 2J Consequently, the diagonal part of / is given by df(x> y, z) = (\x + \y - \z> x + z, -\x + \y + £z), and the nilpotent part of / is given by «/(*» y>z)= ("ix ~\y~ jz» °»\x + 2>+ 2z)- EXERCISES 4.7 Determine the Jordan decomposition of the linear mappings described in Exercises 4.5 and 4.6. 4.8 Consider the linear mapping / : R3 -» R3 whose action on a basis {^1,^2,^3} °f R3 is given by /(^) = -^+2^ f{b2)=3bi+2b2 + b3\ f(b3) = -*3.

56 Further Linear Algebra (a) Show that the minimum polynomial of / is [X + 1)2(X - 2). {b) Determine a basis of F3 with respect to which the matrix of / is upper triangular. (c) Find the Jordan decomposition of /. The Jordan decomposition of a linear mapping / (or of a square matrix A) is particularly useful in computing powers of / (or of A). Indeed, since / = 6 + t] where 6, r\ commute, we can apply the binomial theorem to obtain The powers of £ are easily computed (by considering the powers of the corresponding diagonal matrix), and all powers of r\ from some point on are zero. EXERCISE 4.9 Determine the n-th power of the matrix "0 0 -1" A =

5 Reduction to Jordan Form It is natural to ask if we can improve on the triangular form. In order to do so, it is clearly necessary to find 'better' bases for the subspaces that appear as the direct summands (or primary components) in the Primary Decomposition Theorem. So let us take a closer look at nilpotent mappings. Definition If the linear mapping/ : V -» V is nilpotent then by the index of/ we shall mean the smallest positive integer k such that/* = 0. Example 5.1 As seen in Example 4.1, the mapping /: IR3 -»IR3 given by /(*, y, z) = (0, xt y) is nilpotent. It is of index 3. EXERCISES 5.1 Let/: V -* V be linear and nilpotent of index p. Prove that if x 6 V is such that/^'(x) f 0V then {*, /W /"-'(*)} is a linearly independent subset of V. 5.2 Let V be a vector space of dimension n over a field F. If/ : V —» V is linear prove that/ is nilpotent of index n if and only if there is an ordered basis of V with respect to which the matrix of/ is * 0 0 /,,-, o • Deduce that annxn matrix A over F is nilpotent of index n if and only if A is similar to this matrix.

58 Further Linear Algebra In order to proceed, we observe the following simple facts. Theorem 5.1 Iff: V —*Vis linear then, for every positive integer i, (1) Ker/'CKer/,+l; (2) i/xeKer/'*1 ften/M eKer/'. Proof (1) If x 6 Ker/1' then/'(x) = 0V gives/,+,(x) = f[f' {x)] = f{0v) = 0V and therefore x 6 Ker/'+l. (2) Ifx6Ker/,+l then/'[/W]=/,+,(x)=0v,andso/(x)6Ker/'. □ In general, for a

linear mapping /: V —» V we have, by Theorem 5.1(1), the chain of subspaces {Ov,} C Ker/C Ker/2 C • • • C Ker/' C Ker/,+l C ■ • • . In the case where/ is nilpotent, the following situation holds. Theorem 5.2 Let V be a non-zero vector space over a field F and let f : V —» V be a linear mapping that is nilpotent of index k. Then there is the chain of distinct subspaces {<M C Ker/ C Ker/2 C - • • C Ker/*"1 C Ker/* = V. Proof Observe first that Ker ff {0y}, for otherwise from fk{x) = 0y we would have /*"' {x) - 0V for every x, and this contradicts the hypothesis that/ is of index it. In view of Theorem 5.1, it now suffices to show that (1=1,...,/:-1) Ker/'> Ker/'*1. In fact, suppose that there exists i 6 {1,...,/:-1} such that Ker/; = Ker/,+l. Then, for every x 6 V, we have 0^ = /*W=/,+,l/H,+,)(x)] whence/H,+,)(x)6Ker/,+l = Ker/' and so 0, = /'[/H/+.)(x)]=/*-.W This produces the contradiction/*-1 = 0. □ In connection with the above, the following result will prove to be very useful. Theorem 5.3 Let V be a finite-dimensional vector space and let f : V -» V be linear and such that Ker/' C Ker/,+l. If {v,,..., vs} is a basis of Ker/' that is extended to a basis { v,,..., v4l w,,..., wt) of Ker/,+l then S = {/(w,),... ./(w,)} is a linearly independent subset of Ker/1.

 Proof If x 6 Ker//+I then by Theorem 5.1(2) we have/(x) 6 Ker/'. Thus we see that S C Ker/'. To see that S is linearly independent, suppose that Then /i,w, + -+/irwr6Ker/CKer/,+l. Now observe that we must have /i,»v, + • • • + ntwt = 0Vt for otherwise, since {v,,...,v,, wlt...twt) is a basis of Ker/,+1, there would exist scalars a,- and /¾. not all zero, such that r 5 r E/*.-w.- = Ea.-v.- + E/%w.- 1=1 1=1 1=1 whence we would have a dependence relation between the basis elements of the subspace Ker/,+l and clearly this is not possible. Since w,,..., w, are linearly independent, it now follows from the equality 11^^- +//,^, = 0V that each /i, = 0. Consequently S is linearly independent. D We now introduce a special type of upper triangular matrix. Definition By an elementary Jordan matrix associated with \ 6 F we shall mean either the 1 x 1 matrix [\] or a square matrix of the form ' \ 1 0 ... 0 0" 0 X 1 ... 0 0 0 0 \ ... 0 0 0 0 0 ... \ 1 0 0 0 ... 0 \ _ in which all the diagonal entries are X, all the entries immediately above the diagonal entries are 1, and all other entries are 0. By a Jordan block matrix associated with the eigenvalue \ 6 F we shall mean a matrix of the form h where each J,- is an elementary Jordan matrix associated with \ and all other entries areO.

 Theorem 5.4 Let V be a non-zero finite-dimensional vector space over afield F and let f: V —» V be linear and nilpotent of index k. Then there is a basis of V with respect to which the matrix off is a Jordan block matrix associated with the eigenvalue 0. Proof For i = 0,..., k let V/, = Ker /'. Since/ is nilpotent of index k we have, by Theorem 5.2, the chain {0V) = W0 C W, C W2 C ■ • • C W,_, C Wk = V. Choose a basis £, of W, = Ker/ and extend this by T2 CW2\Wl to a basis B2 = Bx U T2 of W2. Next, extend this by T3 C W3\W2 to a basis B3 = B2 U T3 of W3, and so on. Then Bk = fl, U r2 U ■ • ■ U Tk is a basis of ^. Now let us work backwards, replacing each 7,- = £,\£,_i as we go, except for Tk. To be more specific, let Tk - {*,,...,xa}. By Theorem 5.3 we have that {/(* i)• • • • •/(*€»)} ls a linearly independent subset of Wk_v Moreover, it is disjoint from Wk_2 since/(x,) 6 Wk_2 = Ker/*"2 gives the contradiction x, 6 Ker/*-1 = Wk_{. Consider therefore the set *MU {/(*,),...,/(*,,)}. This is linearly independent in WA_, and so can be extended to a basis of Wk_lt say ^Utflx,) f[xa))V{yu — .yfi} where each y,- 6 W*-i \Wn_2- In this way we have replaced 7^., in the basis Bk by r,, = {fixl)t...tfixa)}v{ylt...tyfi}. Repeating the argument with the role of Tk assumed by Tk_lt we can construct a basis of Wk_2 of the form Bt-iUifHxt) /2(*J}U (fly,),..., f1yfi))V {z,,...,z7} where each z, 6 W*_2\1VA_3. In this way we have replaced Tk_2. Continuing in this way, and using the fact that a basis for W0 = [0V] is 0, we see that we can replace the basis Bk of V by the basis described in the following array: ' k • ■* 11 • ■ • i *a i Tk-i*" /t*i). • ••. /Mi yi. •••. y/j. Tk-2** /2(*i). ...,/2W, /(yi). •••, /(y^), 2i,..,z7, *. W*-'(*,) /*-1W,/*"2(y.), .. ,/*-2(y/s) q ,¾. Note that in this array the elements in the bottom row form a basis for IV, = Ker/, those in the bottom two rows form a basis for W2 = Ker/2, and so on. Also,

 the array is such that every element is mapped by/to the element lying immediately below it, the elements of the bottom row being mapped to Oy. We now order this basis of V by taking the first column starting at the bottom, then the second column, starting at the bottom, and so on. Then, as the reader can easily verify (using a larger sheet of paper!), the matrix of/ relative to this ordered basis is a Jordan block matrix associated with the eigenvalue 0. D The above process is best illustrated by an example. Example 5.2 Consider the linear mapping/: R4 —» R4 given by /(fl|*,M) = (o,M,o). We have/2 = 0 and so/ is nilpotent of index k = 2. Now ^l = Ker/= {(0,6,c10);^lc6R}; W^Ker/^R4. A basis for Wx is Bx = {(0,1,0,0), (0,0,1,0)} which we can extend to a basis B2 = B,u7-2 = {(0,1,0,0), (0,0,1,0)}U{(1,0,0,0), (0,0,0,1)} of W2 - IR4. Now the image of T2 under/ is {(0,1,0,0),(0,0,1,0)} which is independent in W,. In fact, this is the basis £,, and so the array in the theorem becomes (1,0,0,0),(0,0,0,1), (0,1,0,0),(0,0,1,0). Arranging the columns bottom-up and from left to right, we obtain the ordered basis 5=((0,1,0,0),(1,0,0,0),(0,0,1,0),(0,0,0,1)}. In order to compute the matrix of/ relative to the ordered basis Bt we observe that the transition matrix from B to the standard basis is "0 10 0" 10 0 0 0 0 10" 0 0 0 1 Now since P~l = P and the matrix of/ relative to the standard basis is "0 0 0 0" 10 0 0 0 0 0 1 ' 0 0 0 0

62 Further Linear Algebra it follows that the matrix of/ relative to B is the Jordan block matrix rlAP = In practice, we rarely have to carry out the above computation. To see why, let us take a closer look at the proof of Theorem 5.5. Relative to the ordered basis B constructed from the table we see that there are ot > 1 elementary Jordan matrices of size k x kt then (3 ^ 0 of size (k - I) x (k -1), and so on. The number of elementary Jordan matrices involved is therefore a + /? + 7 + — + U;. But, again from the table, this is precisely the number of elements in the bottom row. But the bottom row is a basis of W, = Ker/. Consequently we see that the number of elementary Jordan matrices involved is dim Ker/. Returning to Example 5.2, we see that Ker/ has dimension 2, so there are only two elementary Jordan matrices involved. Since at least one of these has to be of size k x k = 2 x 2, the only possibility for the Jordan block matrix is 0 1 0 0 0 1 0 0 Our objective now is to extend the scope of Theorem 5.4 by removing the restriction that/ be nilpotent. For this purpose, let us return to the Primary Decomposition Theorem. With the notation used there, let us assume that all the eigenvalues of/ lie in the ground field F. Then we observe, by Theorem 3.2 and its Corollary 2, that the minimum polynomial off is (X - \-)r' and consequently the mapping f - \idyf is nilpotent of index exonthe dr dimensional subspace V,. Theorem 5.5 [Jordan Form] Let Vbea non-zero finite-dimensional vector space over afield F and let f : V —» V be linear. If \u...,\karethe distinct eigenvalues off and if each \j belongs to F then there b an ordered basis of V with respect to which the matrix off is a block diagonal matrix J i h in which Jt is a Jordan block matrix associated with X,.

5. Reduction to Jordan Form 63 Proof With the usual notation, by Theorem 5.4 there is a basis of V, = Ker(/"- Xlidv)"* with respect to which the matrix of/, - X,idy. is a Jordan block matrix with 0 down the diagonal (since the only

eigenvalue of a nilpotent mapping is 0). It follows that the matrix J, of/, is a Jordan block matrix with X, down the diagonal. □ Definition A matrix of the form described in Theorem 5.5 is called a Jordan matrix of/. Strictly speaking, a Jordan matrix of/ is not unique since the order in which the Jordan blocks Jt appear down the diagonal is not specified. However, the number of such blocks, the size of each block, and the number of elementary Jordan matrices that appear in each block, are uniquely determined by /. So, if we agree that in each Jordan block the elementary Jordan matrices are arranged down the diagonal in decreasing order of size, and the Jordan blocks themselves are arranged in increasing order of the magnitude of the eigenvalues, we have a form that we can refer to as 'the* Jordan matrix of/. Definition If A £ Matnx,, F is such that all the eigenvalues of A belong to F then by the Jordan normal form of A we shall mean the Jordan matrix of any linear mapping represented by A relative to some ordered basis. At this stage it is useful to retain the following summary of the above: • If the characteristic and minimum polynomials of/ (or of any square matrix k k A representing/) are c = Y[(X - X,)*' and m = J]{X - X,)" then, in the Jordan i=i i=i matrix of/ (or Jordan normal form of A)t the eigenvalue X, appears precisely dt times down the diagonal, the number of elementary Jordan matrices associated with X, is dim Ker(f, - X.idy.), which is the geometric multiplicity of the eigenvalue X,, and at least one of each of these elementary Jordan matrices is of maximum size *, x «,- Example 5.3 Let/: R7 -♦ R7 be linear with characteristic and minimum polynomials cf = (X - 1)3(X - 2)4, m, = (X - \)2(X - 2)\ In any Jordan matrix that represents/ the eigenvalue 1 appears three times down the diagonal with at least one associated elementary Jordan matrix of size 2x2; and the eigenvalue 2 appears four times down the diagonal with at least one associated

64 Further Linear Algebra elementary Jordan matrix of size 3x3. Up to the order of the blocks, there is therefore only one possibility for the Jordan normal form, namely 1 1 1 1 2 1 2 1 2 2 Example 5.4 Let us modify the previous example slightly. Suppose that cy is as before but that now mf={X-\)2{X-2)2. In this case the eigenvalue 2 appears four times in the diagonal with at least one associated elementary Jordan matrix of size 2x2. The possibilities for the Jordan normal form are then 1 1 1 2 1 2 2 1 2 1 1 1 2 1 2 Example 5.5 If/: IR5 —»IR5 has characteristic polynomial c/=(X-2)2(X-3)3 then the possible Jordan normal forms, which are obtained by considering all six possible minimum polynomials, are 2 1 2 B where A is one of 2 and B is one

of "3 1 ' 3 1 3 i "3 1 3 3 i "3 3 3

5. Reduction to Jordan Form 65 Example 5.6 For the matrix A = "1 0 0 3 7 9 -2] -4 -5 we have cA - (X -1)3 and mA = {X -1)2. The Jordan normal form is then 1 1 1 1 Example 5.7 Consider the matrix A = 2 1110 0 2 0 0 0 0 0 2 10 0 0 0 11 0 -1 -1 -1 0 As can readily be seen, the characteristic polynomial of A is cA = (X-\)HX-2)2. The general eigenvector associated with the eigenvalue 1 is [0,0,jc,-jc,0] so the corresponding eigenspace has dimension 1. Likewise, the general eigenvector associated with the eigenvalue 2 is [x,y,-y.O.O] so the corresponding eigenspace has dimension 2. Thus, in the Jordan normal form, the eigenvalue 1 appears three times down the diagonal with only one associated elementary Jordan matrix; and the eigenvalue 2 appears twice down the diagonal with two associated elementary Jordan matrices. Consequently the Jordan normal form of A is 1 1 1 1 1 EXERCISES 5.3 Determine the Jordan normal form of the matrix -13 8 1 2 -22 8 -22 13 0 -5 0 13 5 3 -1 5

66 Further Linear Algebra 5.4 Determine the Jordan normal form of the matrix "5-1-32-5 0 2 0 0 0 10 11-2 0-1031 1-1-11 1 5.5 Find the Jordan normal form of the differentiation map on the vector space of real polynomials of degree at most 3. 5.6 Let V be the vector space of functions q : F2 —»IR of the form q(x, y) = ax2 + bxy + cy2 + dx + ey+f. Let (p : V —» V be the linear mapping given by f(<l) = hlq{x> y)dy. Determine the matrix A/ that represents <p relative to the ordered basis B of V given by B={x2,xy,y2,x,y,\}. Compute (1) the characteristic and minimum polynomials of A/; (2) the Jordan normal form J of A/. Definition By a Jordan basis for/: V —► V we shall mean an ordered basis of V with respect to which the matrix of/ is a Jordan matrix. To obtain a method of computing a Jordan basis, consider first the t x t elementary Jordan matrix \ 1 \ 1 \ 1 \ 1 \ A corresponding basis {v{,..., v,} will be such that fM = Xv,; f{v2) = \v2 + iyf /(v3)= Xv3 + v2; M-i)= ^^,-1 + ^-21 /(vr)= Xv, + vM.

5. Reduction to Jordan Form 67 Thus, for every t x t elementary Jordan matrix associated with \ we require V!,..., v, to be linearly independent with (1) i^elmtf-XidJnKertf-Xid); (2) (/ = 2 f) tf-Xid)(v,.)=Vi- The solution in the general case is then obtained by applying the above procedure to each elementary Jordan matrix and pasting together the resulting ordered bases. Example 5.8 Let/: IR3 -»IR3 be given by /(*, y> z) = {x + >, -x + 3y, -x + y + 2z). Relative to the standard ordered basis of IR3, the matrix of/ is A = We

have cA - [X - 2)2 and mA - (X - 2)2. The Jordan normal form is then " 1 -1 -1 1 3 1 0" 0 2 y = 2 1 2 Now we have tf-2id)(jt,y,z) = {-x + >, -x + >, -x + y) and we begin by choosing Vi 6 Im {f-2 id)nKer {f-2 id). Clearly, ^ = (1,1,1) will do. Next we have to find v2, independent of v1(such that {f-2 id)(v2) = V| ■ Clearly, v2 = (1,2,1) will do. To complete the basis, we now have to choose v3 e Ker {f-2 id) with {vi.vj.vj} independent. Clearly, v3 = (1,1,0) will do. Thus a Jordan basis is B= {(1,1,1),(1,2,1),(1,1,0)}. To determine an invertible matrix P such that P~lAP = 7, it suffices to observe that the transition matrix from B to the standard basis is P = Then a simple calculation reveals that P~lAP = J. "1 1 1 1 2 1 1" 1 0

68 Further Linear Algebra Example 5.9 An ordered basis for IR4[X] is {1,X,X2,X3,X4}. Relative to this the differentiation map D is represented by the matrix A = 0 10 0 0 0 0 2 0 0 0 0 0 3 0 0 0 0 0 4 0 0 0 0 0 The characteristic polynomial of A is Xs % the only eigenvalue is 0, and the eigenspace of 0 is of dimension 1 with basis {1}. So the Jordan normal form of A is 7 = 0 10 0 0 0 0 10 0 0 0 0 10 0 0 0 0 1 0 0 0 0 0 A Jordan basis is {Pi,P2iP3iP4iP$} where Dp,=0, Dp2 = plt Dps=p2t Dp4 = p3l Dp5 = p4. We may choose pl = 1 and p2 = X, then p$ = jX2, p4 = gX3, p5 = jjX4. A Jordan basis is therefore {24, 24X, 12X2,4X3, X4}. EXERCISES 5.7 Determine a Jordan basis for the linear mappings represented by the matrices in (1) Exercise 5.4; (2) Exercise 5.5. 5.8 Referring to Exercise 5.6, determine (1) a Jordan basis for ip\ (2) an invertible matrix P such that P~lMP = J where J is the Jordan normal form of A/. 5.9 Suppose that/ : IR5 -► IR5 is represented with respect to the ordered basis {(1,0,0,0,0),(1,1,0,0,0),(1,1,1,0,0),(1,1,1,1,0),(1,1,1,1,1)}

5. Reduction to Jordan Form 69 by the matrix /i = 18 6 4 1 0 10 0 0 0 12 10 0-1-1 0 1 0-5-4-3 -2 Find a basis of IR5 with respect to which the matrix of/ is in Jordan normal form. An interesting consequence of the Jordan normal form is the following. Theorem 5.6 Every square matrix A over C is similar to its transpose. Proof Since all the eigenvalues of A belong to C it clearly suffices to establish the result for the Jordan form of A. Because of the structure of this, it is enough to establish the result for an elementary Jordan matrix of the form \ 1 \ 1 y = \ l \ Now if B={vu...,vk} is an associated Jordan basis, define (/=1,...,/:) Wi=vk_i+] and consider the ordered basis B*= {*!,...,w4} = {vt,...,i>,}. It is readily seen that the matrix relative to this basis is the

transpose of J. Consequently we see that J is similar to its transpose. D By way of an application, we shall now illustrate the usefulness of the Jordan normal form in solving systems of linear differential equations. Here it is not our intention to become heavily involved with the theory. A little by way of explanation together with an illustrative example is all we have in mind.

70 Further Linear Algebra By a system of linear differential equations with constant coefficients we shall mean a system of equations of the form *; = anxx + 0,2*2 + ■' + alnxn x2 ~ fl21*l + fl22*2 + ■ ■ ■ + <*2nxn *'n ~ anl*l + <>n2x2 + "' ' + am*n where x, ,..., xn are real differentiate functions of f, x\ denotes the derivative of xi% and ai} 6 R for all ij. These equations can be written in the matrix form (1) X' = AX where X = [x, • • ■ xj\ and A = [aiy]nxn. Suppose that A can be reduced to Jordan normal form JA, and let P be an invertible matrix such that P~lAP = JA. Writing Y = P-1X, we have (2) {PY)' = X' = AX = APY and so (3) Y' = P~lX' = P~lAPY = ^Y. Now the form of JA means that (3) is a system that is considerably easier to solve for Y. Then, by (2), PY is a solution of (1). Example 5.10 Consider the system i.e. X' = AX where We have X = c and so the Jordan form of A xj = JCj + x2 x'2 = -*i + 3jc2 *3 = ~"*1 + 4*2 ~ *3 V *2 .*3_ , 4 = "110 -13 0 -1 4 -1 'A = (X+\)(X-2)2 = mA is 4= "-1 2 1 2 .

5. Reduction to Jordan Form 71 We now determine an invertible matrix P such that P~lAP = JA. For this, we determine a Jordan basis. For a change, let us do so with matrices rather than mappings. Clearly, we have to find independent column vectors p,, p2, p3 such that 04 + /3)Pl = 0, 04-2/3)p2 = O, 04-2/3)p3 = pj. Suitable vectors are, for example, Pi = "0" 0 1 , P2 = ■f 1 1 . P3 = "-1" 0 0 Thus we may take /> = '0 1 0 1 1 1 -1" 0 0 [The reader should check that P~lAP = JA or, what is equivalent and much easier, that AP = PJA.] With Y = P~lX we now solve the system Y' = JAY, i.e. y\ = ->n y'i = 2>2+>3. /3 = 2>3- The first and third of these equations give >, = or,*-', >3 = a3<?2'; and the second equation becomes It follows that Consequently we see that)4 = 2)^+0^-y2 = a3fe2' + ot2e2t. Y = a2e2t + a3te2' a$e 2r A solution of the original system of equations is then given by a2e2t-¥ai(t-\)e2t n X = />Y = at2e2' + at3te2' 21 a,e~r + a2e + ot^te

72 Further Linear Algebra EXERCISES 5.10 Find the Jordan normal form J of the matrix 0 10-1 A = -2 3 0-1 -2 12-1 2-103 Find also a Jordan basis and an

invertible matrix P such that P~lAP - J. Hence solve the system of differential equations 4 L*4. 0 -2 -2 2 1 0 -1" 3 0-1 1 2 -1 -10 3 V *2 *3 .*4_ 5.11 Solve the system of differential equations x\ - xx + 3x2 — 2jCi x2 - 7x2 - 4jc3 *3 - 9jc2 ~ 5jc3. 5.12 Show how the differential equation jc'"-2jc"-4jc' + 8jc = 0 can be written in the matrix form X' = AX where " 0 0 -8 1 0 4 0" 1 2 [Hint. Write jc, = xt x2 - x\ x3 = x"\ By using the method of the Jordan form, solve the equation given the initial conditions x(0) = 0, /(0) = 0, jc"(0)= 16.

6 Rational and Classical Forms Although in general the minimum polynomial of a linear mapping/: V -» Vcan be expressed as a product of powers of irreducible polynomials over the ground field F of Vt say mf = pj'pf • -pj', the irreducible polynomials p, need not be linear. Put another way, the eigenvalues of/ need not belong to the ground field F. It is therefore natural to seek a canonical matrix representation for/ in the general case, which will reduce to the Jordan representation when all the eigenvalues of/ do belong to F. In order to develop the machinery to deal with this, we first consider the following notion. Suppose that W is a subspace of the F-vector space V. Then in particular W is a (normal) subgroup of the additive (abelian) group of V and so we can form the quotient group V/W. The elements of this are the cosets x + W={x + w\ w£W}, and the group operation on V/W is given by (x + W) + (y + W) = (x + y) + W. Clearly, under this operation the natural surjection iw'-V —> V/W given by \[w(x) - x + W is a group morphism. Can we define a multiplication by scalars in such a way that V/W becomes a vector space over F and the natural surjection \[w is linear? Indeed we can, and in only one way. In fact, for \\w to be linear it is necessary that we have the identity (VjtgV) (VXgF) xm*)-M**). and so multiplication by scalars must be given by \{x + W) = \x + W. With respect to this operation of multiplication by scalars it is readily verified that the additive abelian group V/W becomes a vector space over F. We call this the quotient space of V by W and denote it also by V/W.

 As we shall now show, if V is of finite dimension then so also is every quotient space of V. Theorem 6.1 Let V be a finite-dimensional vector space and let W bea subspace of V. Then the quotient space V/W is also finite-dimensional. Moreover, if {v,,..., v,,,} is a basis of W and {*, + W xk + W} is a basis of V/W then {v,,..., vmtxlt..., xk} is a basis of V. Proof Suppose that / = {x{ + W,..., xp + W] is any linearly independent subset of V/W. Then the set {xlt... txp} of coset representatives is a linearly independent subset of p V. To see this, suppose that £ \,Xi = 0V. Then, using the linearity of |JW, we have 1=1 Ov/w = M<V) = M t *.*.■) = E *M*i) = t M*.- + "0 t=i i=i t=i and so each \t is 0. Consequently, p < dim V. Since |/| = p it follows that every linearly independent subset of V/W has at most dim V elements. Hence V/W is of finite dimension. Suppose now that {v,,..., vm} is a basis of W and that {jc, + W,..., xk + W] is a basis of V/W. Consider the set B = {v|,..., vm x,,..., xk}. Applying \\w to any linear combination of elements of B we see as above that B is linearly independent. Now for every x 6 V we have x + W 6 V/W and so there exist scalars X, such that x + w=f:\i(xi + w)=(Y;\ixl) + w 1=1 1=1 * ft /11 and hence jc - ^ X,x, 6 W. Then there exist fj,j such that x - £ X,jc, = £ /i,Vy. i=l i=i y=i Consequently x is a linear combination of the elements of B. It follows that the linearly independent set B is also a spanning set and therefore is a basis of V. D Corollary 1 dimV=dimW + dimVyw. Proof dimy= |fi| = m + & = dimW + dimVytV. D Corollary 2 //^= w©Zf/K?nZ~vyw. Proof We have dim V = dim W+dim Z and so, by Corollary 1, dim Z = dim V/W whence it follows that Z~V/W. D

6. Rational and Classical Forms 75 EXERCISES 6.1 Let V and VV be vector spaces over a field F and let/: V —► VV be linear. If Z is a subspace of V prove that the assignment x + Zi->/(jt) defines a linear mapping from V/Z to VV if and only if Z C Ker/. 6.2 If /: V -► VV is linear prove that V/Ker/ ~ Im/. We shall be particularly interested in the quotient space V/W when VV is a sub- space that is /-invariant for a given linear mapping/: V —► V. In this situation we have the following result. Theorem 6.2 Let V be a finite-dimensional vector space and let f: V —► V be linear. If VV is an f-invariant subspace of V then the prescription /+(jc + W) = /(jc) + W defines a linear mapping /*" : V/W —► V/W, the minimum polynomial of which divides the minimum polynomial off Proof Observe that ifjc + W=y + W then x - y 6 VV and so, since VV is /-invariant, we have/(x) -f{y) =f{x-y)6W from which we obtain/(*) + VV = f{y) + VV. Thus the above prescription does indeed define a mapping/^ from V/W to itself That/* is linear follows from the fact that, for all jc, y 6 V and all scalars X, = /(* + y) + w = /*(* +WO+r(y + w); /IM*+ ^)1=/1^ + ^] = /(Xjc) + vv = \f{x) + W « WW + H! = \T(* + wo.

76 Further Linear Algebra To show that the minimum polynomial of /* divides that of/, we now show by induction that (voo) (T)n = (n+. For the anchor point n - 1 this is clear. As for the inductive step, suppose that (T)" = (H+- Then, for

every x 6 V, {r)n+i(x+w)=n(rnx+w)] = r[TM + w] = /lH*)] + w = /"+,(jc) + W = (f+,)+(x + W0, and consequently (/*)"♦' = (Jn+l)\ m It follows from this that for every polynomial p = £ a,X' we have i=0 i=0 i=0 Taking in particular p - mj we obtain mf{f*) - 0 whence Definition We call/* : V/VV —► V/W the linear mapping that is induced by / on the quotient space V/W. EXERCISE 6.3 For the linear mapping/ : F3 -> F3 given by f{xtytz) - {xtxtx) describe the induced mapping/*: F3/Ker/ —► F3/Ker/. We shall now focus on a particular type of invariant subspace. Suppose that V is a finite-dimensional vector space and that/ : V —► V is linear. It is readily seen that the intersection of any family of/-invariant subspaces of y is also an/-invariant subspace of V. It follows immediately from this that for every subset X of V there is a smallest/-invariant subspace that contains X, namely the intersection of all the /-invariant subspaces that contain X. We shall denote this by Zfx. In the case where X = {x} we shall write z{, or simply Zx when/ is clearly understood. The subspace Zx can be characterised as follows.

6. Rational and Classical Forms 77 Theorem 6.3 Let V be a finite-dimensional vector space over afield F and let f : V —* V be linear. Then, for every xeV, z, = {K/)M;peF[x]}. Proof It is readily seen that the set W - {p(J)(x) ; p 6 F[X]} is a subspace of V that contains x. Since /commutes with every p{f), this subspace is/-invariant. Suppose now that T is an /-invariant subspace that contains x. Then clearly T contains fk(x) for all k, and consequently contains p(f)[x) for every polynomial P G F[X]. Thus T contains W. Hence W is the smallest /-invariant subspace that contains x and so coincides with Zx. D Example 6.1 Let/: F3 -+ F3 be given by /(a,fc,c) = (-6 + c, a + c, 2c). Consider the element (1,0,0). We have/(1,0,0) = (0,1,0) and/^(1,0,0) = /(0,1,0) = -(1,0,0), from which it is readily seen that 2(1.0,0) = {(*,y.o); x,yew}. EXERCISE 6.4 Let/: F3 -+ R3 be given by /(a, 6, c) = (0,0,6). Determine Zx where x = (1, -1,3). Our immediate objective is to discover a basis for the subspace Zx. For this purpose, consider the sequence Jf./(jr)./2W...../rW.... of elements of Zx. Clearly, there exists a smallest positive integer k such that/*(x) is a linear combination of the elements that precede it in this list, say fk(x) = \0x + >,/(*) + • • • + W*-' W. and {*,/(.*), • • • ,fk~l(x)} 's men a l»nearly independent subset of Zx. Writing flj = -X,• for / = 0,..., k -1 we deduce that the polynomial mx = a0 + OiX + ■ • + flA-i**'1 + X* is the monic polynomial of least degree such that mx{f)(x) - 0V. Definition We call mx the /-annihilator of x.

78 Further Linear Algebra Example 6.2 Referring to Example 6.1, let x = (1,0,0). Then, as we have seen,/2(x) = -x. It follows that the/-annihilator of x is the polynomial mx - X2 + 1. EXERCISES 6.5 In Exercise 6.4, determine the/-annihilator of x = (1, -1,3). 6.6 Let/ : F3 —► F3 be the linear mapping that is represented, relative to the standard ordered basis of R3, by the matrix "10 1" 0 1 0 -1 0 -1 Determine the/-annihilator of x - (1,1,1). With the above notation we now have the following result. Theorem 6.4 Let V be a finite-dimensional vector space and let f: V —> V be linear. If x 6 V has f-annihilator mx - a0 + axX + • • + ak.xXk~x + Xk then the set is a basis of Zx and therefore dim Zx - deg mx. Moreover, if fx : Zx —► Zx is the induced linear mapping on the f-invariant subspace Zx then the matrix of fx relative to the ordered basis Bx is 0 0 0 ... 0 -a0 1 0 0 ... 0 -a, C = 0 10 ... 0 -a2 0 0 0 ... 1 -fl4_! Finally, the minimum polynomial of fx is mx. Proof Clearly, Bx is linearly independent and/*(x) £ Span Bx. We prove by induction that in fact/"(jc) 6 Span Bx for every n. This being clear for n - 1,...,/: suppose that n > /: and that/"-,(jc) 6 Span Bx. Then/n-,(jc) is a linear combination of jc, /(jc), ... /*'" (x) and so/" (jc) is a linear combination of/(x), f2{x),.... /*(jc), whence it follows that/"(.*) E Span Bx. It is immediate from this observation that

6. Rational and Classical Forms 79 Pif){x) £ Span Bx for every polynomial p. Thus Zx C Span Bx whence we have equality since the reverse inclusion is trivial. Consequently, Bx is a basis of Zx. Now since AM = /M /xM*)W2M fxlfk-2M)=fk-l(x) fxlfk-l(*)]=fk(x) = -a0x-aj{x) ak_J"{x) it is clear that the matrix of fx relative to the basis Bx is the above matrix Cmi. Finally, suppose that the minimum polynomial of fx is mu = b0 + bxX + • • • + br_{Xr~l + *r- Then <V = m/a (fl(x) = b0x + *>,/(*) + -.. + W^x) +/'(*) from which we see that/r(x) is a linear combination of x, /(x), ... /r-1(x) and therefore k < r. But m/f is the zero map on Zx, whence so is mx{fx). Consequently we have m/f \mx and so r ^ k. Thus r = k and it follows that m^ = mx. D Definition If V is a finite-dimensional vector space and /: V -► V is linear then a sub- space W of V is said to be /-cyclic if it is /-invariant and has a basis of the form {*>/(*)> ■ • ■ t/"^*)}. Such a basis is called a cyclic basis, and x is called a cyclic vector for W. In particular, Theorem 6.4 shows that x £ V is a cyclic vector for the subspace Zx with cyclic basis £,. The subspace Zx is called the/-cyclic subspace spanned by {x}. The matrix Cmx of Theorem 6.4 is called the companion matrix of the/-annihilator mx. Our first main objective can now be revealed. Bringing the above results together, we shall show that if /: V -► V

has minimum polynomial of the form p* where p is irreducible then V can be expressed as a direct sum of/-cyclic subspaces. The main consequence of this is that/ then has a block diagonal representation by companion matrices. One more observation will lead us to this goal. Theorem 6.5 Let W be an f-invariant subspace of V. Then for every x£V both the f-annihilator off and the j*-annihilator ofx + W divide the minimum polynomial off Proof By Theorem 6.4, the/-annihilator of x is the minimum polynomial of/x, the mapping induced on Zx by/, which clearly divides the minimum polynomial of/.

 As for the/*-annihilator of jc+IV, this likewise divides the minimum polynomial of /*" which, by Theorem 6.2, divides that of/. □ Theorem 6.6 [Cyclic Decomposition] Let V be a non-zero vector space of finite dimension over afield F and let f: V —► V be linear with minimum polynomial mf = p' where p is irreducible over F. Then there are f-cyclic vectors jrt,..., xk and positive integers nx,..., nk with each n, ^ t such that 0) V=0Zx#; /=1 (2) the f-annihilator of xk is pn'. Proof The proof is by induction on dim V. The result is trivial when dim V = 1. Suppose then that the result holds for all vector spaces of dimension less than n (where n > 1) and let V be of dimension n. Since mf = p' there is a non-zerox, e V with p'-1 (/)(^) f 0V. The/-annihilator of jcj is then mXl = p'. Let W = ZXl and let/*": V/W -> V/W be the induced linear mapping. By Theorem 6.2, the minimum polynomial of /*" divides m^ = p' and so the inductive hypothesis applies to/* and V/IV. Thus there exist/*-cyclic subspaces Z\i+w> ■ • • 121^+nf of V/W such that V7W=0ZV/+lv f=2 and, for 2 <i< kt the/^-annihilator of y, + W is p"* for some n, < t. We now observe that there exists jc, 6 y, + IV such that the/-annihilator of xt is p"'. In fact, since the/*-annihilator of y, + W is pn\ we have p^)"*(y,) £ IV = ZXl. Thus there is a polynomial h such that P(/T'(y.) = *(/)(*,)- It follows from this that o,«i#«y,) = i</n-Mfl(*i). Butpristhe/-annihiIatorof *,,sop'lp'^'fiandhencepn'\h. Consequentlyh = p^q for some polynomial ¢. Now define jc, = y, -^(/)(^). Then y,-*, = <7(/)(*i)eiv and so jcf 6 yf + W whence jc, + IV = y(+ IV. The/*-annihilator of y, + W therefore divides the/-annihilator of x,. But pl/TW = iWh -«(fl(*i)] = /Wfa) - Mfl(*i) = o,. Thus we see that the/-annihilator of jc, is p"\

 Suppose now that deg p = d. Then deg pn' - dn,. Now since p"» is both the /-annihilator of jc, and the/^-annihilator of jc, + Wt it follows by Theorem 6.4 that 4 = {*/./(*,) f*-l(Xi)} is a basis for ZXi%

and that Bi={*,+w, r (*i+ho. • • •, cry*"-' (*,-+wo> is a basis for Zx.+W. But since i=2 1=2 * k it follows that (J Bs is a basis of V/W. Then, by Theorem 6.1, (J A, is a basis of V. i=2 i=| k Consequently we have that V=QZX. and the induction is complete. □ i=i Corollary 1 k With the above notation, relative to the basis (J Ak the matrix of f is of the form r"=l where C, is the companion matrix of mXl = pn'. D Corollary 2 dim V = (n{ + • ■ • + nk) deg p. D Without loss of generality, we may assume that the/-cyclic vectors xx,..., xk of Theorem 6.6 are arranged such that the corresponding integers n, satisfy f = n, ^n2^ ■•■ >*0 1- With this convention we have the following result. Theorem 6.7 The integers nl%...tnkare uniquely determined by / Proof From the above we have, for every /, dim Zx. = deg mx. = deg p* = dnr

82 Further Linear Algebra Observe now that for every,/ the image of Zx. under p{f)J is the/-cyclic subspace Z^/(x#). Since the/-annihilator of jc, is pn', of degree dnh we see that f 0 ifj>fif; I d{n, -j) if/ < nr dim ZMiK Now every jc £ V can be written uniquely in the form x = v, + --- + v4 (v,ezx.) and so every element of Imp(f)J can be written uniquely in the form pWjW = pWj(vi) + - +pW'(v4). Thus, if r is the integer such that nlt...,nr>j and n^., </ then we see that Imp(/r = 0Z,(/)'<*,> and consequently dim ImpW = d £(ni -j) = </ £ (n, - ,). It follows from this that dim Imptf);-1 - dim lmp(f)i = </(£ fa -j + 1) - £ (n, -j)) nj>j-l ni>j = J x (the number of n, ^J). Now the dimensions on the left of this are determined by / so the above expression gives, for each j, the number of n, that are greater than or equal to j. This determines the sequence ' = «i > «2 >*••>«*> 1 completely. D Definition When the minimum polynomial of/ is of the form p' where p is irreducible then, relative to the uniqely determined chain of integers t - n{^ n2^ • • -^ nk ^ 1 as described above, the polynomials p' = pny ,p"2,... ,pn* are called the elementary divisors of/. It should be noted that the first elementary divisor in the sequence is the minimum polynomial of/.

6. Rational and Classical Forms 83 We can now apply the above results to the general situation where the characteristic and minimum polynomials of a linear mapping/: V -► V are '/-nM-lf. rnf = $#...$ where p,,... tpk are distinct irreducible polynomials. We know by the Primary Decomposition Theorem that there is an ordered basis of V with respect to which the matrix of/ is a block diagonal matrix "*i V in which each At is the matrix (of size dt deg p, xd{ deg p,) that represents the induced mapping/ on V) = Kerp,(/)'\ Now the minimum polynomial of/ isp" and so, by the Cyclic Decomposition Theorem,

there is a basis of Vt with respect to which At is the block diagonal matrix in which the Ciy are the companion matrices associated with the elementary divisors of/. By the previous discussion, this block diagonal form, in which each block A, is itself a block diagonal of companion matrices, is unique (to within the order of the Aj). It is called the rational canonical matrix of/. It is important to note that in the sequence of elementary divisors there may be repetitions, for some of the n, can be equal. The result of this is that some companion matrices can appear more than once in the rational form. Example 6.3 Suppose that/: R4 -► IR4 is linear with minimum polynomial Then the characteristic polynomial must be cf = (X2+1)2. SinceX2 + l is irreducible over IR it follows by Corollary 2 of Theorem 6.6 that 4 = (n, + • • • + nk)2. Now since the first elementary divisor is the minimum polynomial we must have n, = 1. Since we must also have each n, ^ 1, it follows that the only possibility is k = 2 and

 Hi = /¾ = 1. The rational canonical matrix of/ is therefore CX2 + , ©Cxi + , = Example 6.4 Suppose now that/: IR6 -* IR6 has minimum polynomial m/ = (X2 + l)(X-2)2. The characteristic polynomial of/ is then one of c, = (X2 + l)2(X-2)2, c2 = (X2 + l)(X-2)4. Suppose first that cy = c,. In this case we have IR6 = Vt®V2 with dim V, = 4 and dim V2 = 2. The induced mapping/, on V, has minimum polynomial m, = X2 + 1 and the induced mapping/2 on V2 has minimum polynomial m2 = (X -2)2. The situation for V, is as in the previous Example. As for V2% it follows by Corollary 2 of Theorem 6.6 that 2 = n, + ■ • • + nk whence necessarily k = 1 since n, = deg p2 = 2. Thus the only elementary divisor of/2 is (X - 2)2. Combining these observations, we see that in this case the rational canonical matrix of/ is Cxi+l ®Cxi+i ©C(X_2)i. Suppose now that cy = c2. In this case we have IR6 = V, © V2 with dim V, = 2 and dim V2 = 4. Also, the induced mapping /2 on V2 has minimum polynomial m2 = (X -2)2. By Corollary 2 of Theorem 6.6 applied to/2, we have 4 = n, + • • •+/¾ with n, = 2. There are therefore two possibilities, namely k-2 with nx = n2- 2; /: = 3 with n, = 2, n2 - n3 = 1. The rational canonical matrix of/ in this case is therefore one of the forms Of J+l ©C(x-2)i ®^(X-2)J*» Q'+i © C(X_2)2 © Cx_2 © Cx_2. Note from the above Example that a knowledge of both the characteristic and minimum polynomials is not in general enough to determine completely the rational form. The reader will not fail to notice that the rational form is quite different from the Jordan form. To illustrate this more drammatically, let us take a matrix in Jordan form and find its rational form. 0 -1 1 0 0 -1 1 0

6. Rational and Classical Forms 85 Example 6.5 Consider the matrix A = "2 0 0 1 0' 2 1 0 2 We have cA = (X - 2)3 = mA and, by Corollary 2 of Theorem 6.6,3 = nx + • • • + nk with n, = 3. Thus k = 1 and the rational form is C<X-2)3 = 0 0 8 ' 1 0 -12 0 1 6 EXERCISES 6.7 Determine the rational canonical form of the matrix a 0 1' 4 = 0 a 1 -1 1 a "1 2 2 2 2] 1 2 2 1 6.8 Let/ : R3 -v IR3 be the linear mapping that is represented, relative to the standard ordered basis of IR3, by the matrix A = Determine the rational canonical matrix of/. 6.9 Let/ : IR3 —► IR3 be the linear mapping that is represented, relative to the standard ordered basis of IR3, by the matrix "5-6-6" A= -1 4 2 3 -6 -4 Determine the rational canonical matrix of/. 6.10 Suppose that/ : F7 -» IR7 is linear with characteristic and minimum polynomials c, = (X - \)\X-2)\ m, = (X- \)2(X -2)3. Determine the rational canonical matrix of/. 6.11 If V is a finite-dimensional vector space and/ : V -» V is linear, prove that V has an/-cyclic vector if and only if cy = mf.

86 Further Linear Algebra The fact that the rational form is quite different from the Jordan form suggests that we are not quite finished, for what we seek is a general canonical form that will reduce to the Jordan form when all the eigenvalues belong to the ground field. We shall now proceed to obtain such a form by modifying the cyclic bases that were used to obtain the rational form. In so doing, we shall obtain a matrix representation that is constructed from the companion matrix of p, rather than from that of p"'. Theorem 6.8 Let x be a cyclic vector of V and let f: V —» V be linear with minimum polynomial ff where p = a0 + atX + • • • + ak_tXk~l + Xk. Then there is a basis of V with respect to which the matrix off is the kn x kn matrix ~Cp M Cp M Cp M in which Cp is the companion matrix of p and M is the kx k matrix f Proof Consider the n x k array /*-'(*) ... f(x) x P(f)lfk-l(x)] ... pVMx)] p(f)(x) p(f)n-l[fk-l(*)] ••• pW'I/WI pW-'M We first show that this forms a basis of V\ For this purpose, it suffices to show that the elements in the array are linearly independent. Suppose, by way of obtaining a contradiction, that this were not so. Then some non-trivial linear combination of these elements would be 0V and consequently there would exist a polynomial h such that h{f)(x) = 0V with deg h < kn = deg p". Since x is cyclic, this contradicts the hypothesis that p" is the minimum polynomial of/. Hence the above array constitutes a basis for V. We order this basis in a row-by-row manner, as we normally read.

6. Rational and Classical Forms 87 Observe now that/ maps each element in the array to its predecessor in the same row, except for those at the beginning of a row. For these elements we have and similarly, writing p(f) as h for convenience, /[/in-m(f *-')(*)] = A"-[f*(x)J = ^-n-**-,/*-'W <*<,* + p(/)(x)] = -ctk-xhn-m\fk-x(x)\ a0*"-"(x) + /^+1 (x). It is now a simple matter to verify that the matrix of/ relative to the above ordered basis is of the form described. D Definition The block matrix described in Theorem 6.8 is called the classical p-matrix associated with the companion matrix Cp. Moving now to general considerations, consider a linear mapping / : V —► V. With the usual notation, the induced mapping / on V, has minimum polynomial pf'. If we now apply Theorem 6.8 to the cyclic subspaces that appear in the Cyclic Decomposition Theorem for/, we see that in the rational canonical matrix of/ we can replace each diagonal block of companion matrices C{) associated with the elementary divisors p'" by the classical p,-matrix associated with the companion matrix of p,. The matrix that arises in this way is called the classical canonical matrix of/. Example 6.6 Let/: IR6 -► IR6 be linear and such that c/ = m/ = (X2-X + l)2(X + l)2. By the Primary Decomposition Theorem we have that the minimum polynomial of the induced mapping/, is (X2 -X + 1)2 whereas that of the induced mapping/2 is (X + 1)2. Using Corollary 2 of Theorem 6.6 we see that the only elementary divisor of/, is (X2 -X +1)2, and the only elementary divisor of/2 is (X +1)2. Consequently the rational canonical matrix of/ is '0 0 0-1 10 0 2 0 10-3 0 0 12 0 -1 1 -2 C/VI_ (X'-X+l)1 ^^(X+l)1 -

88 Further Linear Algebra The classical canonical matrix of/ can be obtained from this as follows. With p, = X2 -X +1 we replace the companion matrix associated with p} by the classical Pi-matrix associated with the companion matrix of p,. We do a similar replacement concerning p2 = X + 1. The result is Cyl_ XJ-X+I M. 2x2 Q'- X+l 'X+l wlxl Cx+,J = "0 1 -1 0 1 0 0 1 1 0 -1 1 -1 1 -1 Finally, let us take note of the particular case of Theorem 6.8 that occurs when k = 1. In this situation we have p = X -aQt and/ -o^idy is nilpotent of index n. Consequently C. reduces to the 1 x 1 matrix [a0] and the classical p-matrix associated with Cp reduces to the n x n elementary Jordan matrix associated with the eigenvalue a0. Thus we see that when the eigenvalues all belong to the ground Meld the classical form reduces to the Jordan form. Example 6.7 Let/: IR7 -> IR7 be linear and such that cf = (X - 1)3(X -2)4, m, = (X - 1)2(X -2)3. The rational canonical matrix is C(X-lp © CX-I © C(x_2)J © Cx-2 = ro l 1 2 l . 1 l * 8 -12 6 2. Here all the eigenvalues belong to the ground

Meld, so the classical canonical matrix coincides with the Jordan normal matrix. The reader can verify that it is n l l l . ■ 2 1 2 1 2 2.

6. Rational and Classical Forms 89 EXERCISES 6.12 Prove that if g e F[X] is monic of degree n and if Cg is the companion matrix of g then det(X/n-Cg) = g. [Hint. Use induction on n. Let g = a0 + a,X + .. •+an_1Xn-,+Xn and let /i 6 F[X] be given by h = a, + a2X + • • • + an_,Xn-2 + X""1. Show that det (X/n - Cg) = Xh + a0.] 6.13 For each of the following matrices over IR determine (a) the characteristic polynomial; (b) the minimum polynomial; (c) the elementary divisors; (d) the rational canonical form; (e) the classical canonical form. 1 1 3 5 2 6 -2 -1 -3 6.14 Let V be a real vector space of dimension 10 and let/: V -» V be linear and such that cf = (X2 -X + 1)3(X + l)4, mf = (X2 -X + 1)2(X + I)2. Determine the rational and classical canonical matrices of/. 6.15 Repeat the previous exercise assuming now that V is a complex vector space. "2 2 r 2 2 1 2 2 1 0 0 11 0 0 11 110 0 110 0

7 Dual Spaces If V and W are vector spaces over the same field F, consider the set Lin(V, W) that consists of all the linear mappings/ : V -» W. Given/, g 6 Lin(V, HO, define the mapping/ + g: V -» W by the prescription (V*6V) (f+ *)(*) = /{*) + *(*). Then it is readily seen that/ + g 6 Lin(V, W). Also, given/6 Lin(V, IV) and A 6 F, define the mapping \f: V -» W by the prescription (V*€V) (>/)(*) = >/(*)• Then it is readily seen that \f 6 Lin(V.lV). Moreover, under these operations of addition and multplication by scalars, Lin(V, W) becomes a vector space over F. A particular case of this is of especial importance, namely that in which for W we take the ground field F (regarded as a vector space over itself)- It is on this vector space Lin(V, F) that we shall now focus our attention. Definition We shall call Lin(V, F) the dual space of V and denote it by Vd. The elements of Vd, i.e. the linear mappings from V to the ground field F, are called linear forms (or linear functional) on V. Example 7.1 The i-ih projection pt: IRn -»IR given by p,(jc,,..., xn) = xtr is a linear form on IR" and so is an element of (IRn)d. Example 7.2 If V is the vector space ofnxn matrices over a field F then the mapping tr: V -» F n that sends each A 6 V to its trace tr A = £ a,, is a linear form on V and so is an element of Vd.

7. Dual Spaces 91 Example 7.3 The mapping / : IR[X] -► IR given by l{p) - JjJ p is a linear form on IR[X] and so is an element of IR[X]d. In what follows we shall often denote a typical element of Vd by xd. The reader should

therefore take careful note of the fact that this notation will be used to denote a linear mapping from V to the ground field F. We begin by showing that if V is of finite dimension then so is its dual space Vd. This we do by constructing a basis for Vd from a basis of V. Theorem 7.1 Let {v,,..., vn} be a basis of V and for / = 1,..., n let vf : V —» F be the linear mapping such that At \ x \Xf ifi = i; l0f ifjfL Then {vf,..., vj} is a basis of Vd. Proof n It is clear that each vf £ V*. Suppose that £ X,vf = 0 in Vd. Then for; = 1,..., n we have 0F = (t V?)(vy) = t KA*j) = t M.7 = V i= i i=i i=i from which we see that {vf,..., vj} is linearly independent. n If now x = 52 xjvj G ^ men we nave y=i (1) V?(X) = 5>;V?(vy) = £Xj6U = Xt j=\ j=l and hence, for every/6 V*. i=| i=l /=1 1=1 Thus we see that (2) (v/eV) /=£/(*.>?. i=i which shows that { vf,..., vj } also spans V*. whence it is a basis. □ Corollary If V b a finite-dimensional vector space then dim Vd = dim V. U

92 Further Linear Algebra Note from the equalities (1) and (2) in the above proof that we have (V*GV) x = fvf(x)v/; 1=1 Definition If {v,,..., vn} is a basis of V then by the corresponding dual basis we shall mean the basis {vf,..., vj[} described in Theorem 6.1. Because of (1) above, the mappings vf,..., vj[are often called the coordinate forms associated with v,,..., v,,. Example 7.4 Consider the basis {v,, v2} of IR2 where v, = (1,2) and v2 = (2,3). Let {vf, x{) be the corresponding dual basis. Then we have l = vf(vl) = vf(1,2) = vf(l,0) + 2vf(0,l); 0 = vf(v2) = vf(2,3) = 2vf(l,0) + 3vf(0,l). These equations give vf (1,0) = -3 and vf (0,1) = 2 and hence vf is given by rf{x,y) = -3x + 2y. Similarly, we have 4{x,y) = 2x-y. Example 7.5 Consider the standard basis {e,,..., en] of IRn. By definition, we have ef (ey) = 6,7 and so n n *?(*i, •••,*,,) = *?(£ xjej) = £ xjef{ej) = x, 7=1 7=1 whence the corresponding dual basis is the set of projections {pu...,pn}. Example 7.6 Let f,,..., rn+, be distinct real numbers and for each i let ¢,, : Rn[X] -» IR be the substitution mapping, given by ¢,, (p) = p{t:). Then B = {¢,,,..., ¢,^,} is a basis for IRnf X]d. To see this, observe that since IRn[X]d has the same dimension as IRn[X]

7. Dual Spaces 93 (namely n +1), it suffices to show that B is linearly independent. Now if £ \£# = 0 then we have /n+I % /=1 o = (EU,,)(*) = V, + M2 + - •• + *,,♦.',,♦. n+l i-i This system of equations can be written in the matrix form A/x = 0 where x = [X i ... X,,+|]' and M is the Vandermonde matrix M = By induction, it can be shown that 1 1 'i h r, t2 'n+l det A/=[](',-';). Since t,,..., fn+, are distinct, it follows that det A/ f 0. The only solution of Mx = 0 is

therefore x = 0 whence X, = • • ■ = \n+l = 0 as required. To determine a basis of IR,,[X] of which B is the dual, let such a basis be ^={Pi,-,P,,+i}- Then we require Cf.(p/) = fi,-y for all ij. It is readily seen that the Lagrange polynomials are the successful candidates. EXERCISES 7.1 If B- [Ejj ; i= l,...,mandy = 1,...,n] is the canonical basis of Matfflxn F, determine the corresponding dual basis. 7.2 Let V be the vector space of n x n matrices over a field F. If B 6 V prove that tpB : V -> F given by p* (A) = tr B'A is a linear form on V. Prove also that every <p 6 V is of this form for some B£V.

94 Further Linear Algebra Theorem 7.2 Let (v,)n, (wi)n be ordered bases of V and let (vf)n, (wf)n be the corresponding dual bases. If P is the transition matrix from (v,)n to (wt)n then the transition matrix from (v?),, to (w?),, is (/>-»)'. Proof Let the transition matrix from (vf)n to (wf)n be Q. Then we have, for each i, n n /=i /=i Consequently, ^=^/) = 1:^(1:^(0) *=i 1=1 *=i i=i n = YLikiPkj *=l = [(21¾. Thus we see that Q'P = /,, and therefore (2 = (P~l)'• E Theorem 7.2 provides a very convenient way of exhibiting dual bases which we shall now describe. Given a basis *={(<*l|.-».fl|n), (021.--.0¾). ---.(0.1.---.0} of I FT, the transition matrix from £ to the standard basis («,-),, is the matrix P whose i-th column is [an ... ain]'. By Theorem 6.2, if we associate with the i-th row of P-1 the mapping [irn,..., 7rin]: IRn -» R defined by ki.--.0(^1.--.^) = ^1+-+7½ then we see that Ki.--. ^. J(fl/I. • • •. fl/n) = I'*" P]ij = *./ and hence we can represent the dual basis of B by ^- {[ff|i,-..,iriJi [*2i.---.0. ••-. Kii.-'-.O}- Example 7.7 Consider again Example 7.4 where we computed the basis that is dual to the basis B = {(1,2), (2,3)} of IR2. The transition matrix from B to the standard basis is P =

7. Dual Spaces 95 and its inverse is [2 ~l The basis that is dual to B can therefore be described as * = {[-3.2], [2,-1]} where the notation is defined such that [-3,2](xty) = -3* + 2y, [2, -\](x,y) = 2x-y. Example 7.8 Consider again Example 7.5 where we computed the basis that is dual to the standard basis B = (e,)n of IRn. Applying the above procedure, we see that P- /n = P~l from which it follows that the dual basis consists of the projections pu...tpn. EXERCISES 7.3 Determine the basis of (IR3)" that is dual to the basis {(1,0,-1),(-1,1,0),(0,1,1)}. 7.4 Determine the basis of (IR4)" that is dual to the basis {(4,5,-2,11),(3,4,-2,6),(2,3,-1,4),(1,1,-1,3)}. 7.5 Let A = {jc, , x2} be a basis of a vector space V of dimension 2, and let Ad = (tpl, <p2} be the corresponding dual basis of V*. Find, in terms of <p{ and p2, the basis of Vd

that is dual to the basis 5= {jc, +2x2, 3jc,+4jc2} ofV. We now introduce the following notation. Given x 6 V and y4 e Vd we shall write With this notation, the following identities are immediate from the linearity of the mappings involved: (*) (*+ 7.^) = (^) + 0^): (0) (jr./+z"> = (*./) + (*.**); (7) (**./) = * (*./); (6) (jr,X/) = X(jcy>.

96 Further Linear Algebra Now it is clear from (/?) and (6) that for every x 6 V the mapping x : Vd -► F defined by ?(/) = (*,/) is a linear form on Vd% i.e. is an element of the dual space V = (Vd)d of V*. Moreover, by (a) and (7) it is quickly verified that the mapping otv : V —» V defined by ctv(x) - * ls dso linear. Definition We call V the bidual of V, and av the canonical mapping from V to V. • Note that the various notations employed above can be summarised by the following identities: ?(/) = (*,/) = /(*) ■ Theorem 7.3 If V is of finite dimension then the canonical mapping av : V —► V is an isomorphism. Proof n Let {v,,..., vn} be a basis of V and consider x = £ xy v, 6 Ker av. Then x is the zero of V and so 3c(yd) = 0 for all yd 6 V. In particular, for i = 1,..., n we have 0 = ?(v?) = (,, v?) = vf M = £ xyv? (v;) = x,- and consequently x = 0V. Thus aty 's injective. Now by the Corollary of Theorem 7.1 we have dim V = dim 1^ = dim V. It follows therefore that otv is an isomorphism. D In the case where V is of finite dimension we shall agree to identify V and V. This we can do only because the isomorphism av is natural, in the sense that it is independent of the choice of basis. Example 7.9 The mappings/,,/2: IR|[X] -> IR given by /.(P)= [p, /2(P)= fp Jo Jo are linear forms and, as is readily seen, {/1,/2} forms a basis of the dual space IR,[X]d. Since R|[X] ~ R2 under the correspondence a0 + a,X <—>(a<,,0|)

7. Dual Spaces 97 we have that R,[X]d ~ (|R2)d. If p = a0 + axX then /i(p) = fl0 + ^i /2(p) = 2a0 + 2a, and so we can associate with {/i,/2} the basis B = {[1, j], [2,2]} of (IR2)*'. By considering the matrix rl 2 whose inverse is P = />-' = 1 2 2 * 2 -2 we see that {(2, -2), (~, 1)} is the basis of IR2 = IR2 that is dual to B. Hence {2 - 2X, -£ + X] is the basis of R,[X] that is dual to the basis {/, ,/2}. EXERCISES 7.6 Determine a basis of IR4 whose dual basis is {[2,-1,1,0),1-1,0,-2,01,1-2,2,1,01,1-8,3,-3,1]}. 7.7 Let <p,, <p2, y>3 be the linear forms on ^[X] defined by ¥>i(p) = / P. ¥>2(p) = *>P(1), ¥>3(P) = P(0)- Jo Show that B = {p,, p2, P3} is a basis of IRJX]' and determine the basis {P1.P2.P3} of ^2W mat is duaI t0 B- 7.8 Let p,,..., <pn 6 (IR")4. Prove that the solution set C of the linear inequalities V>M>0, ¥>2M ^0.-•-.¥>«(*) ^0

satisfies the properties (a) a,/3eC => a + /?eC; (6) aeC, feR.f^O => faeC. Show that if p,,..., <pn form a basis of (IRn)d then C = {f,a, + • • • + fnan ; f, 6 IR, f, ^ 0} where {a,,..., an} is the basis of IR" dual to the basis {p,,..., (pn).

98 Further Linear Algebra Suppose now that V, W are vector spaces over F and that/ : V —» W is linear. Given any yd e W* consider the linear mapping yd of. The diagram / ** V——»W—-—>F shows that ydof& Vd. Thus the assignment yd •-» >d o/ defines a mapping from IV* to Vd. We call this the transpose of/ and denote it by/'. Then we have /V) = /°/ and so/' can be described as composition on the right by/. It is readily seen that /'eLinJW^.V). • In terms of the notation introduced previously, we have {/wy>=/i/wi=(/ o/)w=[/V)iw=(*./'(/)>. Consequently we have the identity wo,/) = (*,/v». The above use of the word 'transpose' is suggested by the following result. Theorem 7.4 Let V and W be vector spaces over F of dimensions m and n respectively. Let (a,-)m, {b;)n be ordered bases of V, W and let f : V -» W be linear. If the matrix off relative to (a,-)m, (£>,),, is A = [a,;]nxni then relative to the corresponding dual ordered bases (bd)n, (ad)m the matrix off is A'. Proof Let the matrix representing/' be B - [bjj]mxn. Consider the identity {fM^)- (ahf'(bd)). The left hand side is U=l I *=l whereas the right hand side is Thus we have b^ - fl;i- for all /, j and therefore B = A'. D The main properties of transposition are the following. Theorem 7.5 (1) idy= idytf. (2) IfftgeUn(VtW)then tf + s)'=/' + *'. (3) If feUn(VtW)and gelin(WtX)then (gof)' = /'og'.

7. Dual Spaces 99 Proof All are immediate from the fact that for every linear mapping h the effect of h' can be described as composition on the right by h. D Corollary If f:V->W is an isomorphism then so is f iW* -> Vdt and (J')-1 = {fl)'. Proof This follows from (1) and (3) on taking g = /'". □ We can also consider the transpose of/'. We denote this by /" and call it the bitranspose of/. There is a natural connection between bitransposes and biduals. Theorem 7.6 For every linear mapping f: V —► W the diagram is commutative, in the sense that f"oav<=ctwof. Proof We have to show that/"(x) = /(x) for every x 6 V. Now for all yd e V* we have ^(5)1(/)=^0/-)(/) = 5^(/)] = W(/)) = VW,/)=/W(/) from which the result follows. D A consequence of Theorem 7.6 is that when V and W are of finite dimensions (in which case we agree to identify V, V and W, W and therefore also aVt \6V and ctWt \6W) we have/" = /. This then matches the matrix situation, where A" - A. EXERCISES 7.9 If V

is a finite-dimensional vector space prove that a linear mapping f : v -v V is surjective (resp. injective) if and only if its transpose is injective (resp. surjective). 7.10 Let V be a finite-dimensional vector space. Prove that the mapping p:Un(VlV)-»Lin(VrflV'") given by <p{f) = /' is an isomorphism.

100 Further Linear Algebra 7.11 Let/: R3 -» R3 be given by the prescription /(*,>, z) = (2* + y, x + y + z, -z). Given the basis X = {(1,0,0),(1,1,0),(1,1,1)} of R3 and the basis yd = {[1,0,0],[1,1,0],[1,1,1]} of (IRY, determine the matrix of/1 with respect to the bases Yd and Xd. 7.12 Let D : R,,[X] -> Ft be the differentiation mapping. Determine a basis for the kernel of the transpose of D. Definition If x 6 V and yd 6 Vd are such that (x, yd) = 0y then we shall say that x is annihilated by/. Since (x, y*) = y'(jc) we see that the set of elements of V that are annihilated by y6 is Ker y6. Now it is immediate from the identities (/?) and (7) preceding Theorem 7.3 that, for every non-empty subset E of V, the set of elements of Vd that annihilate every element of £ is a subspace of V*. We denote this subspace by E°. Thus £°={yd6Vrf; (VxeE) (*,/) = 0}. We call E ° the annihilator of E. It is clear that {0V}° = V* and that V° = {Oy*}. Theorem 7.7 If V is a finite-dimensional vector space and if W is a subspace of V then W° is also a subspace and dimW°=dimy-dimW. Moreover, identifying V and V, we have W = W00. Proof That W° is a subspace of V is clear. Suppose now that dim V-n. The result is trivial ifW= V so suppose that W f V. Then dim V/-m< n. Let{a,,...,aw} be a basis of W and extend this to a basis {^1,.--,^,^+1,..-,^} n of V. Let {of,..., ad} be the corresponding dual basis. If xd = £ X,-of 6 W° then i=l fory = 1,..., m we have o=(«;,^)=i:m^«?)=v 1=1 It follows that {afn+i,..., oj!} is a basis of W° and consequently dim VV° = n-m- dim V - dim W.

7. Dual Spaces 101 As for the final statement, consider the subspace W00 = (VV°)° of V = V. By definition, every element of W is annihilated by every element of W° and so we have W c W°°. On the other hand, by what we have just proved, dim VT°= n -dim W°= n -{n -ro) = ro = dim W. It therefore follows that W = W°°. D Annihilators and transposes are connected: Theorem 7.8 If V and W are finite-dimensional and iff: V —» W w /mear then (l)(Im/) °=Ker/'; (2)(Ker/)°=Im/'; (3) dim Im/J = dim Im/; (4) dimKer/J = dimKer/. Proof (1) We have / 6 (Im/)° if and only if, for every x e V* 0 = ^),/) = (^(/)) which is the case if and only if/1 (/) 6 V° = {0^}, i.e. if and only if/ 6 Ker/J. (2) Replacing/ by/1 in (1) and using the fact that/" = /, we obtain (Im/')° =

Ker/. Then, by Theorem 7.7, (Ker/)°= (Im/')°°= Im/J. (3), (4) follow from (1), (2), and Theorem 6.7. □ EXERCISES 7.13 Let V be a finite-dimensional vector space and let (V;)l€/ be a family of subspaces of V\ Prove that i£l i'€/ i €/ <€/ 7.14 If V is a finite-dimensional vector space and W is a subspace of V, prove ti\atVd/W°~Wd. [Hint. Show that/ - g 6 W° if and only the restrictions of /,g to W coincide.] 7.15 Prove that if 5 is a subset of a finite-dimensional vector space V then the subspace spanned by S is 5°°. 7.16 If W is the subspace of R4 spanned by {(1,1,0,0),(0,0,1,1)}, find a basis for W°. 7.17 Let A and B be subspaces of a vector space V such that V - A®B and let 7rA, 7¾ be the projections onto A, 5 respectively. Prove that V* = Im7rieIm7ri.

 7.18 By a canonical isomorphism (: V —► V* we mean an isomorphism (such that, for all xty £ V and all isomorphisms/ :V —*Vtwe have the identity (*) (*.C(y)) = {fl*).CMy)l). In this exercise we outline a proof of the fact that if V is of dimension n > 1 over F then there is no canonical isomorphism (: V —» V* except when n = 2 and \F\ = 2. If (is such an isomorphism, show that if y f 0V then the subspace Ker ((>) = {C(>)}°is of dimension n -1. Suppose first that n > 3. If there exists t G Ker ¢(0 for some / ^ 0V, let {f, X|,..., xn_2) be a basis of Ker ¢(0 and extend this to a basis {f, jc, ,..., jcn_2, z} of ^. Let/: V -» y be the linear mapping such that /(0 = ', /(*i) = z> /W-*i. /(*,) = *,('Y0. Show that/ is an isomorphism that does not satisfy (*). [Hint. Takex = jc,,> = f.] If, on the other hand, t $ Ker ¢(0 for all t f Oy let {jc, ,..., *,,_,} be a basis of Ker C(f) so that {*,,..., xn_t, t} is a basis of V. Show that {x,+x2,x2,...,xn_,,f} is also a basis of V. Show further that x2 £ Ker ((*,). Now show that if /: V —» y is the linear mapping such that /(*i) = *2, /(*2) = *.+*2, /(0 = ', /U) = ^,(^1.2) then/ is an isomorphism that does not satisfy (*). Conclude from these observations that we must have n = 2. Suppose now that F has more than two elements and let \ £ F be such that \ f 0,1. If there exists t f 0y such that t £ Ker ((f), observe that {f} is a basis of Ker ((f) and extend this to a basis {f, z} of V. If /: V —» y is the linear mapping such that /« = ', /W = *z show that/ is an isomorphism that does not satisfy (*). [Hint. Take x = z, y = f.] If, on the other hand, f $ Ker ((f) for all f f % let {z} be a basis of Ker ((f) so that {f, z} is a basis of y. If /: y -» y is the linear mapping such that /M = *z. /10 = ' show that/ is an isomorphism that does not satisfy (*). [Hint. Takex-y-z.]

 Conclude from these observations that \F\ = 2. Now examine the vector space F2 where F = {0,1}. [Hint. (F2)d is the set of linear mappings/: F x F -> F. Since F2 has four elements there are 24 = 16 laws of composition on F. Only four of these are linear mappings from F2 to F\ and each of these is determined by its action on the natural basis of F2. Compute (F2)d and determine a canonical isomorphism from F2 to {F2)d] We now pass to the consideration of the dual space of an inner product space. For this purpose, suppose that V is an inner product space and consider the inner product (xty)»-» (x \y). For every y e V the assignment x >-► (x \ y) is linear and therefore is an element of the dual space Vd. We shall write this element as yd, so that we have the following useful amalgamated notation: (x\y) = Ax) = (xJ) ■ As we shall see, V and Vd are related via the following notion. Definition Let V and VV be vector spaces over a field F where F is either IR or C. A mapping /: V —f VV is called a conjugate transformation if (Vjc 6 V) (y\ 6F) f(x + y) = f(x) +/(>), f(\x) = \f(x). If, furthermore, / is a bijection then we say that it is a conjugate isomorphism. Note that when F = IR conjugate transformations are simply linear mappings. Theorem 7.9 If V is a finite-dimensional inner product space then there is a conjugate isomorphism tiv : V -» Vd, namely that given by tiv : x >-► x6. Proof For all x, y, z 6 V we have from which we obtain (y + z)d = / + zd, i.e. dv{y + z) = tf v(y) + <Mz)- Likewise, and so (Xy)" = X/, i.e. ^^(^^) = ^^v(>)- T^s ^vis a conjugate transformation. That dv is injective follows from the fact that if x 6 Ker tiv then xd = 0 and so (jc | jc) = (jc, jcd) = 0 whence x = <V To see that dv is also surjective, let f eVd. If {ex,..., en} is an orthonormal basis of V, let

 Then fory = 1,..., n we have A*j) = (*j I *) = U I EM) = £/(*.)<*; I */> = /(*;)• x I i=I ' i=l Thus jcd and/coincide on the basis {elt...,eH}. Hence/= jr* = tfy(*) and so dv is surjective. Thus tiv is a conjugate isomorphism. D Note from the above that we have the identity (*|y) = (*A(>)>- Since tV is a bijection, we also have (writing tfy'(y) for y) the identity (* I V(>)> = (*,>>. We can now establish the following important result. Theorem 7.10 Let V and W be finite-dimensional inner product spaces over the same field. Then for every linear mapping f: V —*W there is a unique linear mapping f*:W—*V such that Proof With the above notation, we have the equalities (Mb) = (^)^) = <*,/'(/)> = <*IW(/)]> = (x|(Vo/'o^)(y)) from which it follows immediately that /*= V°/'°tf»v is the only linear mapping with the stated property. D Corollary f*: W —*V isthe unique linear mapping such that the diagram W *w » W1 f yd is commutative, in the sense that tivof* = f otiw. □ Definition The unique linear mapping of Theorem 7.10 will be called the adjoint of/.

 Example 7.10 Consider the vector space V = Mat,,xn C with the inner product (A\B) = trB*A where B* denotes the transpose of the complex conjugate of B. For every M £ V the mapping 0M : V -» V given by @M(A) = MA is clearly linear. Since G%# M) IB) = tr B*MA = tr(ATO)M = (A \ M*B) = {A \ MB)) we see that the adjoint of 0M exists and is /¾^.. EXERCISES 7.19 Let V be a complex inner product space. Forallјt,yeyiet/xv: V -* V be given by (VzGV) fx,y(z) = (z\y)x. Prove that/ is linear and that (a) (Vx.y.zGVO/^o/.^W2/,/, (b) the adjoint of/iy is/yJt. 7.20 If V is an inner product space then a linear mapping /: V —» V is said to be self-adjoint if/* exists and /* - /; normal if/* exists and /°/* = /*o/. Prove that if x f 0V and y f 0V then the linear mapping /x v of the previous exercise is (a) normal if and only if there exists \ G C such that x = \y\ (b) self-adjoint if and only if there exists \ 6 R such that x = \y. 7.21 Let C[0,1] be the inner product space of real continuous functions on [0,1] with the integral inner product if \ 8)= ffg- Jo Let K : C[0,1] -» C[0,1] be the integral operator given by *(/)= fxyf(y)dy. Jo Prove that K is self-adjoint. 7.22 If/: IR3 -»IR3 is given by f(x,y,z) = {x + ytytx+y + z)t determine/*.

 7.23 Let V be a finite-dimensional inner product space. Prove that for every / 6 V there is a unique 0 e V such that (Vxev) f(x) = (x\p). [Hint. Let {a,,..., a,,} be an orthonoimal basis of V and consider the n element 0 = £/(<*>,.] i=l Find such a)9 when y = R3 and f[xt y, z) = x - 2y + 4z. 7.24 Show as follows that the result of the previous exercise does not hold for an inner product space V of infinite dimension. Let z 6 C be fixed and define the 'evaluation at z map/z e Vd by (tyev) /» = p(z). Show that noqeV exists such that (Vpev)/<(p) = (H<?>- [Hint. Suppose that such a q exists and let r 6 V be given by r = X - z. Show that, for every p 6 V\ rp?=0. Now take p- fq and deduce the contradiction <? = 0.] 7.25 Let y = C[X], the vector space over C of polynomials with complex coefficients. Show that Ip,Q)"{p\q) = / P? Jo is an inner product on V. If P = E fl*x* ^ ^ define p = £ apt*. Let/, : V - y be given by (v*ey) /,(<?) = /*?. Show that f£ exists and is/p. Now let D: V -> V be the differentiation mapping. Show that D does not admit an adjoint. [Hint. Suppose that D* exists and show that (p\Dq + D>q) = p(\)q(\)-p{0)q(0)- Suppose now that q is a fixed element of V such that ¢(0) = 0 and q(\) = 1. Use the previous exercise with z = 1 to obtain the required contradiction.] .

7. Dual Spaces 107 The principal properties of the assignment/ »-► /*are described in the following two results. Theorem 7.11 Let Vt Wt X be finite-dimensional inner product spaces over the same field and let f,g:V -» W and h:W-*Xbe linear mappings. Then (1) (/"+*)*=/*+**; (2) (*/)*= \r; (3) (hof)* = roh* (4) (/y = /. Proof (1) is immediate from/* = /^1 o/r o /^ and the fact that {f + g)'=f' + g'. (2) (\f(x) | y) = \ (f(x) | >)_= X (x | /*(>)) = (* I \TM) and therefore, by the uniqueness of adjoints, (X/)*= X/*. (3) (h[f{x)] | >) = {f{x) | **(>)) = (jc | /Wy)]) and therefore, by the uniqueness of adjoints, (hof)* = f*o h*. (4) Taking complex conjugates in the identity of Theorem 7.10 we obtain the identity (f(y)l*) = (y|/W>, from which it follows by the uniqueness of adjoints that (/**)* = /. D Theorem 7.12 Let V and W be finite-dimensional inner product spaces over the same field with dim V = dim W. If f :V —*W is linear then the following are equivalent: (1) f is an inner product space isomorphism; (2) / is a vector space isomorphism and /"' = /*; (3) /o/*- idw; (4)ro/=id^ Proof (1) =^ (2): If (1) holds then/-1 exists and we have the identity V{x)\y) = W*)\f\rx{y)\) = (x\r\y)) from which it follows by the uniqueness of adjoints that/-1 = /*. It is clear that (2) =*► (3) and that (2) =*► (4). (4) => (1): If (4) holds then/ is injective, hence bijective, and/"1 = /*. Thus (Vxty<EV) (f(x)\f(y)) = (x\f[f(y))) = (*\y) and so/ is an inner product isomorphism. The proof of (3) =>• (1) is similar. D

108 ___ Further Linear Algebra EXERCISES 7.26 If V and W are finite-dimensional inner product spaces over the same field, prove that the assignment/1-> /* defines a conjugate isomorphism from Lin (V, IV) to Lin (W,V). 7.27 Let V be a finite-dimensional inner product space and let/: V -► V be linear. If W is a subspace of V, prove that IV is/-invariant if and only if W1 is/*-invariant. We have seen how the transpose/' of a linear mapping/ is such that Ker/' and Im/' are the annihilators of Im/ and Ker/ respectively. In view of the connection between transposes and adjoints, it will come as no surprise that Ker f* and Im/* are also related to the subspaces Im/ and Ker/. Theorem 7.13 If V is a finite-dimensional inner product space and iff: V —► V is linear then Irnr=(Ker/y-, Ker/* = (Im/)1. Proof Let z 6 Im/*, say z = /*(y). Then for every x 6 Ker/ we have (x\z)-(x\rb))-V(x)\y) = (Ov\y) = o and consequently z 6 (Ker/)1. Thus Im/*C (Ker/)1. Now let y 6 Ker/*. Then for z = /(jc) 6 Im/ we have (z|y>-{fljr)|y>-(jr./ty)>-(jr|ov>-o and consequently y 6 (Im/)1. Thus Ker/*C (Im/)1. Using Theorem 2.7, we then have dimlm/ = dim V-dimflm/)-1- ^ dim y- dim Ker/* = dimlm/* ^ dim(Ker/)1 = dim V-dim Ker/ =

dimlm/. The resulting equality gives dim Im/* = dim (Ker /)1, dim (Im/)1 = dim Ker /*, from which the result follows. D

7. Dual Spaces 109 We now investigate the natural question of how the matrices that represent/ and f* are related. Definition UA = lau] 6 Mat mxn C then by the adjoint (or conjugate transpose) of A we mean the n x m matrix v4*such that [A%j - aj[. Theorem 7.14 Let V and W be finite-dimensional inner product spaces over the same field. If, relative to ordered orthonormal bases (v,)n and (w,)m a linear mapping f:V-*W is represented by the matrix A then the adjoint mapping f*:W -*Vis represented, relative to the bases (w^ and (v-),,, by the adjoint matrix A*. Proof m For) = 1,..., n we have, by Theorem 1.5, f(vj) - J^{f{vj) | w,)*,. Thus, if A = i=l n [aij] we have at] - (fty) \ w{). Likewise, we have/*(w;) = £(/^) | v-)v,. Then, since it follows that the matrix that represents f* is A*. D EXERCISE 7.28 If A 6 Matnxn F prove that det A*= deFX It is clear from Theorems 7.12 and 7.14 that a square matrix M represents an inner product space isomorphism if and only if A/-1 exists and is M*. Such a matrix is said to be unitary. It is readily seen by extending the corresponding results for ordinary vector spaces to inner product spaces that if A, B are n x n matrices over the ground field of V then AtB represent the same linear mapping with respect to possibly different ordered orthonormal bases of V if and only if there is a unitary matrix U such that B = U*AU = U~lAU. We describe this situation by saying that B is unitarily similar to A. When the ground field is R, the word orthogonal is often used instead of unitary. In this case A/ is orthogonal if and only if A/-' exists and is A/', and B is said to be orthogonally similar to A. It is clear that the relation of being unitarily (or orthogonally) similar is an equivalence relation on the set of n x n matrices over C (or IR). Just as with ordinary similarity, the problem of locating particularly simple representatives (or canonical forms) in certain equivalence classes is important from both the theoretical and practical points of view. We shall consider this problem later.

8 Orthogonal Direct Sums In Theorem 2.6 we obtained, for an inner product space V and a finite-dimensional subspace W of V, a direct sum decomposition of the form V = W © W1 We now consider the following general notion. Definition Let V,,..., Vn be non-zero subspaces of an inner product space V. Then V is said to be the orthogonal direct sum of V,,..., Vn if (1) V=0V,.; (2)(i = 'r'...,n) V^EVy. ^•' In order to study orthogonal direct sum decompositions in an

inner product space V let us begin by considering a projection p:V—>V and the associated decomposition V- Imp©Kerp as established in Theorem 2.4. For this to be an orthogonal direct sum according to the above definition, it is clear that p has to be an ortho-projection in the sense that Ker p = (Imp)1, Imp = (Ker p)1. Note that, by Theorem 2.7, when V is of finite dimension either of these conditions will do. Example 8.1 As observed in Example 2.1, we have IR2 = Y © D where Y = {(0,>) ; y e IR} and D - {(jc, x) ; x 6 R}. Let p be the projection on D parallel to Y. Then for all (jt.y) e IR2 we have p{xty) - {x,x)t so that Imp = D and Kerp = Y. But relative to the standard inner product on IR2 we have YL - {(xt 0); x 6 IR}. Hence (Ker p)1 f Imp and so p is not an ortho-projection.

8. Orthogonal Direct Sums 111 In order to discover precisely when a projection is an ortho-projection, we require the following result Theorem 8.1 If W and X are subspaces of a finite-dimensional inner product space V such that V=W@Xthen V= W1©^. Proof By Theorem 2.8, we have {0V} = VX = (W + X)1 = W1DX1 and V =(0^ = (^ HX)1 = W1 + X1 whence we see that V = W1 © X1. D Corollary If V - W © X and if pis the projection on W parallel to X then p* is the projection on X1 parallel to W1. Proof Since p is idempotent it follows by Theorem 7.11(3) thatp*op*= (pop)*= p*. Hence p* is idempotent and so is the projection on Imp*parallel to Ker p*. By Theorem 2.3, Imp = W and Ker p = X, so by Theorem 7.13 we have W1 = [Imp)1 - Ker p*and X1 = (Kerp)-L = Imp*. D Definition If V is an inner product space then a linear mapping /: V -► V is said to be self- adjoint if/ = /*. We can now characterise ortho-projections. Theorem 8.2 Let V be a finite-dimensional inner product space. Ifp:V-*Visa projection then p is an ortho-projection if and only if pis self-adjoint. Proof By the Corollary of Theorem 8.1, p*is the projection on Imp* = (Ker p)1 parallel to Ker p*= (Imp)1. Thus, if p is an ortho-projection we must have Imp*= Imp. It then follows by Theorem 2.3 that for every x 6 V we have p{x) = p^pix)). Consequently p = p* o p and therefore, by Theorem 7.11 (3)(4), p*=(p*op)*=p*op** = p*op = p. Hence p is self-adjoint.

112 Further Linear Algebra Conversely, if p - p* then clearly Imp = Imp* = (Ker p)1, which shows that p is an ortho-projection. D If now W is a finite-dimensional subspace of an inner product space V then by Theorem 2.6 we have V = W&W1. Now Theorem 8.1 and its corollary guarantee the existence

of an ortho-projection on W. There is in fact precisely one ortho-projection on W. To see this, suppose that p and q are ortho-projections on W. Then Imp = W - Imq and Ker p = (Imp)1 = (Im^)1 = Ker q. Since projections are uniquely determined by their images and kernels, it follows that p- ^. We may therefore talk of the ortho-projection on W. It may be characterised as follows. Theorem 8.3 Let W be a finite-dimensional subspace of an inner product space V. If Pw is the ortho-projection on W then for every x^V the element pw(x) is the unique element of W that is nearest x, in the sense that (V*6W0 Ix-ArMUIx-wl. Proof Since pyy is idempotent, for every w^Wwe have that x - pw{x) 6 Ker pw - W1. Consequently, \\x-M\2=\\Pw(x)-W + x-Pw(x)\\i = \\Pw(x)-w\\2+*- Pw(x)\\2 >\\x-Pw(x)\\2 whence the required inequality follows. D Corollary If {et,..., en} is an orthonormal basis of W then (Vx<EV) p*(*) = £(*k>,. Proof This is immediate from the above and the discussion that precedes Example 2.9. D EXERCISES 8.1 In IR4 consider the subspace W - Span {(0,1,2,0), (1,0,0,1)}. Determine the ortho-projection of (1,2,0,0) on W. 8.2 Let W be a finite-dimensional subspace of an inner product space V. Prove that p^x = idy - p^.

8. Orthogonal Direct Sums 113 8.3 Let V be a finite-dimensional inner product space. If W is a subspace of V, prove that (v*ev) IM*)K||*||. 8.4 In IR2 consider a line L passing through the origin and making an angle tf with the x-axis. For every point (jc, y) we have thatpL(jt, y) = (jc*, y*) where (**, >*) is the foot of the perpendicular from (jc, y) to L. Show by simple geometrical considerations that cos2 tf sin tf cos tf sin tf cost? sin2tf 8.5 Consider the linear mapping/: R3 -» R3 given by f{x>ytz) = {-ztx + zty + z). Determine the ortho-projections on the eigenspaces of/ Concerning direct sums of subspaces in an inner product space we have the following result. Theorem 8.4 Let V be a finite-dimensional inner product space and letVlt...,Vkbe subspaces k of V such that V = ® Vt. Then this direct sum is an orthogonal direct sum if and i=i only if, for each i, every element of Vk is orthogonal to every element of Vj (J f i). Proof The necessity is clear. As for sufficiency, if every element of V) is orthogonal to every element of Vy 0' f 0 men we nave Y, Vj £ *V"- But if' dim ^ V} = dim V - dim V, = dim V/\ it* Consequently J2^ = V/-. D Definition Let V be a finite-dimensional inner product space and let/ : V -► V be linear. Then we say that/ is ortho-diagonalisable if there is an orthonormal basis of V that consists of eigenvectors of/; equivalently, if there is an orthonormal basis of V with respect to which the matrix of/ is diagonal. Our objective now is to determine

under what conditions a linear mapping is ortho-diagonalisable. In purely matrix terms, this problem is that of determining when a given square matrix (over IR or C) is unitarily similar to a diagonal matrix.

114 Further Linear Algebra For this purpose, we recall that if/ is diagonalisable then by Theorem 3.3 its minimum polynomial is nfr = (X-A,)(X->2)...(X-AJ where >,,...,Xk are the distinct eigenvalues of/. Moreover, by Corollary 2 of it Theorem 3.2 we have that V=0VX(where the subspace VAi = Ker(/" - >,idy) is f=i the eigenspace associated with >,. If now p,-: V -» V is the projection on VXi parallel * to £ Vx then from Theorem 2.5 we have £p,- = idv and p, o py = 0 when i fj. ifi i=l Consequently, for every x 6 V, we see that /(*) =/(£pM) = tflpM) = X>,aW = (E>/p,)W. i=l i=l i=l /=1 it which gives/ = £ >,p/. i=l As we shall now show, if V is an inner product space then the condition that each Pi be self-adjoint is the key to/ being ortho-diagonalisable. Theorem 8.5 Let V be a non-zero finite-dimensional inner product space and let f : V —► V be linear. Then f is ortho-diagonalisable if and only if there are non-zero self-adjoint projections p,,..., pk : V —> V and distinct scalars >,,...,>* such that i=i (2)£> = idv; i=i (3)(i^) Pi°py = 0. Proof it =»: It suffices to note that if 0 VAj is an orthogonal direct sum then each pt is /=i an ortho-projection and so, by Theorem 8.3, is self-adjoint. k *= : If the conditions hold then by Theorem 2.5 we have V = 0 Imp,-. Now the i=l >,- that appear in (1) are the distinct eigenvalues of/. To see this, observe that f°Pj = (£*ift) °Pj = EMft°Pj) = V; i=i /= i and so (J - \j'\dv) opj-0 whence {0V} 7* Imp, C Kertf - \idv). Thus each Xj is an eigenvalue of/.

8. Orthogonal Direct Sums 115 On the other hand, for every scalar A, /->idv=x: \iPi - e \Pi=£(>, - \)Ph i=l r=l i=| so that, if x is an eigenvector of/ associated with the eigenvalue \, we have k E(\->)PiW = 0v. k But V = 0 Impj, and so we deduce that (>,- - >)p,(jt) = 0V for / = 1,..., k. If now i=i \ f \t for every i then we would have p,(x) = 0V for every i, whence there would k followthecontradiction x - EftOO - %■ Consequently \ - >, for some/. Hence i=i \ |,..., \k are the distinct eigenvalues of/. We now observe that, for eachy, Imp; = KeT(f-\j\dv). k For, suppose that x G Ker (/"->, idy), so that/(x) = \jx. Then 0V = E(*i-\)aM i=i and therefore (>, - ^)ft(x) = 0V for all i whence p,{x) = 0V for all i £j. Then k ^ = EA(jr) = Py(jr)eImP> and so Ker(/" - >,idv) C Impy. The reverse inclusion was established above. k k Since V = 0 Imp; = 0 Ker(/" - >,idy) it follows that V has a basis consisting i= I i= I of eigenvectors

of/ and so/ is diagonalisable. Now by hypothesis the projections ps are self-adjoint so, for j f i, {pM i ?, w> = (pM i #w> = toiPiM] i x> = (ov i x)=o. It follows that the above basis of eigenvectors is orthogonal. By normalising each vector in this basis we obtain an orthonormal basis of eigenvectors. Hence/ is ortho- diagonalisable. D Definition k For an ortho-diagonalisable mapping/ the equality / = E ^iPi of Theorem 8.5 is /=i called the spectral resolution of/. Applying the results of Theorem 7.11 to the conditions in Theorem 8.5 we obtain (i*) /*=E\p?=X>a; (2-) = (2), (3-) = (3). i=l i=l

 It follows immediately by Theorem 8.5 that/* is also ortho-diagonalisable and that (14) gives its spectral resolution. Note as a consequence that >|,..., \k are the distinct eigenvalues of/*. A simple calculation now reveals that /°r=D>.i2Pi=r°/ i=\ from which we deduce that ortho- diagonalisable mappings commute with their ad- joints. This observation leads to the following notion. Definition If V is a finite-dimensional inner product space and if/: V -► V is linear then we say that/ is normal if it commutes with its adjoint. Similarly, a square matrix A over a field is said to be normal if AA* - A*A. Example 8.2 For the matrix A 6 Mat2x2 C given by we have Then A = A* = 2 i 1 2 2 1 -i 2 A4* = 5 2 + 2/ 2-2i 5 = A*A and so A is normal. EXERCISES 8.6 If the matrix A is normal and non-singular prove that so also is A '. 8.7 If A and B are real symmetric matrices prove that A + iB is normal if and only if A and B commute. 8.8 Show that each of the matrices 1 -i / 1 , B = 0 i i 0 is normal, but neither A + B nor AB is normal.

 8.9 Let V be a complex inner product space and let/ : V -► V be linear. Define /. = #+/*), A=w-n Prove that (1) /, and/2 are self-adjoint; (2)/ = /,+&; (3) if/ = g, + ig2 where g, and g2 are self-adjoint then gt = /, and £2= fo (4) / is normal if and only if/, and/2 commute. 8.10 Prove that the linear mapping/: C2 -► C2 given by /(a, /3) = (2a + //?, a + 2/3) is normal and not self-adjoint. 8.11 Let V be a finite-dimensional inner product space and let/: V -* V be linear. Prove that if a, 0 6 C are such that |a| = |/3| then a/ + 0f* is normal. 8.12 Show that A 6 Mat 3x3 C given by A = l + i -1 2-i -i 2 + 2i 0 -1 + 2/ 0 3 + 3i is normal. We have seen above that a necessary condition for a linear mapping to be ortho- diagonalisable is that it be normal. It is quite remarkable that when the ground field is C this condition is also sufficient. In order to establish this, we require the following properties of normal mappings. Theorem 8.6 Let V bea non-zero finite-dimensional inner product space and let f: V ->Vbea linear mapping that is normal. Then 0)(V* €V) lfl*)l = lfMII: (2) p{f) is normal for every polynomial p\ (3) Im/nKer/= {0v}. Proof (1) Since/ o/*=/*o/ we have, for all x e V, </w i/w> = <* i/vw» = (* \fin*))) = (fw \rt*)) from which (1) follows.

 n n (2) If p = £ fljX" then we have p{f) = Y,af and lt follows by Theorem 7.11 i=0 i=0 n that [p(/0]* = £ OfCD'- Since/ and/* commute, it follows that so do p{f) and <=o Wffl* Hence p(f) is normal. (3) If x e Im/n Ker/ then there exists y 6 V such that jc = /(>) and/(x) = 0V. By (1) we have/*(x) = 0V and therefore o = (f(x)\y) = (x\f(y)) = (x\x) whence x - 0V. D We now characterise the normal projections. Theorem 8.7 Let V be a non-zero finite-dimensional inner product space. If p : V —* V is a projection then p is normal if and only if it is self-adjoint. Proof Clearly, if p is self-adjoint then p is normal. Conversely, suppose that p is normal. By Theorem 8.6 we have \\p{x)\\ = ||p*(x)|| foreveryxG Vandsop(x) = Ovifandonly if p*(x) = 0V. Now, given x 6 V* let y = x -p(jt). We have p{y) = p{x) -p(x) = Oy and so 0V = p*{y) = p*[x) -p*\p{x)] which gives p* - p* o p. Thus, by Theorem 7.11, p - p** = (p* o p)* = p* o p** = p* o p = p*. i.e. p is self-adjoint. □ We can now solve the ortho-diagonalisation problem for complex inner product spaces. Theorem 8.8 Let Vbea non-zero finite-dimensional complex inner product space. Iff : V —* V is linear then f is ortho-diagonalisable if and only if it is normal. Proof We have already seen that the condition is necessary. As for sufficiency, suppose that/ is normal. To show first that/ is diagonalisable it suffices to show that its minimum polynomial mf is a product of distinct linear factors. For this purpose, we shall make use of the fact that the field C of complex numbers is algebraically closed, in the sense that every polynomial of degree at least 1 can be expressed as a product of linear polynomials. Thus mf is certainly a product of linear factors. Suppose, by way of obtaining a contradiction, that a £ C is such that X -a is a multiple factor of mf. Then mf = (X - a)2g for some polynomial g. Thus for every x 6 V we have 0, = [mf{f))(x) = [{f -<*idv)2 o g[f))(x) and consequently [(/" - oridv) o g{f)](x) belongs to both the image and the kernel of/ - oridy. Since, by Theorem 8.6(2),/- \6V is normal we deduce from Theorem 8.6(3) that (V*6V) [(f-"idv)o5(/)]W = 0v.

 Consequently {f-aidv) og{f) is the zero mapping on V, so/ is a zero of the polynomial (X -a)g. This contradicts the fact that (X -ct)2g is the minimum polynomial of/. Thus we see that/ is diagonalisable. To show that/ is ortho-diagonalisable, it suffices to show that the corresponding projections pt onto the eigenspaces are ortho-projections, and by Theorem 8.2 it is enough to show that they are self-adjoint. k Now since/ is diagonalisable we have/ = £ A^p,-. Since nop; = 0 for i f /, it is /= i * * readily seen that/2 = £>2fl. A simple inductive argument now gives/" = £>"# i=l i=l for all n ^ 1. An immediate consequence of this is that for every polynomial g we have i=l Consider now the Lagrange polynomials Llt...tLk associated with \,,..., \k. We recall (see Example 7.6) that If} Aj *i Taking g = Lj we see from the above that k k *•/(/) = E LMi)Pi = E fyfi = Py i=l i=l It therefore follows by Theorem 8.6(2) that each p} is normal and so, by Theorem 8.7, is self-adjoint. □ Corollary A square matrix over C is unitarily similar to a diagonal matrix if and only if it is normal. D EXERCISES 8.13 If A is a normal matrix show that so also is g[A) for every polynomial g. 8.14 Prove that A*= p{A) for some polynomial p if and only if A is normal. 8.15 For each of the following linear mappings determine whether or not it is ortho-diagonalisable: 0)/ (2)/ (3)/ (4)/ C2 -► C2; /(jc,y) = (x +1>, -ix + y). C2-C2; /tjfiy) = (jf + (I+i)y,(l-^+y). C3-C3; /(^,2)=(7,2,/(0: + 7 + 2)). C3-C3; /(jc,y,2)=(jc + i>,-ijc+y + (l-i)z,(l + 0y + z)-

2 . 2 -*1 -*J 120 Further Linear Algebra It should be noted that in the proof of Theorem 8.8 we made use of the fact that the Meld C of complex numbers is algebraically closed. This is not true of IR and so we might expect that the corresponding result fails in general for real inner product spaces (and real square matrices). This is indeed the case: there exist normal linear mappings on a real inner product space that are not diagonalisable. One way in which this can happen is when all the eigenvalues of the mapping in question are complex. Example 8.3 In the real inner product space R2 consider the matrix ^2*/3 = which represents an anti-clockwise rotation through an angle 2n/3 of the coordinate axes. It is readily seen that /?2*/3xs normal- But its minimum polynomial is X2 +X+1, and this has no zeros in IR. Hence this matrix is not diagonalisable over IR. Thus, in order to obtain an analogue of Theorem 8.8 in the case where the ground Meld is IR we are led Mrst to the consideration of normal linear mappings whose eigenvalues are all real. These can be characterised as follows. Theorem 8.9 Let V be a non-zero finite-dimensional complex inner product space. Iff : V —» V is linear then the following conditions are equivalent: (1) f is normal and all its eigenvalues are real; (2) / is self-adjoint. Proof k (1) => (2): By Theorem 8.8,/ is ortho-diagonalisable.

Let/ = £ X,p, be its spectral resolution. We know that/* is also normal, with spectral resolution/* = k _ £ \jP;. Since each \: is real, it follows that/*=/ and so/ is self-adjoint. i'=i k k _ (2) => (1): If (2) holds then/is normal. If then/= £ \p, and/*= X>,P, i=i i=i k are the spectral resolutions of/ and/* then/ = /* gives £(X, - X,)p, = 0 and so k _ YX^i ~ ^i)pM - 0v f°r everv x e ^ wnence (^i - \)P, = 0 for every i since i=i k V = 0 Imp,. Since none of the p, can be zero, we deduce that X, = A, for every i. i=i Thus every eigenvalue of/ is real. D

8. Orthogonal Direct Sums 121 Corollary All the eigenvalues of a self-adjoint matrix are real. D EXERCISES 8.16 A complex matrix A is such that A*A = -A. Prove that A is self-adjoint and that its eigenvalues are either 0 or -1. 8.17 Prove that a complex matrix is unitaiy (A* = A~l) if and only if it is normal and every eigenvalue is of modulus 1. The analogue of Theorem 8.8 for real inner product spaces is the following. Theorem 8.10 Let V be a non-zero finite-dimensional real inner product space. If f : V —* V is linear then f is ortho-diagonalisable if and only if it is self-adjoint. Proof k =» : If/ is ortho-diagonalisable let/ = £ \p, be its spectral resolution. Since i=i the ground field is IR, every A, is real. Thus, taking adjoints and using Theorem 8.2, we obtain/*=/. <=: Conversely, suppose that/ is self-adjoint and let A be the (n x n say) matrix of/ relative to some ordered orthonormal basis of V\ Then clearly A is symmetric. Now let /* be the linear mapping on the complex inner product space C" whose matrix relative to the standard (orthonormal) basis is A. Then/1* is self-adjoint. By Theorem 8.9, the eigenvalues of f* are all real and, since/4" is diagonalisable, the minimum polynomial of /*" is a product of distinct linear polynomials over IR. Since this is then the minimum polynomial of A, it is also the minimum polynomial of/. Thus we conclude that/ is diagonalisable. That it is in fact ortho-diagonalisable is established precisely as in the proof of Theorem 8.7. D Corollary If A is a square matrix over IR then A is orthogonally similar to a diagonal matrix if and only if A is symmetric. D Example 8.4 By the above corollary, the real symmetric matrix 4 2 2 A = 2 4 2 2 2 4

122 Further Linear Algebra is orthogonally similar to a diagonal matrix D. Let us find an orthogonal matrix P such that P~lAP = D. It is readily seen that the characteristic polynomial of A is cA = (X-2)2(X-%)t so the eigenvalues of A are 2 and 8. A basis for the eigenspace associated with the eigenvalue 2 is ' t -1" 1 0 1 -1" 0 1 1 J Applying the Gram-Schmidt process, we see that an

orthonormal basis for this eigenspace is I 7? 0 v L i •< As for the eigenspace associated with the eigenvalue 8, a basis is ' < k, rii i [ij > j so an orthonormal basis of this eigenspace is 75 11.75. Pasting these orthonormal bases together, we see that the matrix /> = 7? "75 75 0 7S 75. is orthogonal and such that P~lAP = diag{2,2,8}. EXERCISE 8.18 Find an orthogonal matrix P such that P~lAP is diagonal where A = Hence find a matrix B such that Er = A 4i 0 -3i"> 0-10 3 _

8. Orthogonal Direct Sums 123 We shall now derive a useful alternative characterisation of a self-adjoint linear mapping on a complex inner product space. In order to do so, we require the following result. Theorem 8.11 Let V be a complex inner product space. If f : V —► V is linear and such that (f{x) | x) = Ofor allxeV then / = 0. Proof Using the condition we see that, for all y, z € V, 0=(«y + «)|y + «> = y(y)|«) + yW|y); 0 = if(iy + z) | iy + z) = i(f(y) \ z) - i(f(z) | y)t from which it follows that (f(y) \ z) = 0. Consequently,/(y) = 0V for all y 6 V and therefore/ = 0. D Theorem 8.12 Let V be a complex inner product space and let f : V -+ V be linear. Then the following statements are equivalent: (1) / is self-adjoint; (2)(WxeV) y(jc)|jc)6R Proof (1) => (2): If/ is self-adjoint then for every x 6 V we have from which (2) follows. (2) =>• (1): If (2) holds then for every x 6 V we have (fw i *) = (* 1/(^)) = vww = ow I ^) and consequendy Since this holds for all x 6 V it follows by Theorem 8.11 that/* = /. D This characterisation now leads to the following notion. Definition If V is an inner product space then a linear mapping/: V -> V is said to be positive if it is self-adjoint and such that (f(x) \x)^0 for every x 6 V\ and positive definite if it is self-adjoint and such that (f{x) \ x) > 0 for every non-zero x 6 V.

124 Further Linear Algebra Example 8.5 Consider the linear mapping f : f\2 — > f\2 whose matrix, relative to the standard ordered basis, is the rotation matrix cos tf sin tf -sintf cost? i.e./ is given by /(xty) = (xcostf+ysin tf, -xsin tf+ycostf). For/to be positive it is necessary that A be symmetric, so that sin ti = 0; i.e. tf = /m for some integer n. Also, we must have (f(xty)\(x,y)) = (*2 +y2)cosOO. This requires (4* - l)f < tf < (4* + 1)|. Consequently we see that/ is positive if and only if tf = 2/i7r for some integer n, i.e. if and only if/ = id. Theorem 8.13 If V is a non-zero finite-dimensional inner product space and f : V —* V is linear then the following statements are equivalent: (1) f is positive \ (2) / w self-adjoint and every eigenvalue is real and greater than or equal to

0; (3) there is a self-adjoint g £ Lin (V, V) such that g2 = /; (4) there exists h 6 Lin (V, V) such that h^oh-f. Proof (1) =» (2) : Let \ be an eigenvalue of/. By Theorem 8.9, \ is real. Then 0 < (f{x) | x) = (\x | x) = X(x | jc) gives X ^ 0 since (x | x) > 0. (2) => (3): Since/ is self-adjoint it is normal and hence is ortho-diagonalisable. k Let its spectral resolution be/ = £ X,p, and define g : V -► V by i=i i=l Since thep,- are ortho-projections, and hence are self-adjoint, we see from (2) that g is self-adjoint Also, sincep, o py = 0 when i ^y, it follows readily that g2 = /. (3)=>(4):Take/i = g. (4) => (1): Observe that {h*o h)* = h*o h** = h*o h and that, for all jc 6 V. (h\h(x)\ | jc) = (h{x) | /i(jc)> ^ 0. Thus we see that/ = /i*o /i is positive. D It is immediate from Theorem 8.13 that every positive linear mapping has a square root. We shall now show in fact that/ has a unique positive square root. Theorem 8.14 Let f be a positive linear mapping on a non-zero finite-dimensional inner product space V. Then there is a unique positive linear mapping g : V —► V such that g2 = /. Moreover, g = q{f) for some polynomial q.

8. Orthogonal Direct Sums 125 Proof ft Let/ = £ ^iPi be *he spectral resolution of/ and define g as before, namely /=1 Since this must be the spectral resolution of g, it follows that the eigenvalues of g are y^ for i = 1,..., k. It follows by Theorem 8.13 that g is positive. Suppose now that h : V -*■ V is also positive and such that h2 = /. If the spectral resolution of h is £ /iyfy where the qj are orthogonal projections then we have '=■ £\p, =/= a2 = !>>%• 1=1 7=1 Now, as in the proof of Theorem 8.5, the eigenspaces of/ are the subspaces Imp, for i = 1,... tkt and also Imfy for,/ = 1,.. .m. It follows that m = it and that there is a permutation cr on {1 k} such that qQ^ = p,, whence we have that /i^ = X,. Thus /i0(,) = V\ and we deduce that h - g. For the final statement, consider the Lagrange polynomials it* *' Ai Since L,(\) = 6,-,, the polynomial £ = £ \AA is then such that q{f) = g. D i'=i There is a corresponding result to Theorem 8.13 for positive definite linear mappings, namely: Theorem 8.15 If V is a non-zero finite-dimensional inner product space and iff: V —* V is linear then the following statements are equivalent: (1) fis positive definite; (2) / is self-adjoint and all the eigenvalues off are real and strictly positive; (3) there is an invertible self-adjoint g such that g2 = /; (4) there is an invertible h such that h*oh = f Proof This follows immediately from Theorem 8.13 on noting that g is invertible if and only if 0 is not one of its eigenvalues. D Corollary If f is positive definite then f is invertible. D

126 Further Linear Algebra Of course, the above results have matrix analogues. A matrix that represents a positive (definite) linear mapping with respect to some ordered orthonormal basis in a finite-dimensional inner product space is called a positive (definite) matrix. A positive matrix is often also called a Gram matrix. By Theorem 8.13, we have the following characterisation. Theorem 8.16 For a square matrix A the following statements are equivalent: (1) A is a Gram matrix; (2) A is self-adjoint and all its eigenvalues are greater than or equal to 0; (3) there is a self adjoint matrix B such that B2 - A\ (4) there is a matrix C such that C*C = A D Corollary A real symmetric matrix is positive definite if and only if all its eigenvalues are real and strictly positive. D EXERCISES 8.19 Let/ : R3 -► R3 be the linear mapping that is represented, relative to the standard ordered basis, by the matrix £ -1 /i = r 3 2 i 2 2 -1 -1 I -1 Prove that/ is positive definite. 8.20 If V is a finite-dimensional inner product space and if /, g : V —► V are positive definite, prove that so also is/ + g. 8.21 If V is a finite-dimensional inner product space and if ftg : V —> V are positive definite, prove that/o g is positive definite if and only if/and g commute. 8.22 Let V be a finite-dimensional vector space with a given inner product (-1 -). Prove that every inner product on V can be expressed in the form if (-) | -) for a unique positive definite/6 Lin {V, V). [Hint. Let ((-1 -)) be an inner product on V. Use Exercise 7.23 to produce, for the linear form ((-1/3)) 6 Vdt a unique element 0' £V such that ((-10)) = (-1 ff). Define/ :V -+V by/(0) = 0'.]

9 Bilinear and Quadratic Forms In this chapter we shall apply some of the previous results in a study of certain types of linear forms. Definition Let V and W be vector spaces over a field F. A bilinear form on V x W is a mapping /: V x W - Fsuch that, for all x 6 V, all y 6 W, and all X e F, (l)/(* + *\y) = /(*,?)+/(*'.?); (2)/(*,? + /) = /(*,y)+/(*,/); (3)/{>*.>) = */{*.>) =/(*.*>). Example 9.1 The standard inner product on IR", namely /((*, jcn), (y, y,,)) = £ *,>, is a bilinear form on R" x IR". Example 9.2 Let V be the vector space of real continuous functions. Let K(x,y) be a given continuous function of two real variables and let a,b 6 IR- Define H : V x V -> \R by "M)= f f K(xty)g(x)h(y)dxdy. Then standard properties of integrals show that H is bilinear. EXERCISE 9.1 Determine if /: IR2 x IR2 -► IR is bilinear when f({x 1,^1),(^2.^2)) is (1) xiyi-x2yx\ (2) (*,+>i)2 -*2>2-

128 Further Linear Algebra Definition Let V be a vector space of dimension n

over F and let (v,)n be an ordered basis of V. If/: V x V —► F is a bilinear form then by the matrix of / relative to the ordered basis (v,)n we shall mean the matrix A = [a{j]nxn given by aki - /(v,, vy). Suppose now that V is of dimension n and that/: Vx V -► F is bilinear. Relative n n to an ordered basis (v,)n let x = £ x,v, and y = £ y^. If the matrix of/ relative i=i i=i to (v,)n is A = [a,j] then, by the bilinearity of/, we see that n n n i=ly=l i,y=l Conversely, given any nxn matrix /l = [a,y] over F it is easy to see that (*) defines a bilinear form/ on V x V; simply observe that the above can be written as /(x,y) = xri4y= [x, ■■ xn]A yi yn. Moreover, with respect to the ordered basis (v,),, we have 0 /(v,,vy)=[0.-. l(i) • 0]A (/) = fl/y. and so the matrix of/ is A. Example 9.3 In Example 9.1 the matrix of/ relative to the standard ordered basis is /,,. Example 9.4 The matrix A = 1 2 5 -2 0 1 0-6 6 gives rise to the bilinear form xty = *i>i + 2x,y2 + 5*i>3 - 2*2>i + *2>3 " 6*3>2 + 6*3y3.

9. Bilinear and Quadratic Forms 129 Example 9.5 The matrix a h g A= h b f .8 f c. gives rise to the bilinear form x*/iy = axxyx +^2^2 + ^3^3 + ^^2+^2^1)+^1^3 + ^3^1)+/^3 + ^3^2)- Example 9.6 The bilinear form xl (y2 + 73) + x2y$ is represented, relative to the standard ordered basis, by the matrix A = r° 0 0 11" 01 0 0 EXERCISE 9.2 Determine the matrix of each of the bilinear forms (1) 2^,-3.^3 + 2.^2-5^3 + 4^,-, (2) 3.^,+2.^2 + ^3- It is natural to ask how a change of basis affects the matrix of a bilinear form. Theorem 9.1 Let V be a vector space of dimension n over afield F. Let (v,)n and (w^ be ordered bases of V. Iff: V x V —► Fis bilinear and if A is the matrix off relative to (v,)n then the matrix of f relative to (wt)n is PAP where P is the transition matrix from (w,)n to (v,)n. Proof n We have wy = J^/tyv,- f°TJ - 1, ••, n and so, by the bilinearity of/, n n f(^hwj) = /(2>,,v/, 5>wvJ 1=1 *=1 n n = EEAiAyM.v*) 1=1 k=l n n i=\k=i 1=1 = [P'AP]iJt from which the result follows. D

130 Further Linear Algebra Example 9.7 Consider the bilinear form 2xxyx - 3x2y2 + ^3 relative to the standard ordered basis of IR3. The matrix that represents this is /i = 2 0 0 0 0-3 0 0 1 To compute the form resulting in a change of reference to the ordered basis {(1,1,1),(-2,1,1),(2,1,0)), we observe that the transition matrix from this new basis to the old basis is /> = Consequently the matrix of the form that results from the change of basis is "0-6 4" P>AP = -6 6-8 1 -11 8 The form relative to the new basis can be read off from this matrix. ri 1 1 -2 2" 1 1 1 0 EXERCISE 9.3 Consider the bilinear

form / : IR3 standard ordered basis of IR3 by IR3 given with reference to the 2xlyl + x2y3-3x3y2 + x3y3. Determine its equivalent form with reference to the basis {(1,2,3),(-1,1,2),(1,2,1)). Definition If A and B are n x n matrices over a field F then we say that B is congruent to A if there is an invertible matrix P such that B - P'AP. It is readily seen that the relation of being congruent is an equivalence relation onMat.,VMF. EXERCISE 9.4 Prove that any matrix that is congruent to a symmetric matrix is also symmetric.

9. Bilinear and Quadratic Forms 131 Definition A bilinear form / : V x V -► V is said to be symmetric if f{x, y) = /(y, x) for all *, y € ^-, and skew-symmetric if f(x, y) = -/(y, jc) for all x, y 6 V. EXEflC/SE 9.S Prove that every bilinear form can be expressed uniquely as the sum of a symmetric and a skew-symmetric bilinear form. Do so for (1) xxyx + xxy2 + 2x2yx + x2y2\ (2) 2xlyl-3x2y2 + x3y3. It is clear that a matrix that represents a symmetric bilinear form is symmetric; and. conversely, that every symmetric matrix gives rise to a symmetric bilinear form. Definition If V is a vector space over a field F then by a quadratic form on V we mean a mapping Q : V -► F given, for some symmetric bilinear form/: V x y -► F, by (V*6V) COO = /(*,*)• Example 9.8 The mapping g : R2 -> R2 given by Q(xfy)=x2-xy + y2 is a quadratic form on R2, as can be seen on writing the right hand side as x x'i4x= [x y] i 1 Example 9.9 The mapping Q : R3 -► R3 given by C(*,y,z) = *2+y2-z2 is a quadratic form on R3. as can be seen on writing the right hand side as ldAx =[x y z] "1 0 0 1 0 0 0" 0 -1 X y z EXERCISES 9.6 Prove that the set of quadratic forms on a vector space V forms a sub- space of the vector space Map(Vf F).

132 Further Linear Algebra 9.7 Relative to the symmetric bilinear form provided by the standard inner product on IR", describe the associated quadratic form. 9.8 Prove that the following expressions are quadratic forms: (1) 3x2-5xy-7y2; (2) 3jc2 - Ixy + 5xz + 4y2 - 4yz - 3z2. In what follows we shall restrict the ground Meld F to be R Given a symmetric bilinear form/: V x V -* R by the associated quadratic form we shall mean the mapping Qf : V -► R defined by Qf{x) = /(*, x). Theorem 9.2 Let V be a vector space over R. Iff : V xV -> Risa symmetric bilinear form then the following identities hold: (\)Qf(\x)=\2Qf(x); (2) f(x,y) = \[Qf{x + y) -Qf(x) -Qf(y)]; (3)/(xly)=i[C/(x + y)-C/(x-y)]. Proof (1) <?/(>*) = /(>*,>*) = >2/(*,*) = >2C/M. (2) Since/ is symmetric, we have Qf{x + y) = f{x+ytx + y) = Qf{x) + 2/(x, y) + Qf{y). (3) By (1) we have Qf{-x) = Qf{x) and so, by (2), Qf(x-y) = Qf(x)-2f(xty) + Qf(y).

This, together with (2), gives (3). D An important consequence of the above is the following. Theorem 9.3 Every real quadratic form is associated with a uniquely determined symmetric bilinear form. Proof Suppose that Q : y -► R is a quadratic form. Suppose further that/, g : V xV -*R are symmetric bilinear forms such that Q-Qf- Qg. Then by Theorem 9.2 we see that/(x, y) = g(x, y) for all x, y 6 V and therefore/ = g. D Because of this fact, we define the matrix of a real quadratic form on a finite- dimensional vector space to be the matrix of the associated symmetric bilinear form.

9. Bilinear and Quadratic Forms 133 Example 9.10 The mapping Q : IR2 -► IR given by Q(xt y) = Ax2 + 6xy + 9y2 = [x y] is a quadratic form. The matrix of Q is 4 3* 3 9 and the associated bilinear form is /((*.>). (*\30) = ***' + W + x'y) + 9y/. EXEflC/SE 9.9 Given a quadratic form Q : IR" -► IR, let i4 be the matrix of Q and let A be an eigenvalue of A. Prove that there exist a, a,,, not all zero, such that 1=1 It is clear from Theorem 9.1 that symmetric matrices A,B represent the same quadratic form relative to different ordered bases if and only if they are congruent. Our objective now is to obtain a canonical form for real symmetric matrices under congruence, i.e. to obtain a particularly simple representative in each congruence class. This will then provide us with a canonical form for the associated real quadratic form. The results on orthogonal similarity that we have obtained in Chapter 8 will put us well on the road. For our immediate purposes, we note that if A and B are matrices that are congruent then A and B have the same rank. In fact, B = P'AP for some invertible matrix P. Since an invertible matrix is a product of elementary matrices, multiplication by which can be expressed in terms of elementary row and column operations, all of which leave the rank invariant, the rank of P'AP is the same as the rank of A. Theorem 9.4 If A is an n x n real symmetric matrix then A is congruent to a unique matrix of the form 0 4 3 3 9

134 Further Linear Algebra Proof Since A is real symmetric it follows by the Corollary to Theorem 8.10 and the Corollary to Theorem 8.9 that A is orthogonally similar to a diagonal matrix and all the eigenvalues of A are real. Let the positive eigenvalues be >,,..., >r and let the negative eigenvalues be ->r+,,..., -\r+s. Then there is an orthogonal matrix P such that /^ = ^().,,...,).,,-)^,,...,-)^,,0 0}, there being n-(r + s) entries 0. Let D = [djj]nxn be the diagonal matrix such that 1 otherwise. <£,.= Then it is readily seen that {PD)'APD=DP'APD = lr -h Now since P and D are both invertible, so is PD.

Thus we see that A is congruent to a matrix of the stated form. As for uniqueness, it suffices to suppose that L = "/, M = -/. are congruent and show that r-J and s = J. Now if L and M are congruent then they have the same rank, sor + s = r1 +sf. Suppose now, by way of obtaining a contradiction, that r < r' (in which case d < s). Let W be the real vector space Mat nx, IR. Clearly, IV is an inner product space under the definition (x|y) = x/y. Consider the mapping fL : W -* W given by /L(x) = Lx. Since L is symmetric we have (Lx|y) = (Lx)'y=x'Ly=(x|Ly) and so fL is self-adjoint. Likewise, so also is the mapping fM : W -► W given by /M(x) = Mx. Consider now the subspaces X= {xeW; *, = •--=^ = 0, Xr+s+l = ..= xn = 0}; K={x6W;^+, = ... = x^ = 0}.

9. Bilinear and Quadratic Forms 135 Clearly, X is of dimension s, and for every non-zero x 6 X we have (1) x'Lx=-x2r+l x2r+s<0. Also, Y is of dimension n - s\ and for every x 6 Y we have x'A/x = x\ + • • • + x\ ^ 0. Now since L and A/ are congruent there is an invertible matrix P such that M - P'LP. Defining fP : W -> W by fP(x) = Px and observing that/p is also self-adjoint, we then have, for all x £ K, 0 ^ x'A/x = (Mx | x) = (P'LPx | x) = ({fL ofP)(x) |/, (x))f from which we see that if Z = {fP{x); x £ Y} then (2) (Vz6Z) {/!,(«) 11)>0. Now since/P is an isomorphism we have dim Z = dim Y = n-J and so dimZ + dimX=n-s' + s>n^ dim(Z + X). It follows that the sum Z + X is not direct (for otherwise we would have equality), and so ZnX f {0W}. Consider now a non-zero element z of ZnX. From (1) we see that (fL{z) | z) is negative, whereas from (2) we see that (4(z) Iz) >s non-negative. This contradiction shows that we cannot have r1 < r. Similarly, we cannot have r < r\ We therefore conclude that r- r1 whence also s - $\ □ The above result gives immediately the following theorem which describes the canonical quadratic forms. Theorem 9.5 [Sylvester] Let Vbe a vector space of dimension n over R and let Q : V —» IR be a quadratic form on V. Then there is an ordered basis (v,)n of V such that if n x - Yl XjVi then e(x)=*; + ■•+x2r-x2r+l—x2+s. Moreover, the integers r and s are independent of such a basis. D Definition The integer r + s in Theorem 9.5 is called the rank of the quadratic form Q, and the integer r - s is called the signature of Q. Example 9.11 Consider the quadratic form Q : IR3 -► IR given by Q{x, y, z) = x2 - Ixy + Ayz - 2y2 + 4Z2.

136 Further Linear Algebra The matrix of Q is the symmetric matrix '1-10' A= -1-2 2 0 2 4 We can use the method outlined in the proof of Theorem 9.4 to effect a matrix reduction and thereby obtain the canonical form of Q. In many

cases, however, it is often easier to avoid such calculations by judicious use of the method of 'completing the squares'. For example, for the above Q it is readily seen that Q(x,y,z) = (x-y)2 + (y + 2z)2-4y2 which is in canonical form. Thus Q is of rank 3 and of signature 1. Example 9.12 The quadratic form Q : IR3 -» R given by Q[x, y, z) = 2xy + 2yz can be reduced to canonical form either by the method of completing squares or by a matrix reduction. The former method is not so easy in this case, but can be achieved as follows. Define \/2x = X + Yt \Jly=X-Y, \Jlz = Z. Then Q(xtyfz)=(X + Y)(X-Y) + (X-Y)Z = X2-Y2 + {X-Y)Z = (X + iZ)2-(K+iZ)2 = 5(* + y + z)2-5(*-:y + z)2, which is of rank 2 and signature 0. Example 9.13 Consider the quadratic form Q : IR3 -»IR given by Q{x>y,z) = 4xy + 2yz. Here we have Axy + 2yz = {x + yf - {x- y)2 + 2yz = X2 -Y1 + (X -K)Z [X = x + y, K= x -y] = (X + IZ)2-r2-KZ-iz2 = (X + iZ)2-(K+IZ)2

9. Bilinear and Quadratic Forms 137 where £ = x + y + ±z and rj = and so if we define = x - y + jz. Taking £ = z we then have *=iK + u-fl y=|K-u) z = C, P = ■ i i 2 2 2 J 4 o 0 0 1 then we have V y z_ = P re] V LcJ To see that P does the reduction, observe that the matrix of Q is "0 2 0" A= 2 0 1 0 1 0 A simple check gives P'AP = diag{l, -1,0}. EXERCISES 9.10 A quadratic form Q : IR3 -► IR is represented, relative to the standard ordered basis of IR3 by the matrix /i = 11-1 1 1 0 -1 0 -1 Determine the matrix of the canonical form of Q. 9.11 By completing squares, determine the rank and the signature of each of the quadratic forms (1) 2y2-z2 + xy + xz\ (2) 2xy-xz-yz\ (3) yz + xz + xy + xt + yt + zf. 9.12 For each of the following quadratic forms determine the canonical form, and an invertible matrix that does the reduction: (1) x2 + 2y2 + 9z2 - 2xy + Axz - 6yz\ (2) x2 + Ay2 + z2 - At2 + 2xy - 2xt + 6yz - 8yr - 14zf.

138 Further Linear Algebra Definition A quadratic form Q is said to be positive definite if Q[x) > 0 for all x f 0V. If Q is a quadratic form on a finite-dimensional vector space and if A is its matrix then Q is positive definite if and only if 0 < Q{x) = x'Ax = (i4x|x), which is the case if and only if the corresponding symmetric bilinear form is positive definite or, equivalently, the matrix A is positive definite. Theorem 9.6 A real quadratic form on a finite-dimensional vector space is positive definite if and only if its rank and signature are the same. Proof Clearly, such a form is positive definite if and only if there are no negative terms in its canonical form. This is the case if and only if the rank and the signature are the same. D Example 9.14 Let/: IR x IR

-»IR be a function whose partial derivatives fx ,fy are zero at the point (*o.>o)- Then the Taylor series at (x0 + h,y0 + h) is f(xo.yo) + tytfxx + 2M/« + *2/yv](*o,yo) + • ■ • • For small values of htk the significant term in this is the quadratic form [h2fxx+2hkfxy + k%](x0ty0) in htk. If it has rank 2 then its canonical form is ±H2 ± K2. If both signs are positive (i.e. the form is positive definite) then/ has a relative minimum at (v0, v0), and if both signs are negative then/ has a relative maximum at {x0,y0). If one sign is positive and the other is negative then/ has a saddle point at (jr0>>o)- Thus tne geometry is distinguished by the signature of the quadratic form. EXERCISES 9.13 Consider the quadratic form Q: IR3 -»IR that is represented, relative to the standard basis, by the matrix A = Is Q positive definite? 1 1 -1 1 1 0 -I 0 -1

9. Bilinear and Quadratic Forms 139 9.14 Show that neither of the quadratic forms defined on IR2 by 0) C(*,;y)=*2 + 3xy + y2; (2) Q(x,y) = 2x2-4xy + 3y2- z2, is positive definite. 9.15 Show that for the quadratic form Q defined on IR" by C(*i,.-.,*,,) = E(*r-*,)2 r<s both the rank and the signature are n -1. [Hint. Show that the matrix of Q relative to the standard ordered basis is «-1 -1 -1 .. -1 n-\ -1 .. -1 -1 n-\ .. . -1 . -1 . -1 /i = -1 -1 -1 ...n-1 Show that A is congruent to diag{n -1, n - 2,..., 2,1,0}.] 9.16 Consider the quadratic form J3(Xr5 + r + 5)jfrjf, r,*=l in which \ 6 IR is arbitrary. Show that with yt - £ rxr and y2 = £ xr r=\ i=l the form becomes ^5 + 2y,y2. Hence show that for all values of \ the rank is 2 and the signature is 0.

10 Real Normality We have seen in Theorem 8.8 that the ortho-diagonalisable linear mappings on a complex inner product space are precisely those that are normal; and in Theorem 8.10 that the ortho-diagonalisable linear mappings on a real inner product space are precisely those that are self-adjoint. It is therefore natural to ask what can be said about normal linear mappings on a real inner product space; equivalently, we may ask about real square matrices that commute with their transposes. Our objective here is to obtain a canonical form for such a matrix under orthogonal similarity. As we shall see, the main results that we shall obtain stem from further applications of the Primary Decomposition Theorem. The notion of minimum polynomial will therefore play an important part in this. Now, as we are assuming throughout that the ground field is IR, it is clear that we shall require a knowledge of what the monic irreducible polynomials over IR look like. This is the content of the following result. Theorem 10.1 A monic polynomial f over IR is irreducible if

and only if it is of the form X-afor some a 6 IR, or of the form (X - a)2 + b2 for some afb£R with b^O. Proof It is clear that X - a is irreducible over IR. Consider now the polynomial f=(X-a)2 + b2 = X2-2aX + a2 + b2 where at b 6 IR with 6^0. By way of obtaining a contradiction, suppose that/ is reducible over IR. Then there exist p, q 6 IR such that f=(X-p)(X-q) = X2-(p + q)X+pq. It follows that p + q - la and pq- a1 + b2. These equations give pt-2ap + a2 + b2 = 0 and so we have that p= \[2a ± yjAa2 -4(a2 + b2)] = a ± y/^b1.

 Since b ^ 0 by hypothesis, we have the contradiction p$ IR. Conversely, suppose that / 6 IR[X] is monic and irreducible. Suppose further that / is not of the form X -a where a 6 IR. Let z = a + ib £ C be a root of/, n noting that this exists by the fundamental theorem of algebra. If/ = £ a.X' then /=0 n 0 = /(z)= £ fl|Z'« Taking complex conjugates, we obtain i=0 i=0 /=0 Thus we see that z 6 C is also a root of/. Consequently, (X-z)(X-z)=X2-(z + z)X + zz = X2-2aX + a2 + fc2elR[X] is a divisor of/. Since by hypothesis/ is irreducible over IR, we must have z 6 C\IR whence 6^0, and then/ = X2 - 2aX + a2 + b2 = (X - a)2 + 62. □ Corollary The irreducibles of IR[X] are the polynomials of degree 1, ami the polynomials pX2 + aX + r wtere a2 - 4pr < 0. Proof If pX2 + qX + r were reducible over IR then it would have a linear factor X - or with a 6 IR, whence pa2 + qot + r = 0. For this to hold, it is necessary that q2 -Apr ^ 0. Consequently, if q2 -Apr < 0 then the polynomial pX2 + qX + r is irreducible. Conversely, if pX2 + qX + r is irreducible over IR then so is x' + ix^. P P It follows by Theorem 10.1 that there exist a, b 6 IR with fc ^ 0 such that ^ =-2a, - = a2+fc2. P P Consequently, f -4pr = Aa2? -4p2(a2 + b2) = -Vfc2 < 0 as required. D Example 10.1 The zeros of X^-l 6 IR[X] are **w,> = cos f +isin f for* = 0,1,. ...2p-l. As a product of irreducibles over IR, X2?-1 = (X - 1)(X + 1) n (X2 -2Xcos — + l). fc=i P

 If V is a real inner product space then a linear mapping /: V -* V is said to be skew-adjoint if/* = -/. Correspondingly, a real square matrix A is said to be skew-symmetric if A' = -A. Theorem 10.2 Let Vbe a non-zero finite-dimensional real inner product space and let f : V —» V be linear. Then there is a unique self-adjoint mapping g : V -► V and a unique skew-adjoint mapping h : V -» V such that f=g + h. Moreover, f is normal if and only if g and h commute. Proof Clearly, we have the decomposition /=^+/1 + ^-/1 in which 5(/ +/1 is self-adjoint, and \{f -/*) is skew-adjoint. Suppose now that/ = g + h where g is self-adjoint and h is skew-adjoint. Then we have/* = g* + h* = g - h. It follows from these equations that g = 5 (/"+/*) and h= \{f -/*), which establishes the uniqueness of such a decomposition. If now / is normal then/ 0/*=/* 0/ gives which reduces to go h = hog. Conversely, if g and h commute then it is readily seen that/0/*= g2 -h2 =/*o/ whence/ is normal. □ The reader will be familiar with the matrix equivalent of Theorem 10.2, that a real square matrix A can be expressed uniquely as the sum of a symmetric and a skew-symmetric matrix, namely A = \(A + A') + ^(A - A'). A characterisation of skew-adjoint linear mappings on a real inner product space is the following. Theorem 10.3 If Visa non-zero finite-dimensional real inner product space then a linear mapping f: V —♦ V is skew-adjoint if and only if (VxGV) (f(x)\x) = 0. Proof =» : If/ is skew-adjoint then for every x 6 V we have (f(x) | x) = (x I -/(*)) = -(* |/M) = -(fix) IX) and therefore (/(x) \ x) - 0. <= : If the condition holds then for all x, y 6 V we have 0={f(x + y)\x + y)=(f(x)\y) + (f(y)\x)

 which gives VMI y) = ~{f(y) I *) = -<* \f(y)) = (* I -/(>)>• It now follows by the uniqueness of adjoints that /* = -/ and consequently / is skew-adjoint. D EXEflC/SES 10.1 If V is a non-zero finite-dimensional real inner product space prove that for every linear mapping/ :V-»V there is a unique self-adjoint linear mapping g: V-»V such that (V*6V) (f(x)\x)=(g(x)\x). 10.2 If A is a real skew-symmetric matrix and if g is a polynomial such that g(/l) = 0 prove that g[-A) = 0. Deduce that the minimum polynomial of A contains no terms of odd degree. We now make the following observation. Theorem 10.4 If V is a non-zero finite-dimensional real inner product space and if f : V —» V is k normal then its minimum polynomial is of the form mj = JJ p, where plt...,pkare distinct irreducible polynomials. Proof * We know that the minimum polynomial of/ is of the general form mf = JJpJ' where p,,... tpk are distinct irreducibles. What we have to show here is that when / is normal every e, = 1. Suppose, by way of obtaining a contradiction, that we have e, ^ 2 for a particular index i. If we write as usual Vj = Ker p,(/)" then we have, for every x e V,, A-W^Welmp.-WnKcrftW. But since/ is normal so also isp,(/), by Theorem 8.6(2). It then follows by Theorem 8.6(3) that the restriction of p, (/)"" to V, is the zero mapping. If as usual we let f.: y. _> vt.be the mapping induced by/ on the/-invariantsubspaceV^, we thus have that Piifi)e'~l = 0. But this contradicts the fact that, by Theorem-^, the minimum polynomial of/ is pj'. □ Concerning the minimum polynomial of a skew-adjoint mapping, we have the following description.

144 Further Linear Algebra Theorem 10.5 Let V be a non-zero finite-dimensional real inner product space and f : V -* V a skew-adjoint linear mapping. If p is an irreducible factor of the minimum polynomial of f then either p= X or p= X2 + b2 for some bj^O. Proof Since skew-adjoint mappings are clearly normal it follows by Theorem 10.4 that mf k is of the form fjp, where pu... ,pk are distinct monic irreducible polynomials. We i-i also know, by Theorem 10.1, that each pt is either linear or of the form (X -a,)2+b2 with bt f 0. Suppose first thatp,- is linear, say pt = X - ait and let/ be the mapping induced by / on the primary component V; = Kerp,(/). Then we have/ = fl,id^. and consequently/* = /. Since/ is also skew-adjoint, it follows that 0 = / = fl/id^. Hence a{ = 0 and we have px =* X. Suppose now thatp, is not linear. Then we have 0 = ^ = /7^^ +(a2+ b2)idVl. Since/ is skew-adjoint, we deduce that 0 = /2 + 2^ + K2 + 62)idVj. These equalities give Aaj{ = 0. Now /^0 since otherwise we would have p,- = mf. = X, in contradiction to the hypothesis. Hence we must have at = 0 whence p/=X2+ 6? wherefc.^O. □ Corollary If f is skew-adjoint then its minimum polynomial mf is given as follows: (1) iff= 0 then mj= X; (2) if f b invertible then mf=l[(X2 + b2) i=i for distinct non-zero real numbers bx,..., bk\ (3) iffis neither 0 nor invertible then mf=xfl(X2 + b2) i=2 for distinct non-zero real numbers b2,. •., bk. Proof This is immediate from Theorems 10.4 and 10.S on recalling that/ is invertible if and only if the constant term in m, is non-zero. D

10. Real Normality 145 We now observe how orthogonality enters the picture. Theorem 10.6 If V is a non-zero finite-dimensional real inner product space and if f : V -*V is skew-adjoint then the primary components of fare pairwise orthogonal. Proof Let Vj and Vj be primary components of/ with i ^j. With the usual notation, if/Jy are the mappings induced on Vh Vjbyf suppose first that wy(= X + 6,, m* = X + 6, where 6,,6,- are non-zero. Observe first that, by Theorem 10.4, we have 6- f bj. Then, for x, 6 V-, and Xj 6 Vjt 0 =((/^ + 6¾)^) I *;> «(«*,) I xy)+ *?(*,!*,) = (fl*,) I-/(*;)> + *?(*, I *,) = (*,|/Vy)>+ *?<*, I *y> = (x,|-6^y) + 6?(x,|xy) = (6^-6/)(,,1^). Since 6? ^ 6?, it follows that (x, | xy) = 0. Suppose now that mj. = X2 + 6? with 6, ^ 0, and my = X. In this case the above array becomes 0=((/i-2+6?id,,,)(x,)|xy) = (/(*,)! -/(*,)>+ tfM*/> •*(»il*y> whence again we see that (x, | xj) = 0. D In order to establish the main theorem concerning skew-adjoint mappings, we require one further result. Theorem 10.7 Let V be a non-zero finite-dimensional inner product space and let W be a subspace of V. Then W is f-invariant if and only if W1 is ^-invariant.

146 Further Linear Algebra Proof Suppose that W is /-invariant. Since V = W © W1 we have (V* G W)(Vy G W*) (jr |/*(y)> = {/{*) | y> = 0. It follows that/*(y) G W1 for all y G W1 and so W1 is/*-invariant. Applying this observation again, we obtain the converse; for if WL is/Mnvariant then W = W11 is/**=/-invariant. D Theorem 10.8 Let V bea non-zero finite-dimensional real inner product space and let f: V —► V be a skew-adjoint linear mapping with minimum polynomial mj = X2 + b2 where b^O. Then dim V is even, and there is an ordered orthonormal basis of V with respect to which the matrix off is of the form "0 -b M{b) = b 0 -6 0 0 -b b 0 Proof We begin by showing that dim V is even and that V is an orthogonal direct sum of /-cyclic subspaces each of dimension 2. For this purpose, let y be a non-zero element of V. Observe that/(y) f \y for any scalar \; for otherwise, since/2(y) = -b2y% we would have \2 = -b2 and hence the contradiction b= 0. Let W, be the smallest/-invariant subspace containing y. Since/2(y) = -~b2y it follows that W, is /-cyclic of dimension 2, a cyclic basis for W, being {y,/(y)}. Consider now the decomposition V = W, © W^. This direct sum is orthogonal; for if p is the projection on W, parallel to Wf- then Imp = W, and Ker p=Wf- and so p is an ortho-projection. By Theorem 10.7, Wf- is/*-invariant and so, since/* = -/, we see that Wf- is also/-invariant, of dimension dim V - 2. Now let V, = IVf and repeat the argument to obtain an orthogonal direct sum Vi = W2 © W^ of/- invariant subspaces in which W2 is an/-cyclic subspace of dimension 2. Continuing in this manner, we note that it is not possible in the final deomposition to have dim W;J- = 1. For, if this were so, then W% would have a singleton basis {z} whence/(z) ¢ W;|-, a contradiction. Thus W;j- also has a basis of the form {z,f{z)} and therefore is likewise/-cyclic of dimension 2. It follows therefore that dim V is even.

10. Real Normality 147 We now construct an orthonormal basis for each of the/-cyclic subspaces Wt. Consider the basis {>,,/(>,)}. Since lrt*)l2 = (fly.-) I -/"(*)) = -{fHyt) Iy,) = *2 b.ll2 it follows by applying the Gram-Schmidt process that an orthonormal basis for W: ' lbil,*bJJ' Sincc/2(y,-) = -b2yi it is readily seen that the matrix of/- relative to B{ is XA Pasting together such bases, we obtain an orthonormal basis of V with respect to which the matrix of/ is of the stated form. D Corollary 1 If V is a non-zero finite-dimensional real inner product space and f : V —» V is a skew-adjoint linear mapping then there is an ordered orthonormal basis of V with respect to which the matrix off is of the form in which each Af,- is either 0 or as described in Theorem 10.8.

Proof Combine the corollary of Theorem 10.5 with Theorems 10.6 and 10.8. □ Corollary 2 A real square matrix is skew-symmetric if and only if it is orthogonally similar to a matrix of the form given in Corollary 1. □ Example 10.2 The real skew-symmetric matrix A = has minimum polynomial X(X2 + 3) and so is orthogonally similar to the matrix * 0 -1 -1 1 f 0 1 -1 0 0 -n/3 n/3 0

146 Further Linear Algebra EXERCISES 10.3 Show that the skew-symmetric matrix "02-2 A= -2 0-1 2 10 is orthogonally similar to the matrix "0 3 0 -3 0] 0 0 0 0 10.4 If A is a real skew-symmetric matrix prove that A and A' are orthogonally similar. We now turn to the general problem of a normal linear mapping on a real inner product space. For this purpose, recall from Theorem 10.2 that such a mapping can be expressed uniquely in the form g + h where g is self-adjoint and h is skew-adjoint. Moreover, by Theorem 8.10, g is ortho-diagonalisable. Theorem 10.9 Let V be a non-zero finite-dimensional real inner product space and let f \V —*V be a normal linear mapping whose minimum polynomial is mf = (X -a)2+b2 where 6^0. If gt hare respectively the self-adjoint and skew-adjoint parts off then (1) h is invertible; (2)m, = X-fl; (3) mh = X2 + b2. Proof (1) Suppose, by way of obtaining a contradiction, that Ker h f {0V}. Since/ is normal we have goh- hogby Theorem 10.2. Consequently, Ker h is g-invariant. Since/ = g + h% the restriction of/ to Ker h coincides with that of g. As Ker h is g-invariant, we can therefore define a linear mapping/* : Ker h -» Ker h by the prescription Since g is self-adjoint, we see immediately that so is/*. By Theorem 8.10,/* is then ortho-diagonalisable, and so its minimum polynomial is a product of distinct linear factors. But mf* must divide mf which, by hypothesis, is of degree 2 and irreducible. This contradiction therefore gives Ker h = {0V} whence h is invertible.

10. Real Normality 149 (2) Since/ = g + h with g* = g and h* - -h we have/* = g-h. Since also, by hypothesis,/2 -2a/+ {a2 + fc2)idv = 0 we have iff -2a/*+ (a2 + b2)\6v = 0 and consequently f2-(ff = 2a(f -/-) = 4^. Thus, since/o/* = /*o/, we see that ^0^=1(/+/^)01(/-/^)=1^-^1 = ^ and so (g -aidy) o/i = 0. Since /i is invertible by (1), we then have that g -a\6v = 0 whence mg-X -a. (3) Since/ -h- g- a\6v we have/ = h + aidy and so 0=/2-2a/ + (a2 + fc2)idv = (/i + aidy)2 -2a(/t + aidv) + (a2 + 62)id^ = /t2 + fc2idv. Since A is skew-adjoint and invertible it now follows by the Corollary to Theorem 10.5 that/^ = ^ + 62. D We can now extend Theorem 10.6 to normal mappings. Theorem 10.10 If V

is a non-zero finite-dimensional real inner product space and if f : V —» Vis normal then the primary components off are pairwise orthogonal Proof By Theorems 10.1 and 10.4 the minimum polynomial of/ has the general form * mf = {X -a,)I](X2 -2a,X + a) + b}) where each bt ^ 0. The primary components of/ are therefore Vl = Ker(/" - fl|idy) and (i = 2 k) V{ = Ker {f -2aJ+ (a2 + b2)\dv). Moreover, for each /, the induced mapping/- on V, is normal with minimum polynomial X - a | if i = 1 and X2 -lafi + aj + b2 otherwise. Now /• = ft + ^ where ft is self-adjoint and ht is skew-adjoint. Moreover, ft and A, coincide with the mappings induced on V; by g and /i. To see this, let these mappings beg'th' respectively. Then for every x 6 Vt we have gi(x) + /,,(*) = /M = /(*) = *M + h(x) = *'(*) + h'(x) and so ft - g' = h' - /t,. Since the left hand side is self-adjoint and the right hand side is skew-adjoint, each must be zero and we see that ft = g' and /i, = h'.

150 Further Linear Algebra Suppose now that ij > 1 with i ^ j. Then the minimum polynomials of fhfj are mfi = X2 -2a{X + a2 + b2t mfi = X2 -2a}X + a] + b) where either ^ f a;- or 6? ^ b2. Then, by Theorem 10.9, we have mgi=X-ah mgJ = X-ajt m,,=X2 + fc?, mhj = X2+b]. Now given *j 6 V-t and x,- 6 Vy we therefore have, precisely as in the proof of Theorem 10.6, 0 = ((/,? + b}\6Vi)(Xi) | Xj) = (b2 -b2)(xt | xj)t so that in the case where b2 ^ b2 we have (jc, | Xj) = 0. Likewise, 0 = ((& -aiJdVl){Xi) | xy) = (g(x,) | xj) -a^ \ x}) = (*il*(*y))-fl|(*fl*y) = (*il*y(*y))-«.-(*il*y) = fly(Xi|X;)-fl,(*,|*;) = («y- <0(*il*y). so that in the case where at f a; we also have (xt | Xj) = 0. We thus see that V2t..., Vk are pairwise orthogonal. That V, is orthogonal to each Vj for / > 2 follows from the above equalities on takingy = 1 and using the fact that/, = at\6Vl is self-adjoint and therefore gt - /j and /i, = 0. D We can now establish the main result. Theorem 10.11 If V is a non-zero finite-dimensional real inner product space and iff: V —» V is a normal linear mapping then there is an ordered orthonormal basis of V with respect to which the matrix off is of the form "A, where each At is either a 1 x 1 matrix or a 2x2 matrix of the form a -0 0 a in which /3^0.

151 Proof With the same notation as above, let mf = (X -o.)f[{X2 -2o,X + a) + b}) i=2 and let the primary components of/ be V{f for i = 1,..., it. Then m^ = X - a, if i = 1 and mfi = X2 - 2a,X + a) + 6? otherwise. For each 1^ with / ^ 1 we have/,- = & + /i, where the self-adjoint part g, has minimum polynomial X - at and the skew-adjoint part has minimum polynomial X2 + b). Now by Theorem

10.8 there is an ordered orthonormal basis Bt of V-, with respect to which the matrix of ht is 0 b, -hi 0 0 bi -hi 0 A#(W = 0 -6,- 6, 0_ Since the minimum polynomial of gt is X -at we have gt(x) = tij* for every x 6 B,- and so the matrix of g, with respect to Bt is a diagonal matrix all of whose entries are a,. It now follows that the matrix of/ = g, + /i, with respect to B, is a,- -6,- 6: fl. A#(fl,,W = -*< a. a. *, -*>, a, In the case where i = 1, we have/, = a{\&Vl so/, is self-adjoint. By Theorem 8.10, there is an ordered orthonormal basis with respect to which the matrix of/, is diagonal. Now since, by Theorem 10.10, the primary components are pairwise orthogonal we can paste together the ordered orthonormal bases in question, thereby obtaining an ordered orthonormal basis of V with respect to which the matrix of/ is of the form stated. D Corollary A real square matrix is normal if and only if it is orthogonally similar to a matrix of the form described in Theorem 10.11. D

 Our labours produce a bonus. Recall that if V is a real inner product space then a linear mapping/: V -» V is orthogonal if and only if/-1 exists and is/*. Likewise, a real square matrix A is orthogonal if A~l exists and is A'. Clearly, an orthogonal mapping commutes with its adjoint and is therefore normal. Correspondingly, an orthogonal matrix commutes with its transpose. We can therfore deduce as a special case of Theorem 10.11 a canonical form for orthogonal mappings and matrices. Theorem 10.12 If V is a non-zero finite-dimensional real inner product space and iff : V —» V is an orthogonal linear mapping then there is an ordered orthonormal basis of V with respect to which the matrix offis of the form in which each Ak is a 2x2 matrix of the form where 0 + 0 and a2 + 02 - 1. Proof With the same notation as in the above, we have that the matrix M(aitbj) which represents/, relative to the ordered basis Bt is an orthogonal matrix (since/- is orthogonal). Multiplying this matrix by its transpose we obtain an identity matrix and, equating entries, we see that a2 + b2 - 1. As for the primary component V,, the matrix of/| is diagonal. Since the square of this diagonal matrix is an identity matrix, its entries must be ±1. We can now rearrange the basis in such a way that the matrix of/ has the stated form. D Example 10.3 If/: IR3 -»IR3 is orthogonal then/ is called a rotation if det A = 1 for any matrix A that represents /. If / is a rotation then there is an ordered orthonormal basis of IR3 with respect to which the matrix of/ is 1 0 0" 0 cost? -sintf J) sintf cost? for some real number tf.

11 Computer Assistance Many applications of linear algebra require careful, and sometimes rather tedious, calculations by hand. As the reader will be aware, these can often be subject to error. The use of a computer is therefore called for. As far as computation in algebra is concerned, there are several packages that have been developed specifically for this purpose. In this chapter we give a brief introduction, by way of a tutorial, to the package 'LinearAlgebra' in MAPLE 7. Having mastered the techniques, the reader may freely check some of the answers to previous questions!

154 Further Linear Algebra Having opened MAPLE, begin with the input: > with(LinearAlgebra): (1) Matrices There are several different ways to input a matrix. Here is the first, which merely gives the matrix as a list of its rows (the matrix palette may also be used to do this). At each stage the MAPLE output is generated immediately following the semi-colon on pressing the ENTER key. input: > ml:=Matrix([[1,2,3],[2,3,-1], [6,-3,-4]]); output: ml := 12 3 2 3-1 6 -3 -4 In order to illustrate how to do matrix algebra with MAPLE, let us input another matrix of the same size: > m2:=Matrix([[-1,4,7],[-2,5,41], [-6,-3,3]]); m2:= -14 7 -2 5 41 -6 -3 3 Then here is one way of adding matrices, using the 'Add' command: > m3 : =Add (ml, m2) ; m3 := 0 6 10 0 8 40 0 -6 -1 As for multiplying matrices, this can be achieved by using the 'Multiply' command. To multiply the above matrices, input: > m4:=Multiply(ml,m2) ; m4 := -23 5 98 -2 26 134 24 21 -93 Now 'Add' also allows linear combinations to be computed. Here, for example, is how to obtain 3ml + 4m2:

11. Computer Assistance 155 > Add(ml,m2,3,4); "-1 22 37" -2 29 161 -6-21 0 (2) A simpler method An more convenient way to input commands is to use algebraic operations. By way of example: > ml+m2; "0 6 10" 0 8 40 0 -6 -1 Multiplication by scalars is obtained by using a V: > 3*ml+4*m2; "-1 22 37" -2 29 161 -6-21 0 Multiplication of matrices is obtained by using a V: > ml.m2; "-23 5 98" -2 26 134 24 21 -93 As for the more complicated expression: ml(4m2 - 5ml2): > ml.(4*m2-5*mlA2); ' 133-360-178" -388 -51 726 -1044 654 803 (3) Inverses Inverses of matrices can be achieved by using either 'Matrixinverse' or as follows (here it is necessary to insert brackets round the -1): > mlA(-l);

156 Further Linear Algebra !i J_ !L 83 83 83 zl *L zl 83 83 83 24 -IS I 83 83 83 (4) Determinants To compute a determinant, use the command 'Determinant':

> Determinant(ml); -83 Of course the determinant of a product is the product of the determinants: each of the commands > Determinant(ml.m2); > Determinant(ml)*Determinant(m2); gives 70218 Note that we can use negative powers in products: > ml*(-3).m2*3; 1841635 4640885 9128894 T 571787 -19800794 571787 8866674 571787 -32833132 571787 19933848 571787 -62530138 571787 45220050 L 571787 571787 571787 (5) More on defining matrices We now look at other ways of defining a matrix. We start with a clean sheet (to remove all previous definitions): > restart; with(LinearAlgebra): We can enter a matrix as a row of columns: M0:= «a,b,ol<d,e,f>l<g,h,i»; M0:- a d g b eh cfi or as a column of rows (this can also be done using the matrix palette):

 Ml:= «a I b I c> , <d I e I f> , <g I h I i»; Ml := a b c def g h i Then, for example, we have > Ml*2; a2 + bd + eg ab + be + eh ac + bf + ci da + ed+fg bd + e*+fh dc + ef+fi ga + hd + ig gb + hc + ih cg+fh + i2 > Determinant(Ml); aei -afh + dch -dbi + gbf -gee Particular types of matrix can be dealt with as follows. A 3 x 3 lower triangular matrix, for example: > M2:=Matrix(3,[[1],[2,3],[4,5,6]], shape=triangular[lower]); M2:= For a symmetric 3x3 matrix, begin with > M3:=Matrix{3,3,shape=symmetric); '1 0 0" 2 3 0 4 5 6 M3:= 0 0 0 0 0 0 0 0 0 then input, for example, > M3[l,l]:=2;M3[1,3]:=23;M3[2,3]:=Pi; M\x M3, MS 2,3 = 2 = 23 = 7T

 > M3; M3:= > Determinant(M3); 2 0 23 00 ir 23 7T 0 -2tt2 Skew-symmetric matrices can be done similarly: > M4:=Matrix(3,3,shape=antisymmetric); M4:= 0 0 0 0 0 0 0 0 0 > M4[1,2]:=2; M4[1,3]:=23;M4[2,3]:=Pi; M\2 A/4,3 ^3 = 2 = 23 = 7T > M4; M4:= > Determinant(M4); 0 2 23 -2 0 7T -23 -7T 0 0 Hermitian matrices can be dealt with as follows (note here the •): > M5:=Matrix(3,3,shape=hermitian); M5:= 0 0 0 0 0 0 0 0 0 >M5[1,1]:=2; M5[1.2]:=5+7*I; M5[1,3]:=23-6*I; M5[2f3]:=I;

 M5 M5 M5, i.i ».2 '».3 M52f3 = 2 = 5 + 7/ = 23-6/ = / > M5; M5:= > Determinant(M5); 5 + 7/ 0 23-6/ 5-7/0 / 23 + 6/ -/ 0 -384 Submatrices can be defined, for example, as follows: > ml:=M5[2..3,2..3]; ml := 0 / -/0 ml := We can also input matrices where the (i j)-th entry is a function of/ andy: > m2:=Matrix(6,S,(i,j)->i*j); 12 3 4 5 6 2 4 6 8 10 12 3 6 9 12 15 18 4 8 12 16 20 24 5 10 15 20 25 30 ^6 12 18 24 30 36 (6) Writing procedures To use a more complicated function we write a procedure. The following example illustrates a very simple procedure in order to define the 6 x 6 identity matrix: > f:=proc(i,j); if i=j then 1 else 0 fi; end; / := proc(/, j) if i = j then 1 else 0 fi end

160 Further Linear Algebra > m3:=Matrix(6,6, f) ; m3 := 10 0 0 0 0 0 10 0 0 0 0 0 10 0 0 0 0 0 10 0 0 0 0 0 10 0 0 0 0 0 1 Here is a more complicated procedure: > f:=proc(i, j); if i> j then x else if i=j then 0 else -y fi fi; end; / := proc(i, j) if j < i then x else if i = j then 0 else - y fi fi end The 6x6 matrix whose (ij')-th entry is given by this function is: > m4:=Matrix(6,6,f); 0 -y -y -y -y -y x 0 -y -y -y -y xx 0 -y -y -y xx x 0 —y -y x x x x 0 —y ^x x x x x 0 > d:=Determinant(m4); d := (jc4 - x3y + y2x2 - y3x + y4)xy m4 := We can simplify this: > expand(d*(x+y)); x6y + xy6 x6y + xy6 So the determinant of m4 is . Examine this for a few values of n: x+y > for n from 2 to 8 do m4:=Matrix(n,n,f); d:=Determinant(m4); print(' size', n, ' gives',expand(d*(x+y))); od:

11. Computer Assistance 161 size, 2, gives, yx2 + y1 x size, 3, gives, -x3y + y3x size, 4, gives, jc4> + yAx size, 5, gives, -jc5> + y5x size, 6, gives tx6y + xy6 size, 7, giv«, -x'y + xy7 «Z€, 8, gives, jc8> + xy* The general theorem should be easy to spot. Try to prove it. Here is a way to input a tri-diagonal matrix using the command 'BandMatrix': > m5:=BandMatrix([sin(x),x"3,-cos(x)],1,6,6); 3 -cos(jc) 0 0 0 0 m5 := jr sin(x) 0 0 0 0 jr -cos(jc) 0 sin(x) 0 0 0 sin(x) 0 0 -cos(jc) 0 0 -cos(jc) sin(jc) x3 -cos(jc) 0 sin(x) jc3 > d:=Determinant(m5); d := jc18 + 5jcl2sin(jc)cos(x) + 6jc6cos(x)2sin(jc)2 + cos(x)3sin(jt)3 We can also differentiate the entries of a matrix using the 'map' command: > m6:=map(diff ,m5,x) ; 3jc2 sin(x) 0 0 0 0 cos(jc) 3jc2 sin(x) 0 0 0 0 cos(jc) 3jc2 sin(x) 0 0 0 0 cos(jc) 3x2 sin(x) 0 0 0 0 cos(jc) 3jc2 sin(jc) 0 0 0 0 cos(jc) 3jc2 md := > dl:=Determinant(m6); d\ :=729jc,2-405jc8sin(jc)cos(jc) + 54x4cos(jc)2sin(jc)2 -cos(x)3sin(jt)3 (7) More on determinants Again start with a clean sheet:

162 Further Linear Algebra > restart; with(LinearAlgebra): We now examine the determinants of some other matrices and try to spot the general results. Here is the first of two examples: > f:=proc(i, j) ; if i = j then x else b fi; end; / := proc(i J) if i = j then x else b fi end Consider a 6 x 6 matrix whose entries are given by this function: > m:=Matrix(6,6,f) ; 'xbbbbb' b xbbbb bb xbbb m: bbb xbb bbbbx b bbbbb x_ > d:=Determinant(m); d := jc6 -15x*b2 + 40*V -

45b4x2 + 24bsx -5b6 Factorise the answer: > factor(d); {x + 5b){x-b)s This looks like a simple result. Let us examine a few cases: > for n from 2 to 8 do m:=Matrix(n,n,f); d:=Determinant(m); print(factor(d)); od: {x-b){x + b) {x + 2b){x-b)2 {x + lb){x-b)3 (jc + 46)(jc - 6)4 {x + 5b){x-b)5 {x + 6b)(x-b)6 {x + nftx-b)1

11. Computer Assistance 163 It is now easy to spot the general theorem. Try to prove it. Here is a second example: > f:=proc(i,j); if i = j then x else if i < j then y else -y fi fi; end; / := proc(i J) if i = j then x else if i < y then y else - y fi fi end > m:=Matrix{6,6,f); " x y y y y y" -y x y y y y -y -y x y y y m .= -y -y -y x y y -y -y -y -y x y _-y -y -y -y -y *_ > d:=Determinant(m); </:=/ + 15jc4y2 + 15yV + y6 It is less easy to spot the general result here. Let us try a few cases: > for n from 2 to 8 do m:=Matrix(n,nf f); d:=Determinant(m); print(sort(d, [x,y])); od: x2+y2 jc3 + 3xy2 x4 + 6x2y2 + y4 x5 + 10jc3y2 + 5xy4 jc6 + 15jc4y2 + 15jc2y4 + y6 x1 + 2\x5y2 + 35x3y4 + 7xy6 x* + 2%x6y2 + 70x4y4 + 28x2y6 + y8 Here it is harder to spot a simple form of the solution. However there is a compact form for the solution: > for i from 2 to 8 do expand(((x+yPi+(x-y) Ai)/2) ; od;

164 Further Linear Algebra x2 +y2 x3 + 3xy2 x4 + 6x2y2 + y4 jc5 + 10jc3>2 + 5xyA x6 + \5x*y2 + \5x2y4 + y6 x1 + 21jcV + 35jc3/ + 7xy6 jc8 + 28x6y2 + 70jc4/ + 28x2y6 + y8 Now try to prove the general result. (8) Matrices with subscripted entries Dealing with these is easy. Consider the Vandermonde matrix > m:=Matrix(4f4, (i,j)-> b[j]A(i-1)); 1111 m := 6j 62 63 64 6, 62 63 64 61 bo b-x bi /2 u3 u4 J > d:=Determinant(m); d 1= 6263 64 —6264 63 —62 6364 +62 6463 +62 6364 —62 6463 —6,63 64 +6,642633+6,62^4^6,6226^-6,62^42+6,62^32+^^3^^^1^33 -6,262643+6,262633+6,262364-6,26^63-6/63642+6^64632+6,362642 -6,362632-6/62264+6/62263 No surprise what we will obtain when we factorise this: > factor(d); (-64 + 63)(62 -63)(62 -64)(-63 + 6,)(6, -64)(6, -62) (9) Eigenvalues and eigenvectors Let us now find the eigenvalues of a matrix. First input a matrix: > a:=Matrix([[-2,-3,-3],[-1,0,-1],[5,5,6]]); a := -2 -3 -3 -1 0 -1 5 5 6 We use the command 'Eigenvalues':

11. Computer Assistance 165 > e:=Eigenvalues(a); e :- > e[ll; > e[2]; 1 > e:=Eigenvalues(a, output='Vector[row]'); * := [2, 1,1] Of course we can find the eigenvalues as the roots of the characteristic polynomial: >

ch:=CharacteristicPolynomial(afX); ch:=5X-4X2-2 + X3 > factor(ch); (X-2)(X-l)2 We can also find the eigenvalues as the solutions of the characteristic equation: > solve(ch=0,X); 2,1,1 Next, we can find the eigenvectors with the 'Eigenvectors' command: > evl:=Eigenvectors(a); ev\ := Oil 1 0 I .-1 -1 -Sj > ev2:=Eigenvectors(a, output='list');

166 Further Linear Algebra - r f 1.2- . L v "1" 0 -1 . i * . 2.1.- V "3iir i ■ .5 j j j _ ev2 := This needs some interpretation. It returns a list of lists. Each list has the form [ei, mi, {v[1, i], ...v[m, i]}] where ei is the eigenvalue, mi is its algebraic multiplicity and the set {v[1, i],..., v[m, i]} gives ni linearly independent eigenvectors where ni is the geometric multiplicity. Hence the eigenvalue 1 has algebraic multiplicity 2 with "1" 0 -1 . "0" 1 -1 as two linearly independent eigenvectors. The corresponding eigenspaces can also be found with a 'NullSpace' command: > id:=IdentityMatrix(3): > kl:=NullSpace(a-l*id); > k2:=NullSpace(a-2*id); *2:= ' k. -1" 0 1 . "-f 1 0 ^ J 3' 1 > p:=Matrix([kl[l],kl[2],k2[l]]); P'= -1 -1 3 0 1 1 1 0 -5 Now compute p~xap and we should obtain a diagonal matrix: > p^(-l).a.p; 1 0 0 0 1 0 0 0 2 (10) Computing Jordan normal forms Exercise 5.4 asks for the Jordan normal form of the following matrix:

11. Computer Assistance 167 > A:=Matrix([[5,-1,-3,2,-5],[0,2,0,0,0], [1,0,1,1,-2], [0,-1,0,3,1], [1,-1,-1,1,1]]); A := 5 -1 -3 2 -5 0 2 0 0 0 10 11-2 0-1031 1-1-11 1 First compute the characteristic polynomial: > factor(CharacteristicPolynomial(A,X)); (-3 + X)2(X-2)3 and then the minimum polynomial: > factor(MinimalPolynomial(A,X)); (X-2)2(-3 + X)2 Now we could write down the Jordan normal form, but MAPLE does it easily: > JordanForm(A); 2 10 0 0 0 2 0 0 0 0 0 3 10 0 0 0 3 0 0 0 0 0 2 To determine the transition matrix P to a Jordan basis: > P:=JordanForm(A,output='Q');; 11-1 3 3 0 10 0 1 110 0 1 0 1 1-10 0 0 0 1 1 A knowledge of P allows the determination of a corresponding Jordan basis (11) The Gram-Schmidt process Let us see how to apply the Gram-Schmidt orthonormalisation process. Consider again Example 1.12. > xl:=<0llll>; x2:=<H0ll>; x3 :=<1111 0>;

168 Further Linear Algebra xl := [0,1,1] x2 := [1,0,1] x3 := [1,1,0] > GramSchmidt([xl(x2,x3]); no, l.i]. [i-i.il [§,mh Now we normalize to find an orthonormal basis: > GramSchmidt([xl,x2fx3]normalized); [[0,iV5,iV5L [Jv«,4vS,i>/9, [^,^,-^ We can compute inner products using the 'DotProduct'

command: > DotProduct(<l/2137/111-15/17 17/9>. <1/1113/512/7I9/16>); 235519 104720 and so we can apply the Gram-Schmidt process as in hand calculations: > yl:=Normalize(xl,Euclidean); yl:=[0,|v5,iv^] > y2:=Normalize(x2-DotProduct(x2,yl)*yl,Euclidean); 72:=1^/5,4^, In/5] > y3:=Normalize(x3-DotProduct(x3,y2)*y2 -DotProduct(x3,yl)*ylfEuclidean); EXERCISES 11.1 Find the eigenvalues and corresponding eigenvectors of "12 3 4" 5 6 7 8 9 10 11 12 " 13 14 15 16

11. Computer Assistance 169 11.2 Find the determinant and the inverse of'12 3 4 /i = 11.3 For which value of x £ R is the matrix 2 0 3 3 3 1 1 1 3 invertible? 11.4 Given the matrices A = 1 -1 3 2 3" 1 2 -2 1 , * = -2 1 0 3 -2 -3 5 1 -2 give MAPLE code to compute the characteristic polynomial of 11.5 Write a MAPLE program to find the maximum value m of the determinant of all 3 x 3 matrices all of whose entries are 1 or 2. Modify your code to find all such 3x3 matrices with determinant m. 11.6 Let 4, = fa/lix- where ■{ +j if* <j; /2 otherwise, and let Bn = [bij]nxn where by = i2 +j2. Write MAPLE code to compute the determinants of An and Bn for n from 1 to 10. If C,, = An - Bnt write code to compute the determinant of C,, for n from 1 to 10. Devise a program that produces lists la and ic in which the i-th entry of ta is the determinant of A(and the i-th entry of ic is the determinant ofC,. Guess the determinant of Cn for arbitrary n and write a program to verify your conjecture for the first 20 values of n.

170 Further Linear Algebra 11.7 Let us compute the determinant of matrices whose (t ,y)-th entry is given by a polynomial in i andy, for example: > fl:= (i,j)->i*2+j*2: Determinant(Matrix (6, 6, f 1)); MAPLE gives the answer 0. Try some more matrices whose (i.y)-th entry is a polynomial in i andy, for example: > f2:=(i,j)->i"6*j"2+j"3+12: f3:=(i,j)->l+i"2*j-i"3*j"5: f4:=(i,j)-> (i+j"3)"3+(i"5-j"2-6*i*j"2)A2: Determinant(Matrix (6,6,f2)); Determinant (Mat rix(5,5,f 3)); Determinant(Matrix(11,11,f4)); What do you observe? There must be a theorem here. Can you prove it?

12 ... but who were they? This chapter is devoted to brief biographies of those mathematicians who were foremost in the development of the subject that we now know as Linear Algebra. We do not pretend that this list is exhaustive, but we include all who have been mentioned in the present book and in Basic Linear Algebra. Those included are: Bessel, Friedrich Wilhelm; Bezout,

Etienne; Cauchy, Augustin Louis; Cayley, Arthur; Fibonacci, Leonardo Pisano; Fourier, Jean Baptiste Joseph; Gram, Jorgen Pederson; Hamilton, Sir William Rowan; Hermite, Charles; Hilbert, David; Jordan, Marie Ennemond Camille; Kronecker, Leopold; Lagrange, Joseph-Louis; Laplace, Pierre-Simon; Lie, Marius Sophus; Parseval des Chenes, Marc-Antoine; Schmidt, Erhard; Schwarz, Hermann Amandus; Sylvester, James Joseph; Toeplitz, Otto; Vandermonde, Alexandre Theophile. We refer the reader to the website http://www-history.mcs.st-andrews.ac.uk/history/ for more detailed biographies of these and more than 1500 other mathematicians

172 Further Linear Algebra Bessel, Friedrich Wilhelm Bom: 22 July 1784 in Minden, Westphalia (now Germany). Died: 17 March 1846 in Konigsberg, Prussia (now Kaliningrad, Russia). Wilhelm Bessel attended the Gymnasium in Minden for four years but he did not appear to be very talented. In January 1799, at the age of 14, he left school to become an apprentice to the commercial firm of Kulenkamp in Bremen. Bessel spent his evenings studying geography, Spanish, English and navigation. This led him to study astronomy and mathematics, and he began to make observations to determine longitude. In 1804 Bessel wrote a paper on Halley's comet, calculating the orbit using data from observations made by Harriot in 1607. He sent his results to Heinrich Olbers who recognised at once the quality of Bessel's work. Olbers suggested further observations which resulted in a paper at the level required for a doctoral dissertation. In 1806 Bessel accepted the post of assistant at the Lilienthal Observatory, a private observatory near Bremen. His brilliant work there was quickly recognised and both Leipzig and Greifswald universities offered him posts. However he declined both. In 1809, at the age of 26, Bessel was appointed director of Frederick William III of Prussia's new Konigsberg Observatory and appointed professor of astronomy. He took up his new post on 10 May 1810. Here he undertook his monumental task of determining the positions and proper motions of over 50,000 stars. Bessel used Bradley's data to give a reference system for the positions of stars and planets and also to determine the positions of stars. He determined the constants of precession, nutation and aberration winning him further honours, such as a prize from the Berlin Academy in 1815. Bessel used parallax to determine the distance to 61 Cygni, announcing his result in 1838. His method of selecting this star was based on his own data for he chose the star which had the greatest proper motion of all the stars he had studied, correctly deducing that this would mean

that the star was nearby. Since 61 Cygni is a relatively dim star it was a bold choice based on his correct understanding of the cause of the proper motions. Bessel announced his value of 0.314" which, given the diameter of the Earth's orbit, gave a distance of about 10 light years. The correct value of the parallax of 61 Cygni is 0.292". Bessel also worked out a method of mathematical analysis involving what are now known as Bessel functions. He introduced these in 1817 in his study of a problem of Kepler of determining the motion of three bodies moving under mutual gravitation. These functions have become an indispensable tool in applied mathematics, physics and engineering. Bessel functions appear as coefficients in the series expansion of the indirect perturbation of a planet, that is the motion caused by perturbations of the Sun. In 1824 he developed Bessel functions more fully in a study of planetary perturbations

12. ... but who were they? 173 and published a treatise on them in Berlin. It was not the first time that special cases of the functions had appeared, for Jacob Bernoulli, Daniel Bernoulli, Euler and Lagrange had studied special cases of them earlier. In fact it was probably Lagrange's work on elliptical orbits that first led Bessel to work on such functions. Bezout, £tienne Bom: 31 March 1730 in Nemours, France. Died: 27 September 1783 in Basses-Loges (near Fontainbleau), France. Family tradition almost demanded that Etienne Bezout follow in his father's and grandfather's footsteps as a magistrate in the town of Nemours. However once he had read Euler's works he wished to devote himself to mathematics. In 1756 he published a memoir on dynamics and later two papers investigating integration. In 1758 Bezout was appointed an adjoint in mechanics of the Academie des Sciences, then he was appointed examiner of the Gardes de la Marine in 1763. One important task that he was given in this role was to compose a textbook specially designed for teaching mathematics to the students. Bezout is famed for the texbooks which came out of this assignment. The first was a four volume work which appeared in 1764-67. He was appointed to succeed Camus becoming examiner of the Corps d' Artillerie in 1768. He began work on another mathematics textbook and as a result he produced a six volume work which appeared between 1770 and 1782. This was a very successful textbook and for many years it was the book which students hoping to enter the Ecole Polytechnique studied. His books came in for a certain amount of criticism for lack of rigour but, despite this, they were understood by those who needed to use mathematics and as a result

were very popular and widely used. Their use spread beyond France for they were translated into English and used in North America. Bezout is famed also for his work on algebra, in particular on equations. He was much occupied with his teaching duties after appointments in 1763 and he could devote relatively little time to research. He made a conscientious decision to restrict the range of his work so that he could produce worthwhile results in a narrow area. The way Bezout went about his research is interesting. He attacked quite general problems, but since an attack was usually beyond what could be achieved with the mathematical knowledge then available, he attacked special cases of the general problems which he could solve. His first paper on the theory of equations examined how a single equation in a single unknown could be attacked by writing it as two equations in two unknowns. He made the simplifying assumption that one of the two equations was of a particularly simple form; for example he considered the case when one of the two equations had only two terms, a term of degree n and a constant term. Already this paper had introduced the topic to which Bezout would make his most important contributions,

174 Further Linear Algebra namely methods of elimination to produce from a set of simultaneous equations a single resultant equation in one of the unknowns. He also did important work on the use of determinants in solving equations and as a result Sylvester, in 1853. called the determinant of the matrix of coefficients of the equations the 'Bezoutiant*. These and further papers published by Bezout in the theory of equations were gathered together in Thiorie ginirale des Equations algebriques which was published in 1779. Cauchy, Augustin Louis Bom: 21 August 1789 in Paris, France. Died: 23 May 1857 in Sceaux (near Paris). France. In 1802 Augustin-Louis Cauchy entered the Ecole Centrale du Pantheon where, following Lagrange's advice, he spent two years studying classical languages. He took the entrance examination for the Ecole Poly technique in 1805 and the examiner Biot placed him second. At the Ecole Polytechnique he attended courses by Lacroix. de Prony and Hachette and was tutored in analysis by Ampere. In 1807 he graduated from the Ecole Polytechnique and entered the Ecole des Ponts et Chaussees, an engineering school. In 1810 Cauchy took up his first job in Cherbourg working on port facilities for Napoleon's English invasion fleet. In addition to a heavy workload he undertook mathematical research. Encouraged by Legendre and Malus. he submitted papers on polygons and polyhedra in 1812. Cauchy felt

that he had to return to Paris if he was to make an impression with mathematical research. In September of 1812 he returned after becoming ill. It appears that the illness was not a physical one and was probably of a psychological nature resulting in severe depression. Back in Paris. Cauchy investigated symmetric functions and in a 1812 paper he used 'determinant' in its modem sense. He reproved the earlier results on determinants and gave new results of his own on minors and adjoints. This 1812 paper gives the multiplication theorem for determinants for the first time. He was supposed to return to Cherbourg in February 1813 when he had recovered his health but this did not fit with his mathematical ambitions. He was allowed to work on the Ourcq Canal project rather than return to Cherbourg. Cauchy wanted an academic career but failed with several applications. He obtained further sick leave and then, after political events prevented work on the Ourcq Canal, he was able to devote himself entirely to research for a couple of years. In 1815 Cauchy lost out to Binet for a mechanics chair at the Ecole Polytechnique, but then was appointed assistant professor of analysis there. In 1816 he won the Grand Prix of the French Academy of Sciences for a work on waves. He achieved real fame however when he submitted a paper to the Institute solving one of Fermat's claims on polygonal numbers made to Mersenne. In 1817 when Biot left Paris for an expedition to the Shetland Islands in Scotland, Cauchy filled his post at the College de France. There he lectured on methods

 of integration which he had discovered, but not published, earlier. He was the first to make a rigorous study of the conditions for convergence of infinite series in addition to his rigorous definition of an integral. His text Cours a"Analyse in 1821 was designed for students at £cole Polytechnique and was concerned with developing the basic theorems of the calculus as rigorously as possible. He began a study of the calculus of residues in 1826. and in 1829 he defined for the first time a function of a complex variable. In 1826 Cauchy. in the context of quadratic forms in n variables, used the term 'tableau'for the matrix of coefficients. He found its eigenvalues and gave results on diagonalisation of a matrix in the context of converting a form to the sum of squares. He also introduced the idea of similar matrices (but not the term) and showed that if two matrices are similar then they have the same characteristic equation. He also proved, again in the context of quadratic forms, that every real symmetric matrix is diagonalisable. By 1830 Cauchy decided to take a break. He left Paris in September 1830, after the revolution in July, and spent a short time in Switzerland. Political events in France meant that Cauchy was now required to swear an oath of allegiance to the new regime and when he failed to return to Paris to do so he lost all his positions there. In 1831 Cauchy went to Turin and after some time there he accepted an offer from the King of Piedmont of a chair of theoretical physics. He taught in Turin from 1832. In 1833 Cauchy went to Prague to tutor Charles X's grandson. He returned to Paris in 1838 and regained his position at the Academy but not his teaching positions because he still refused to take an oath of allegiance. He was elected to the Bureau des Longitudes in 1839 but, after refusing to swear the oath, was not appointed and could not attend meetings or receive a salary. When Louis Philippe was overthrown in 1848 Cauchy regained his university positions. He produced 789 mathematics papers, an incredible achievement. Cayley, Arthur Bom: 16 August 1821 in Richmond, Surrey, England. Died: 26 January 1895 in Cambridge, Cambridgeshire, England. Arthur Cayley's father came from an English family, but lived in St Petersburg, Russia. It was there that Cayley spent the first eight years of his childhood before his parents returned to England. At school he showed great skill in numerical calculations and, after he moved to King's College School in 1835, his ability for advanced mathematics became clear. His mathematics teacher advised him to pursue mathematics rather than enter the family merchant business as his father wished. In 1838 Cayley began his studies at Trinity College, Cambridge from where he graduated in 1842 as Senior Wrangler. While still an undergraduate he had three

 papers published. After winning a Fellowship he taught for four years at Cambridge, publishing twenty-eight papers in the Cambridge Mathematical Journal. A Cambridge fellowship had a limited tenure so Cayley had to find a profession. He chose law and was admitted to the bar in 1849. He spent 14 years as a lawyer but, although very skilled in conveyancing (his legal speciality), he always considered it as a means to let him pursue mathematics. One of Cayley's friends was Sylvester who also worked at the courts of Lincoln's Inn in London and there they discussed deep mathematical questions throughout their working day. During his 14 years as a lawyer Cayley published about 250 mathematical papers. In 1863 Cayley was appointed Sadleirian Professor of Pure Mathematics at Cambridge. This involved a very large decrease in income but he was happy to have the chance to devote himself entirely to mathematics. This he certainly did, publishing over 900 papers and notes covering nearly every aspect of mathematics. In 1841 he published the first English contribution to the theory of determinants. Sylvester introduced the idea of a matrix but Cayley quickly saw its significance and by 1853 published a note giving, for the first time, the inverse of a matrix. In 1858 he published a memoir which is remarkable for containing the first abstract definition of a matrix. In it Cayley showed that the coefficient arrays studied earlier for quadratic forms and for linear transformations are special cases of this general concept. He defined a matrix algebra defining addition, multiplication, scalar multiplication and gave inverses explicitly in terms of determinants. He also proved that a 2 x 2 matrix satisfies its own characteristic equation. He stated that he had checked the result for 3 x 3 matrices, indicating its proof. The general case was later proved by Frobenius. As early as 1849 Cayley set out his ideas on permutation groups. In 1854 he wrote two papers which are remarkable for their insight into abstract groups. Before this only permutation groups had been studied and even this was a radically new area, yet Cayley defined an abstract group. Significantly he realised that matrices and quaternions could form groups. Cayley developed the theory of algebraic invariance, and his development of n- dimensional geometry has been applied in physics to the study of the space-time continuum. Cayley also suggested that euclidean and non-euclidean geometry are special types of a more general geometry and he united projective geometry and metrical geometry. Fibonacci, Leonardo Pisano Bom: 1170 in (probably) Pisa (now in Italy). Died: 1250 in (possibly) Pisa (now in Italy). Fibonacci, or Leonardo Pisano to give him his correct name, was bom in Italy but was educated in North Africa where his father held a diplomatic post. His father's

12. ... but who were they? 177 job was to represent the merchants of the Republic of Pisa who were trading in Bugia, a Mediterranean port in northeastern Algeria. Fibonacci travelled until around the year 1200 when he returned to Pisa. There he wrote a number of important texts which played an important role in reviving ancient mathematical skills and he made significant contributions of his own. Of his books Liber abbaci (1202), Practica geometriae (1220), Flos (1225), and Liber quadratorum have survived. Fibonacci's work was made known to Frederick II, the Holy Roman Emperor, through the scholars at his court These scholars suggested to Frederick that he meet Fibonacci when the court met in Pisa around 1225. At the court a number of problems were presented to Fibonacci as challenges. Fibonacci gave

solutions to some in Flos, a copy of which he sent to Frederick II. After 1228 only one reference to Fibonacci has been found. This was in a decree made in 1240 in which a salary was awarded to Fibonacci in recognition for the services that he had given to Pisa, advising on matters of accounting and teaching the people. Liber abbaci was based on the arithmetic and algebra that Fibonacci had collected during his travels. The book, which went on to be widely copied and imitated, introduced the Hindu-Arabic place-valued decimal system and the use of Arabic numerals into Europe. Indeed, although mainly a book about the use of Arab numerals, which became known as algorism, simultaneous linear equations are also studied in this work. Many of the problems that Fibonacci considers in Liber abbaci were similar to those appearing in Arab sources. The second section of liber abbaci contains a large collection of problems aimed at merchants. They relate to the price of goods, how to calculate profit on transactions, how to convert between the various currencies in use in Mediterranean countries, and problems which had originated in China. A problem in the third section led to the introduction of the Fibonacci numbers and the Fibonacci sequence for which Fibonacci is best remembered today. 'A certain man put a pair of rabbits in a place surrounded on all sides by a wall. How many pairs of rabbits can be produced from that pair in a year if it is supposed that every month each pair begets a new pair which from the second month on becomes productive?' The resulting sequence is 1,1,2,3,5,8,13,21,34,55,... (Fibonacci omitted the first term in Liber abbaci). This sequence, in which each number is the sum of the two preceding numbers, has proved extremely fruitful and appears in many different areas of mathematics and science. There are also problems involving perfect numbers, the Chinese remainder theorem and summing arithmetic and geometric series. Apart from his role in spreading the use of the Hindu-Arabic numerals and the rabbit problem, his contribution to mathematics has been largely overlooked.

178 Further Linear Algebra Fourier, Jean Baptiste Joseph Born: 21 March 1768 in Auxerre, Bourgogne, France. Died: 16 May 1830 in Paris, France. Joseph Fourier's first schooling was at Pallais's school, then he proceeded in 1780 to the Ecole Royale Militaire of Auxerre where at first he showed talents for literature. By the age of thirteen, however, mathematics became his real interest and within a year he had completed a study of the six volumes of Bezout's Cours de Mathematiques. In 1787 Fourier decided to train for the priesthood and entered the Benedictine abbey of St Benoit-sur-Loire. He was

unsure if he was making the right decision in training for the priesthood and he corresponded with the professor of mathematics at Auxerre. He submitted a paper on algebra to Montucla in Paris and his letters suggest that he really wanted to make a major impact in mathematics. Fourier did not take his religious vows. Having left St Benoit in 1789, he visited Paris and read a paper on algebraic equations at the Acad£mie Royale des Sciences. In 1790 he became a teacher at the Benedictine college, Ecole Royale Militaire of Auxerre, where he had studied. Up until this time there had been a conflict inside Fourier about whether he should follow a religious life or one of mathematical research. However in 1793 a third element was added to this conflict when he became involved in politics and joined the local Revolutionary Committee. He defended members of one political faction while in Orleans then returned to Auxerre to continue work on the revolutionary committee and to teach at the College. In July 1794 he was arrested, the charges relating to the Orleans incident, and imprisoned. Fourier feared that he would be put to the guillotine but, after Robespierre himself suffered that fate, political changes allowed him his freedom. Later in 1794 Fourier was nominated to study at the Ecole Normale in Paris. He was taught by Lagrange, Laplace and Monge then began teaching at the College de France where he again undertook research. He was appointed to the Ecole Polytechnique but repercussions of his earlier arrest remained. Again he was arrested and imprisoned but fortunately soon released. In 1797 he succeeded Lagrange to the chair of analysis and mechanics at the £cole Polytechnique. He was renowned as an outstanding lecturer but he does not appear to have undertaken research around this time. In 1798 Fourier joined Napoleon's army in its invasion of Egypt as a scientific adviser. He returned to France in 1801 with the remains of the expeditionary force and resumed his post as Professor of Analysis at the Ecole Polytechnique. However Napoleon had other ideas about how Fourier might serve him and he was sent to Grenoble where his duties as Prefect were many and varied. It was during his time in Grenoble that Fourier did his important mathematical work on the theory of heat. His work on the topic began around 1804 and by 1807

 he had completed his important memoir which today is very highly regarded but at the time it caused controversy. Referees were unhappy with the work because of an objection, made by Lagrange and Laplace in 1808, to Fourier's expansions of functions as trigonometrical series, what we now call Fourier series. With a rather mixed report there was no move to publish the work. When Napoleon escaped from Elba and marched into Grenoble, Fourier left in haste. However he was able to talk his way into favour with both sides and Napoleon made him Prefect of the Rhone. After Napoleon was defeated, Fourier returned to Paris. During his final eight years he resumed his mathematical research. Gram, Jorgen Pedersen Bom: 27 June 1850 in Nustrup, Denmark. Died: 29 April 1916 in Copenhagen, Denmark. Jorgen Gram, the son of a fanner, entered the Ribe Katedralskole secondary school in 1862. He began his university education in 1868. In 1873 Gram graduated with a Master's degree in mathematics. He had published his first important mathematics paper before he graduated. This was a work on modern algebra which provided a simple and natural framework for invariant theory. In 1875 Gram was appointed as an assistant in the Hafnia Insurance Company. Around the same time he began working on a mathematical model of forest management. His career in the Hafnia Insurance Company progressed well and his work for the company led him back into mathematical research. He began working on probability and numerical analysis, two topics whose practical applications in his day to day work in calculating insurance made their study important to him. Gram's mathematical career was always a balance between pure mathematics and very practical applications of the subject His work on probability and numerical analysis involved both the theory and its application to very practical situations. He published a paper on methods of least squares and for this work he was awarded the degree of Doctor of Science in 1879. Gram was led to study abstract problems in number theory. In 1884 he won the Gold Medal of the Videnskabernes Society for his paper Investigations of the number of primes less than a given number. Although he continued to work for the Hafnia Insurance Company in more and more senior roles, Gram founded his own insurance company, the Skjold Insurance Company, in 1884. He was the director of this company from its founding until 1910. From 1895 until 1910 Gram was also an executive of the Hafnia Insurance Company, and from 1910 until his death in 1916 was Chairman of the Danish Insurance Council. Despite not teaching mathematics in a university and as a consequence never having any students, Gram still managed to influence the next generation of Danish mathematicians in a very positive way. He often lectured in the Danish Mathematical Society, he was an editor of Tidsskriftfor Mathematik from 1883 to 1889, and he also reviewed papers written in Danish for German mathematical journals.

180 Further Linear Algebra Gram received honours for his mathematical contributions despite being essentially an amateur mathematician. The Videnskabemes Society had awarded him their Gold Medal in 1884 before he became a member, but in 1888 he was honoured with election to the Society. He frequently attended meetings of the Society and published in the Society's journals. For many years he was its treasurer. Gram is best remembered for the Gram-Schmidt orthogonalisation process which constructs an orthogonal set from an independent one. He was not however the first to use this method. The process seems to be a result of Laplace and it was essentially used by Cauchy in 1836. Gram met his death in a rather strange and very sad way. He was on his way to a meeting of the Videnskabemes Society when he was struck and killed by a bicycle. He was sixty-five years old when he met his death in this tragic accident. Hamilton, Sir William Rowan Bom: 4 August 1805 in Dublin, Ireland. Died: 2 September 1865 in Dublin, Ireland. By the age of five, Hamilton had already learned Latin, Greek, and Hebrew. He was taught these subjects by his uncle with whom he lived with in Trim for many years. He soon mastered additional languages but a turning point came in his life at the age of 12 when he met the American Zerah Colbum, who could perform amazing mental arithmetical feats and with whom Hamilton joined in competitions of arithmetical ability. It appears that losing to Colbum sparked Hamilton's interest in mathematics. Hamilton's introduction to mathematics came at the age of 13 when he studied Clairaut's Algebra. At age 15 he started studying the works of Newton and Laplace. In 1822 Hamilton found an error in Laplace's Michanique Cileste. He entered Trinity College, Dublin at the age of 18 and in his first year he obtained an optime in Classics, a distinction awarded only once in 20 years. Hamilton made remarkable progress for an undergraduate and before the end of 1824 submitted his first paper to the Royal Irish Academy on caustics. After an unhappy love affair he considered suicide, but instead turned to poetry, which he returned to throughout his life in times of anguish. In his final year as an undergraduate he presented a memoir Theory of Systems of Rays to the Royal Irish Academy. In 1827 he was appointed Professor of Astronomy at Trinity College while he was still an undergraduate aged twenty-one years. This appointment brought a great deal of controversy as Hamilton did not have much experience in observing. Before beginning his duties in this prestigious position, Hamilton toured England and Scotland (from where the family name originates). He met the poet Wordsworth and they became friends. One of Hamilton's sisters also wrote poetry and when

 Wordsworth came to visit them he indicated a preference for her poems rather than his. The two men had long debates over poetry versus science, the culmination of which was that Worsdworth had to tell Hamilton that his talents lay in the latter: You send me showers of verses which I receive with much pleasure.... yet have we fears that this employment may seduce you from the path of science.... Again I do venture to submit to your consideration whether the poetical parts of your nature would not find afield more favourable to their nature in the regions ofprose% not because those regions are humbler, but because they may be gracefully and profitably trod, with footsteps less careful and in measures less elaborate. In 1832 Hamilton published a third supplement to Theory of Systems of Rays which is essentially a treatise on the characteristic function applied to optics. Near the end of the work he applied the characteristic function to study Fresnel's wave surface. From this he predicted conical refraction and asked the Professor of Physics at Trinity College to try to verify his theoretical prediction experimentally. This happened two months later and brought him great fame. In 1833 Hamilton read a paper to the Royal Irish Academy expressing complex numbers as algebraic couples, or ordered pairs of real numbers. After the discovery of algebraic couples, he tried to extend the theory to triplets, and this became an obsession that plagued him for many years. On 16 October 1843 (a Monday) Hamilton was walking in along the Royal Canal with his wife to preside at a Council meeting of the Royal Irish Academy. Although his wife talked to him now and again Hamilton hardly heard, for the discovery of the quaternions, the first noncommutative algebra to be studied, was taking shape in his mind. He could not resist the impulse to carve the formulae for the quaternions j2 =/= ** = «;/* =-1 in the stone of Brougham Bridge as he passed it. Hamilton felt this discovery would revolutionise mathematical physics and spent the rest of his life working on quaternions. He published Lectures on Quaternions in 1853 but he soon realised that it was not a good book from which to learn the theory. Perhaps Hamilton's lack of skill as a teacher showed up in this work. Determined to produce a book of some lasting quality, Hamilton began to write Elements of Quaternions which he estimated would be 400 pages long and take two years to write. The title suggests that he modelled his work on Euclid's Elements. The book was twice the intended length and took seven years to write, the final chapter being incomplete when he died. It was published with a preface by his son. Hamilton's personal life was not a happy one, and was exacerbated by the suicide of his colleague James

MacCullagh at Trinity College. His ever-increasing depen- dancy on alcoholic stimulants brought little respite. He died from a severe attack of gout, shortly after receiving the news that he had been elected the first Foreign Member of the National Academy of Sciences of the United States of America.

182 Further Linear Algebra Hermite, Charles Born: 24 December 1822 in Dieuze, Lorraine, France. Died: 14 January 1901 in Paris, France. Hermite's parents did not take much personal interest in their children's education, although they did provide them with good schooling. He was something of a worry to his parents for he had a defect in his right foot which meant that he moved around only with difficulty. He attended the College de Nancy, then went to Paris where he attended the College Henri. In 1840-41 he studied at the College Louis-le-Grand where some fifteen years earlier Galois had studied. He preferred to read papers by Euler, Gauss and Lagrange rather than work for his formal examinations but he showed remarkable research ability, publishing two papers while at Louis-le-Grand. Hermite wanted to study at the Ecole Polytechnique and he took a year preparing for the examinations being tutored by Catalan in 1841-42. He passed but only attained sixty-eighth place in the ordered list. After one year at the Ecole Polytechnique he was refused the right to continue his studies because of his disability. The decision was reversed but strict conditions which Hermite did not find acceptable were imposed and he decided not to graduate from the Ecole Polytechnique. Hermite made friends with important mathematicians at this time and frequently visited Joseph Bertrand, whose sister Louise he later married. He exchanged letters with Jacobi which show that Hermite had discovered some differential equations satisfied by theta-functions and he was using Fourier series to study them. After spending five years working towards his degree he passed in 1847. In the following year he was appointed to the Ecole Polytechnique, the institution which had tried to prevent him continuing his studies some four years earlier. Hermite made important contributions to number theory and algebra, orthogonal polynomials, and elliptic functions. He discovered his most significant mathematical results over the ten years following his appointment to the Ecole Polytechnique. One topic on which Hermite worked was the theory of quadratic forms. This led him to study invariant theory and he found a reciprocity law relating to binary forms. With his understanding of quadratic forms and invariant theory he created a theory of transformations in 1855. The next mathematical result by Hermite which we must mention is one

for which he is rightly famous. Although an algebraic equation of the fifth degree cannot be solved in radicals, Hermite showed in 1858 that an algebraic equation of the fifth degree could be solved using elliptic functions. He applied these results to number theory, in particular to class number relations of quadratic forms. The year 1869 saw Hermite become a professor of analysis when he succeeded Duhamel both at the Ecole Polytechnique and at the Sorbonne. The 1870s saw Hermite return to problems concerning approximation and interpolation in which he had

 been interested earlier in his career. In 1873 he published the first proof that e is a transcendental number. He is now best known for Hermite polynomials, Hermite's differential equation, Hermite's formula of interpolation and hermitian matrices. He resigned his chair at the Ecole Polytechnique in 1876 but continued to hold the chair at the Sorbonne until he retired in 1897. Hilbert, David Born: 23 January 1862 in Konigsberg, Prussia (now Kaliningrad, Russia). Died: 14 February 1943 in GOttingen, Germany. David Hilbert attended the gymnasium in his home town of Konigsberg. After graduating from this school, he entered the University of Konigsberg. There he went on to study under Lindemann for his doctorate which he received in 1885. Hilbert was a member of staff at Konigsberg from 1886 to 1895, being a Privatdozent until 1892, then as Extraordinary Professor for one year before being appointed a full professor in 1893. In 1895 Klein appointed Hilbert to the chair of mathematics at the University of Gottingen, where he continued to teach for the rest of his career. Hilbert's eminent position in the world of mathematics after 1900 meant that other institutions would have liked to tempt him to leave Gottingen and, in 1902, the University of Berlin offered him a chair. Hilbert turned down the offer, but only after he had used it to bargain with Gottingen and persuade them to set up a new chair to bring his friend Minkowski there. Hilbert contributed to many branches of mathematics, including invariant theory, algebraic number fields, functional analysis, integral equations, mathematical physics, and the calculus of variations. His first work was on invariant theory and, in 1888, he proved his famous Basis Theorem. He discovered a completely new approach which proved the theorem for any number of variables in an entirely abstract way. Although this proved that a finite basis existed, his methods did not construct such a basis. Hilbert's revolutionary approach was difficult for others to appreciate and he had to argue his case strongly before Klein agreed that the work could be published. In 1893 while still at Konigsberg Hilbert began his Zahlberichton algebraic number theory. The German Mathematical Society had requested this major report three years after the Society was created in 1890. The ideas of the present day subject of class field theory are all contained in this work. Hilbert's work in geometry had the greatest influence in that area after Euclid. A systematic study of the axioms of Euclidean geometry led Hilbert to propose 21 such axioms and he analysed their significance. He published Grundlagen der Geometric in 1899, putting geometry in a formal axiomatic setting. The book continued to appear in new editions and was a major influence in promoting the axiomatic approach to mathematics which has been one of the major characteristics of the subject throughout the 20th century.

184 Further Linear Algebra In his famous speech, delivered at the Second International Congress of Mathematicians in Paris, his 23 problems challenged (and still today challenge) mathematicians to solve fundamental questions. The ideas of a Hilbert space came out of his work on integral equations. It was in this context that he introduced the words 'eigenvalue'and 'eigenfunction*. In his work with his student Schmidt, the ideas of abstract infinite dimensional spaces evolved around 1904. However, the fully axiomatic approach did not appear until Banach's 1920 doctoral dissertation. In 1930 Hilbert retired and the city of Konigsberg made him an honorary citizen. His address ended with words showing his devotion to solving mathematical problems: 'We must know, we shall know*. Jordan, Marie Ennemond Camille Born: 5 January 1838 in Lyon, France. Died: 22 January 1922 in Milan, Italy. Camille Jordan entered the Ecole Polytechnique to study mathematics in 1855. His doctoral thesis was examined in January 1861, after which he worked as an engineer, first at Privas, then at Chalon-sur-Saone, and finally in Paris. From 1873 he was an examiner at the Ecole Polytechnique where he became professor of analysis in November 1876. He was also a professor at the College de France from 1883 although until 1885 he was at least theoretically still an engineer by profession. It is significant, however, that he found more time to undertake research when he was an engineer. Jordan was a mathematician who worked in a wide variety of different areas essentially contributing to every mathematical topic which was studied at that time: finite groups, linear and multilinear algebra, the theory of numbers, the topology of polyhedra, differential equations, and mechanics. Topology (called analysis situs at that time) played a major role in some of his first publications which were a combinatorial

approach to symmetries. He introduced important topological concepts in 1866 building on his knowledge of Rie- mann's work. He introduced the notion of homotopy of paths looking at the deformation of paths one into the other. He defined a homotopy group of a surface without explicitly using group terminology. Jordan was particularly interested in the theory of permutation groups. He introduced the concept of a composition series and proved the Jordan-Holder theorem. Jordan clearly saw classification as an aim of the subject, even if it was not one which might ever be solved. He made some remarkable contributions to how such a classification might proceed setting up a recursive method to determine all soluble groups of a given order. His work on group theory done between 1860 and 1870 was written up into a major text which he published in 1870. This treatise gave a comprehensive study of

 Galois theory as well as providing the first ever group theory book. The treatise contains the 'Jordan normal form* theorem for matrices, not over the complex numbers but over a finite field. Jordan's use of the group concept in geometry in 1869 was motivated by studies of crystal structure. He went on to produce further results of fundamental importance. He studied primitive permutation groups and, generalising a result of Fuchs on linear differential equations, he was led to study the finite subgroups of the general linear group of n x n matrices over the complex numbers. Although there are infinite families of such finite subgroups, Jordan found that they were of a very specific group theoretic structure which he was able to describe. Another generalisation, this time of work by Hermite on quadratic forms with integral coefficients, led Jordan to consider the special linear group of n x n matrices of determinant 1 over the complex numbers acting on the vector space of complex polynomials of degree m in n indeterminates. Jordan is best remembered today among analysts and topologists for his proof that a simply closed curve divides a plane into exactly two regions, now called the Jordan curve theorem. Among Jordan's many contributions to analysis should also be mentioned his generalisation of the criteria for the convergence of a Fourier series. Kronecker, Leopold Bom: 7 December 1823 in Liegnitz, Prussia (now Legnica, Poland). Died: 29 December 1891 in Berlin, Germany . Leopold Kronecker was taught mathematics at Liegnitz Gymnasium by Kum- mer who stimulated his interest in mathematics. Kummer immediately recognised Kronecker's talent for mathematics and he took him well beyond what would be expected at school, encouraging him to undertake research. Kronecker became a student at Berlin University in 1841 and there he studied under Dirichlet and Steiner. He did not restrict himself to mathematics, however, for he studied other topics such as astronomy, meteorology, chemistry and philosophy. After spending the summer of 1843 at the University of Bonn, where he went because of his interest in astronomy rather than mathematics, he travelled to the University of Breslau for the winter semester of 1843-44. He went there because he wanted to study mathematics with his old school teacher Kummer who had been appointed to a chair at Breslau in 1842. He spent a year there before returning to Berlin for the winter semester of 1844-45 where he worked on his doctoral thesis on algebraic number theory under Dirichlet's supervision. Just as it looked as if he would embark on an academic career, Kronecker left Berlin to manage the banking business of his mother's brother. He also managed a family estate but still found the time to continue working on mathematics, although he did this entirely for his own enjoyment.

186 Further Linear Algebra In 1855 Kronecker came to Berlin but not to a university appointment. He did not lecture at this time but was remarkably active in research, publishing a large number of works in quick succession. These were on number theory, elliptic functions and algebra, but, more importantly, he explored the interconnections between these topics. Kronecker was elected to the Berlin Academy which gave him the right to lecture at the University and this he did beginning in 1862. The topics on which he lectured were very much related to his research: number theory, the theory of equations, the theory of determinants, and the theory of integrals. As we have already indicated, Kronecker's primary contributions were in the theory of equations and higher algebra, with his major contributions in elliptic functions, the theory of algebraic equations, and the theory of algebraic numbers. However the topics he studied were restricted by the fact that he believed in the reduction of all mathematics to arguments involving only a finite number of steps. Kronecker is well known for his remark: 'God created the integers, all else is the work of man*. He was the first to doubt the significance of non-constructive existence proofs. It appears that, from the early 1870s, he was opposed to the use of irrational numbers, upper and lower limits, and the Bolzano-Weierstrass theorem, because of their non- constructive nature. Kronecker had no official position at Berlin until Kummer retired in 1883 when he was appointed to the chair. By 1888 Weierstrass felt that he could no longer work with Kronecker and decided to go to Switzerland, but then,

realising that Kronecker would be in a strong position to influence the choice of his successor, he decided to remain in Berlin! Lagrange, Joseph-Louis Born: 25 January 1736 in Turin, Sardinia-Piedmont (now Italy). Died: 10 April 1813 in Paris, France. Joseph-Louis Lagrange studied at the College of Turin where his favourite subject was classical Latin. At first he had no enthusiasm for mathematics, finding Greek geometry rather dull. Lagrange decided to make mathematics his career after reading Hal ley's 1693 work on the use of algebra in optics. In 1754, despite having no training in advanced mathematics, he published his first mathematical work. It was no masterpiece and suffered because Lagrange was working without the advice of a mathematical supervisor. Next he studied the tautochrone, the curve on which a particle will always arrive at a fixed point in the same time independent of its initial position. By the end of 1754 his discoveries on this topic made a major contribution to the new subject of the calculus of variations. Lagrange was appointed professor of mathematics at the Royal Artillery School in l\irin in 1755. It was well deserved for he had already shown great originality and depth of thinking.

12. ... but who were they? 187 After Euler proposed him for the Berlin Academy, Lagrange was elected in 1756. The following year he was a founding member of a scientific society, later the Royal Academy of Science, in l\irin. This new Society began publishing a scientific journal Melanges de Turin with Lagrange a major contributor to the first few volumes with papers on the calculus of variations, calculus of probabilities, foundations of dynamics, propagation of sound, and the integration of differential equations with various applications to topics such as fluid mechanics. The libration of the Moon was announced as the topic for the 1764 prize competition of the Acad£mie des Sciences. Lagrange entered the competition, sending his entry to Paris in 1763. It arrived there not long before Lagrange himself but he took ill shortly after. Returning to Turin in early 1765, he submitted an entry for the 1766 prize on the orbits of the moons of Jupiter. D'Alembert, who had visited the Berlin Academy and was friendly with Frederick II of Prussia, arranged for Lagrange to be offered a position in the Berlin Academy. Despite Lagrange's poor position in Turin, he at first refused but later accepted. He succeeded Euler as Director of Mathematics at the Berlin Academy of Science in 1766. For 20 years Lagrange worked at Berlin, producing a stream of top quality papers and regularly winning the prize from the Academie des Sciences in

Paris. He shared the 1772 prize on the three body problem with Euler, won the prize for 1774, another one on the motion of the moon, and he won the 1780 prize on perturbations of the orbits of comets by the planets. His work covered many topics: astronomy, the stability of the solar system, mechanics, dynamics, fluid mechanics, probability, and the foundations of the calculus. He also worked on number theory proving in 1770 that every positive integer is the sum of four squares. In 1771 he proved Wilson's theorem that /lis prime if and only if (n-1)! + 1 is divisible by n. His 1770 paper Reflexions surla risolution algibrique des Equations made a fundamental investigation of why equations of degrees up to 4 could be solved by radicals. For the first time the roots of an equation are considered as abstract quantities rather than having numerical values. He studied permutations of the roots and, despite not composing permutations in the paper, it was a first step in the development of group theory. Lagrange, in a paper of 1773, studied identities for 3 x 3 functional determinants. However this comment is made with hindsight since Lagrange himself saw no connection between his work and that of Laplace and Vandermonde. This paper contains the volume interpretation of a determinant for the first time. Lagrange showed that the tetrahedron formed by 0(0,0,0) and the three points M{xt y, z), M'(x't y', z'), M"{x\ y", z") has volume i[z(jt'y" - y'jc") + z'(yjc" -xy") + z"(jty' -yx')]. In 1787 Lagrange left Berlin for the Academie des Sciences in Paris, where he remained for the rest of his career. His Micanique Analytique, written in Berlin, was published in 1788. It summarised the field of mechanics and transformed the topic

 into a branch of mathematical analysis. Lagrange was the first professor of analysis at the Ecole Poly technique, appointed for its opening in 1794. In 1795 the Ecole Normale was founded with the aim of training school teachers and Lagrange taught courses on elementary mathematics there. Napoleon named Lagrange to the Legion of Honour and Count of the Empire in 1808. He was awarded the Grand Croix de 1'Ordre Imperial de la Reunion a week before he died. Laplace, Pierre-Simon Born: 23 March 1749 in Beaumont-en-Auge, Normandy, France. Died: 5 March 1827 in Paris, France. Pierre-Simon Laplace attended a Benedictine priory school in Beaumont-en- Auge. His father expected him to make a career in the Church and, at the age of 16, he entered Caen University and enrolled in theology. However Laplace soon discovered his love of mathematics. He left Caen without taking his degree, and went to Paris with a letter of introduction to d'Alembert from his teacher at Caen. Although only 19 years old he quickly impressed d'Alembert who tried to find him a position so he could support himself. Laplace was soon appointed as professor of mathematics at the Ecole Militaire and he began producing a steady stream of remarkable mathematical papers. In 1773 he was elected to the Academie des Sciences. By this time he had read 13 papers to the Academy in less than three years on: difference equations, differential equations, mathematical astronomy and the theory of probability. In 1772 Laplace claimed that the methods introduced by Cramer and Bezout were impractical and, in a paper where he studied the orbits of the inner planets, he discussed the solution of systems of linear equations using determinants. Rather surprisingly Laplace used the word 'resultant'for what we now call the determinant; surprising since it is the same word used by Leibniz yet Laplace must have been unaware of Leibniz's work. Laplace gave the expansion of a determinant which is now named after him. His papers in the 1780s would make Laplace one of the most important and influential scientists that the world has seen. It was not achieved, however, with good relationships with his colleagues. Although d'Alembert had been proud to have considered Laplace as his protege\ he certainly began to feel that Laplace was rapidly making much of his own life's work obsolete and this did nothing to improve relations. In 1784 Laplace was appointed as examiner at the Royal Artillery Corps, and in this role in 1785, he examined and passed the 16-year-old Napoleon Bonaparte. Laplace was made a member of the committee of the Academie des Sciences to standardise weights and measures in 1790. This committee worked on the metric

 system and advocated a decimal base. In 1793 the Reign of Terror commenced and the Academie des Sciences, along with the other learned societies, was suppressed on 8 August. The weights and measures commission was the only one allowed to continue but soon Laplace was thrown off the commission. With his wife and two children, he left Paris, not returning until after July 1794. Laplace presented his famous nebular hypothesis in 1796 which viewed the solar system as originating from the contracting and cooling of a large, flattened, and slowly rotating cloud of incandescent gas. He put all his work in this area into his great work the Traiti du Micanique Cileste published in 5 volumes, the first two in 1799. Under Napoleon, Laplace was a member, then chancellor, of the Senate, and admitted to the Legion of Honour in 1805. However Napoleon, in his memoirs written on St Helene, says he removed Laplace from the office of Minister of the Interior after only six weeks '... because he brought the spirit of the infinitely small into the government'. Laplace became Count of the Empire in 1806 and he was named a marquis in 1817 after the restoration of the Bourbons. The first edition of Laplace's Thiorie Analytique des Probabilites was published in 1812. In 1814 Laplace supported the restoration of the Bourbon monarchy and cast his vote in the Senate against Napoleon. The Hundred Days were an embarrassment to him the following year and he conveniently left Paris for the critical period. After this he remained a supporter of the Bourbon monarchy and became unpopular in political circles. When he refused to sign the document of the French Academy supporting freedom of the press in 1826, he lost his remaining friends in politics. Lie, Marius Sophus Born: 17 December 1842 in Nordfjordeide, Norway. Died: 18 February 1899 in Kristiania (now Oslo), Norway. Sophus Lie first attended school in the town of Moss then, in 1857, he entered Nissen's Private Latin School in Christiania. He decided to take up a military career, but his eyesight was not sufficiently good so he gave up the idea and entered the University of Christiania. At university Lie studied a broad science course, and he attended lectures by Sylow in 1862. Surprisingly he graduated in 1865 without having shown any great ability for mathematics, or any real liking for it. He considered a career in astronomy or botany or zoology or physics and in general became rather confused. From 1866 he began to read more and more mathematics and his interests steadily turned in that direction. In 1867 Lie had his first brilliant new mathematical idea. It came to him in the middle of the night and, filled with excitement, he rushed to tell a friend. Lie wrote a short mathematical paper in 1869, which he published at his own expense, based on

 this inspiration. He wrote up a more detailed exposition, but the world of mathematics was too cautious to quickly accept Lie's revolutionary notions. The Academy of Science in Christiania was reluctant to publish his work, and at this stage Lie began to despair that he would become accepted in the mathematical world. The breakthrough came later in 1869 when Crelle's Journal accepted his paper. Setting off near the end of 1869, Lie visited Gottingen and then Berlin. In Berlin he met Kronecker, Kummer and Weierstrass but he was not attracted to Weierstrass's mathematics. However he met Felix Klein and the two instantly found common ground in mathematics. In the spring of 1870 Lie and Klein were together again in Paris where Camille Jordan succeeded in a way that Sylow did not, for Jordan made Lie realise how important group theory was for the study of geometry. Lie's new ideas later appeared in his work on transformation groups. While Lie and Klein thought deeply about mathematics in Paris, the political situation between France and Prussia deteriorated. Lie decided to remain but became anxious as the German offensive met with an ineffective French reply. When the German army trapped part of the French army in Metz, Lie decided it was time to leave and he decided to hike to Italy. He reached Fontainebleau but there he was arrested as a German spy, his mathematics notes being assumed to be top secret coded messages. Only after the intervention of Darboux was he released from prison. In 1871 Lie became an assistant at Christiania and also taught at Nissen's Private Latin School. He submitted a dissertation On a class of geometric transformations for his doctorate which was duly awarded in July 1872. Lie started examining partial differential equations, hoping that he could find a theory which was analogous to the Galois theory of equations. This led to combining the transformations in a way that Lie called an infinitesimal group, what is today called a Lie algebra. When Klein left the chair at Leipzig in 1886, Lie was appointed to succeed him. In Leipzig, his life was rather different from that in Christiania. He was now in the mainstream of mathematics and students came from many countries to study under him. However his health was already deteriorating when he returned to a chair in Christiania in 1898, and he died of pernicious anaemia soon after taking up the post. Parseval des Chenes, Marc-Antoine Born: 27 April 1755 in Rosieres-aux-Saline, France. Died: 16 August 1836 in Paris, France. Very little is known of Antoine Parseval's life. We know that he was born into a family of high standing in France and he describes himself as a squire, certainly suggesting that his family were wealthy land-owners. One of the few pieces of information which exists is that he married Ursule Guerillot in 1795. The marriage certainly did not last long and the pair were soon divorced. The starting point of the historical events which were to play a major role in Parseval' life was the storming of the Bastille on 14 July 1789. Parseval, perhaps

 not surprisingly since he was of noble birth, was a royalist so for him the increasing problems for the monarchy meant that his life was more and more in danger. In 1792 Louis 16th gave up attempts at a compromise with opponents of the monarchy and tried to flee from France. He did not make it but was arrested and brought back to Paris. It was a time of great danger for royalist supporters and indeed it proved so for Parseval who was imprisoned in 1792. Following the execution of the King on 21 January 1793 there followed a reign of terror with many political trials. Parseval, despite being freed from prison, remained in fear of his life. Napoleon became 1st Consul in 1800 and then Emperor in 1804. Parseval should have kept his head down if he wanted to avoid trouble but it was a time when it was almost impossible not to get drawn into the political events. Parseval published poetry against Napoleon's regime and, not surprisingly, had to flee from France when Napoleon ordered his arrest. He was successful in avoiding arrest and was able to return to Paris. Parseval had only five publications, all presented to the Academie des Sciences between 5 May 1798 and 7 May 1804. It was the second of these, dated 5 April 1799, which contains the result known today as Parseval's theorem. Today this theorem is seen in the context of Fourier series, and often also in more abstract settings which are quite far removed from Parseval's original ideas. The original theorem was concerned with summing infinite series and had some restrictive conditions which Parseval removed in a note appended to an 1801 paper. The theorem was not published until his five papers were all published by the Academie des Sciences in 1806. Before that it was known by members of the Academy and appeared in works by Lacroix and Poisson before Parseval's papers were printed. Parseval was never honoured with election to the Academie des Sciences. He was proposed on five separate occasions, namely in 1796, 1799, 1802, 1813 and 1828. He was never particularly close although he did come third in 1799, the year that Lacroix was elected. It would not be unfair to say that Parseval has fared well in having a well known result, which is quite far removed from his contribution, named after him. However he remains a somewhat shadowy figure and it is hoped that research will one day provide a somewhat better understanding of his life and achievements. Schmidt, Erhard Born: 13 January 1876 in Dorpat, Russia (now Tartu, Estonia). Died: 6 December 1959 in Berlin, Germany. Erhard Schmidt attended his local university in Dorpat before going to Berlin where he studied with Schwarz. His doctorate was obtained from the University of Gottingen in 1905 under Hilbert's supervision. His doctoral dissertation concerned

 integral equations. The main ideas of this thesis appeared in a paper in 1907 of which we give more details below. After obtaining his doctorate he went to Bonn where he was awarded his habituation in 1906. Schmidt went on to hold positions in Zurich, Erlangen and Breslau before he was appointed to a professorship at the University ofBerlininl917. Clearly Schmidt's organisational abilities were recognised outside mathematics for he was appointed Dean for the academic year 1921-22 and the Vice-Chancellor of the University of Berlin during the years 1929-30. The inaugural address that he gave when taking up the post of Vice-Chancellor was entitled On certainty in mathematics. After the end of World War II Schmidt was appointed as Director of the Mathematics Research Institute of the German Academy of Science, remaining in that role until 1958. By that time he had retired from his chair, which he did in 1950. Another role which he took on after the end of the war was as the first editor of Mathematische Nachrichten, a journal he co-founded in 1948. Schmidt's main interest was in integral equations and Hilbert space. He took various ideas of Hilbert on integral equations and combined these into the concept of a Hilbert space around 1905. Hilbert had shown that integral equations with symmetric kernels had real eigenvalues, and the solutions corresponding to these eigenvalues he called eigenfunctions. He also expanded functions related to the integral of the kernel function as an infinite series in a set of orthonormal eigenfunctions. Schmidt published a two part paper on integral equations in 1907 in which he reproved Hilbert's results in a simpler fashion, and also with less restrictions. In this paper he gave what is now called the Gram-Schmidt orthononnalisation process for constructing an orthonormal set of functions from a linearly independent set. He then went on to consider the case where the kernel is not symmetric and showed that in that case the eigenfunctions associated with a given eigenvalue occurred in adjoint pairs. We should note, however, that Laplace presented the Gram-Schmidt process before either Gram or Schmidt. In 1908 Schmidt published an important paper on infinitely many equations in infinitely many unknowns, introducing various geometric notations and terms which are still in use for describing spaces of functions and also in inner product spaces. Schmidt's ideas were to lead to the geometry of Hilbert spaces and he must certainly be considered as a founder of functional analysis. Schmidt defined a space H whose elements are square summable sequences of complex numbers, defining a norm in terms of the inner product. He also defined orthogonal elements, showing that a set consisting of pair-wise orthogonal elements is linearly independent. Again he gave the Gram-Schmidt orthononnalisation process in this setting. He also studied projections and spectral resolutions. What are today called Hilbert-Schmidt operators also appear in this 1908 paper.

12. ... but who were they? 193 Schwarz, Hermann Amandus Born: 25 January 1843 in Hermsdorf, Silesia (now Sobiecin, Poland). Died: 30 November 1921 in Berlin, Germany. Hermann Schwarz studied at the Gymnasium in Dortmund where his favourite subject was chemistry. He intended to take a degree in chemistry and he entered the Gewerbeinstitut, later called the Technical University of Berlin, with this aim. However, it wasn't long before Kummer and Weierstrass had influenced him to change to mathematics. Through Karl Pohlke, one of his teachers, Schwarz became interested in geometry. He attended Weierstrass's lectures on the integral calculus and his interest in geometry was soon combined with Weierstrass's ideas of analysis. He continued to study in Berlin, being supervised by Weierstrass, until 1864 when he was awarded his doctorate. His doctoral dissertation was examined by Kummer, his future father-in-law. While in Berlin, Schwarz worked on minimal surfaces, a characteristic problem of the calculus of variations. He had made an important contribution in 1865 when he discovered what is now known as the Schwarz minimal surface. This minimal surface has a boundary consisting of four edges of a regular tetrahedron. Schwarz continued studying in Berlin for his teacher's training qualification and completed this by 1867 when was appointed as a Privatdozent to the University of Halle. In 1869 he became professor of mathematics at the Eidgenossische Technis- che Hochschule in Zurich then, in 1875, was appointed to the chair of mathematics at Gottingen University. Schwarz accepted the professorship in Berlin in 1892, but after this the university lost its leading role for mathematics. That this happened shouldn't have come as a complete surprise to those making the appointment, for Schwarz had published his Complete Works in 1890, two years earlier. Schwarz continued teaching at Berlin until 1918. Outside mathematics he was the captain of the local Voluntary Fire Brigade and, more surprisingly, he assisted the stationmaster at the local railway station by closing the doors of the trains! One important area which Schwarz worked on was that of conformal mappings. His most important work was a Festschrift for Weierstrass's 70th birthday in which he answered the question of whether a given minimal surface really yields a minimal area. An idea in this work, where he constructed a function using successive approximations, led Emile Picard to his existence proof for solutions of differential equations. It also

contains the inequality for integrals now known as the 'Schwarz inequality'. The fact that Schwarz should have come up with a special case of the general result now known as the Cauchy-Schwarz inequality (or the Cauchy-Bunyakovsky- Schwarz inequality) is not surprising, for much of his work is characterised by looking at rather specific and narrow problems but solving them using methods of great

 generality which have since found widespread applications. That he found such general methods says much for his great intuition which was perhaps based on a deep feeling for geometry. The form in which the inequality is usually presented today with its standard modern proof was first given by Weyl in 1918. Schwarz's demeanour has been described as naive, dramatic, and coarse. In spite of giving the impression of self-confidence, he was, in fact, rather insecure and besides, not efficient in business matters. Sylvester, James Joseph Born: 3 September 1814 in London, England. Died: 15 March 1897 in London, England. James Joseph Sylvester attended two primary schools in London, then his secondary schooling was at the Royal Institution in Liverpool. In 1833 he became a student at St John's College, Cambridge. At this time it was necessary for a student to sign a religious oath to the Church of England before graduating and Sylvester, being Jewish, refused to take the oath. For the same reason he was not eligible for a Smith's prize nor for a Fellowship. From 1838 Sylvester taught physics for three years at the University of London, one of the few places which did not bar him because of his religion. At the age of 27 he was appointed to a chair in the University of Virginia but he resigned after a few months. A student who had been reading a newspaper in one of his lectures insulted him and Sylvester struck him with a sword stick. The student collapsed in shock and Sylvester, believing (wrongly) that he had killed him, fled and boarded the first ship back to England. On his return Sylvester worked as an actuary and lawyer but gave mathematics tuition. Cayley was also a lawyer and they worked together at the courts of Lincoln's Inn in London. Sylvester tried hard to return to being a professional mathematician and, after several failed applications, became professor of mathematics at Woolwich. Sylvester did important work on matrix theory, discussing the topic with Cayley while they were at the courts of Lincoln's Inn. In particular he invented the term 'matrix' in 1850, defining a matrix to be an oblong arrangement of terms which he saw as something which led to various determinants from square arrays contained within it. He used matrix theory to study higher dimensional geometry. In 1851 he discovered the discriminant of a cubic equation and first used the name 'discriminant' for such expressions in higher order equations. He defined the nullity of a square matrix in 1884. Being at a military academy Sylvester had to retire at age 55. His only book at this time was on poetry. Clearly Sylvester was proud of this work for he sometimes signed himself 'J. J. Sylvester, author of The Laws of Verse'. For three years Sylvester did no mathematical research but then Chebyshev visited London and

 the two discussed mechanical linkages which can be used to draw straight lines. Sylvester then lectured on this topic at evening lectures at the Royal Institution entitled On recent discoveries in mechanical conversion of motion. In 1877 Sylvester accepted a chair at Johns Hopkins University in the United States and in 1878 he founded the American Journal of Mathematics, the first mathematical journal in the USA. In 1883 Sylvester, although 68 years old at this time, was appointed to the Sav- ilian chair of Geometry at Oxford. However he liked to lecture only on his own research and this was not well liked at Oxford. In 1892 Sylvester, by this time partially blind and suffering from loss of memory, returned to London where he spent his last years at the Athenaeum Club. Toeplitz, Otto Born: 1 August 1881 in Breslau, Germany (now Wroclaw, Poland). Died: 15 February 1940 in Jerusalem (then under British Mandate). After completing his secondary education in Breslau, Otto Toeplitz entered the university there to study mathematics. After graduating, he continued with his studies of algebraic geometry at the University of Breslau, being awarded his doctorate in 1905. In 1906 Toeplitz went to Gottingen and studied under Hilbert who was working on the theory of integral equations at the time. Toeplitz began to rework the classical theories of forms on spaces of finite dimension to spaces of infinite dimension. He wrote five papers directly related to Hilbert's spectral theory. Also during this period he published a paper on summation processes and discovered the basic ideas of what are now called the 'Toeplitz operators'. When Toeplitz arrived in Gottingen, Hellinger was a doctoral student there. The two quickly became friends and they would collaborate closely for many years. It was not until 1913 that Toeplitz was offered a teaching post as extraordinary professor at the University of Kiel. He was promoted to ordinary professor at Kiel in 1920. A joint project with Hellinger to write a major encyclopaedia article on integral equations, which they worked on for many years, was completed during this time and appeared in print in 1927. In 1928 Toeplitz accepted an offer of a chair at the University of Bonn. Toeplitz was dismissed from his chair by the Nazis in 1935. He did not leave Germany but rather remained and worked hard to support the Jewish community who were finding life increasingly difficult. He emigrated to Palestine (as it then was) in 1939 and helped in the building up of Jerusalem University. He had great plans for modernizing the university but unfortunately he became very ill and died a year after his arrival. Toeplitz worked on infinite linear and quadratic forms. In the 1930's he developed a general theory of infinite dimensional spaces and criticised Banach's work

196 Further Linear Algebra as being too abstract. In a joint paper with Kothe in 1934, Toeplitz introduced, in the context of linear sequence spaces, some important new concepts and theorems. Kothe has described how they would find the same idea cropping up in different contexts. Toeplitz was also very interested in the history of mathematics. For example he wrote The Calculus: A Genetic Approach, an excellent book on the history of the calculus. It was originally published posthumously in German in 1949 edited by Kothe. A historical topic which interested him deeply was the relation between Greek mathematics and Greek philosophy. He was a frequent visitor to the Frankfurt Mathematics Seminar in the 1920s and 30s, where his friend Hellinger worked from 1914, and there the history of mathematics played a large role. Toeplitz believed that only a mathematician of stature is qualified to be a historian of mathematics. It was not only the history of mathematics that interested him outside his research area. He wrote a popular book on mathematics in collaboration with Rademacher. This work, The Enjoyment of Mathematics, has been reprinted many times over the years. Toeplitz was also greatly interested in school mathematics and devoted much time to it. Vandermonde, Alexandre Theophile Born: 28 February 1735 in Paris, France. Died: 1 January 1796 in Paris, France. Alexandre-Theophile Vandermonde graduated licencie in 1757. His first love was music and his instrument was the violin. He pursued a music career and he only turned to mathematics when he was 35 years old. It was Fontaine des Berlins whose enthusiasm for mathematics rubbed off on Vandermonde. He was elected to the Acad6mie des Sciences in 1771 after his first paper was read to the Academy. This and three further papers which he presented between 1771 and 1772 represent his total mathematical output. In 1777 Vandermonde published the results of

experiments he had carried out with Bezout and Lavoisier on the effects of a very severe frost which had occurred in 1776. Ten years later, with Monge and Bertholet, he published two papers on manufacturing steel. He was so close to Monge that he was often called 'la femme de Monge'. In 1778 he put forward the idea that musicians should ignore all theory and rely solely on their trained ears when judging music. Positions which Vandermonde held include director of the Conservatoire des Arts et Metiers in 1782 and chief of the Bureau de rHabillement des Armees in 1792. In the same year of 1792 he sat with Lagrange on a committee of the Acad6mie des Sciences which had to examine a newly invented musical instrument. He was involved with designing a course in political economy for the Ecole Normale, after it was founded in October 1794.

12. ... but who were they? 197 The name of Vandermonde is best known today for the Vandermonde determinant, yet it does not appear in his four mathematical papers. It is rather strange, therefore, that this determinant should be named after him. Lebesgue conjectured that it resulted from someone misreading Vandermonde's notation, and therefore believing that this determinant was in his work. The first of Vandermonde's papers presented a formula for the sum of the m-lh powers of the roots of an equation. It also presented a formula for the sum of the symmetric functions of the powers of such roots. Kronecker claimed in 1888 that the study of modem algebra began with this paper. Cauchy stated that Vandermonde had priority over Lagrange for ideas which eventually led to the study of group theory. In his second paper Vandermonde considered the problem of the knight's tour on the chess board. He considered the intertwining of curves generated by the moving knight, an early example of topological ideas. In his third paper Vandermonde studied combinatorial ideas, and in the final paper he studied the theory of determinants. Muir claims Vandermonde was 'the only one fit to be viewed as the founder of the theory of determinants'. The reason for Muir's claim is that, although determinants had been studied earlier by Leibniz, all earlier work had simply used the determinant as a tool to solve linear equations. Vandermonde, however, thought of the determinant as a function and gave functional properties. He showed the effect of interchanging two rows and of interchanging two columns. From this he deduced that a determinant with two identical rows or two identical columns is zero.

13 Solutions to the Exercises 1.1,1.2 Routine verification of the axioms. 13 ||x + y||2 = {x+>|x+y) = (x|^ + (y|x) + ^|y) + (y|>)=||x||2+||y||2+2(x|y).If xty£ IR2 are linearly independent and if L is the line joining x to the origin then the line through y perpendicular to L meets L at the point Xx where (y -Xx \ x) - 0. Clearly, X = (x \ y)/ \\x\\2. If i? is the angle xOy then rcj- IMI.NIHI _ \(*\y)\ 11)11 11)11 1*1 M and so we have the cosine law ||x + yfl2 = ||x||2 + ||y||2 + 2 ||x|| ||>|| cos tf. 1.4 Expanding ||x + y||2 - i ||ix + yfl2 in a complex inner product space we obtain Ml2 + \M 2 + (*\y) + (y*)- H Ml2 + b 2 + <" I >> + (? I "*>] = IMI2 + >l12-'(IN2+IMI2) + 2(x|y>. 1.5 |x + y||2 + H*-y||2= IW|2+ M2 + (x\y) + (y\x)+ IW|2+ H-y||2 + (x| -y) + {-y | x) = 2 ||x ||2 + 2 ||y||2. In IR2 this result says that the sum of the squares of the lengths of the diagonals of a parallelogram is the sum of the squares of the lengths of its sides. 1.6 If IN = llvll then (x + y\x-y) = (x|x) + (y\x) + <x| -v) + (y\ -y) = |x||2 - ||v||2 = 0. In IR2 this says that the diagonals of a rhombus are mutually perpendicular. 1*7 | f'f(x)g(x)dx\ < (// \f(x)\2dx)W(!ba \8(x)\2dx)W 1.8 | tr B*A\ < | tr A*A\lf2\ tr B*B\W. {E,, | En) = tr E*,,E,, = tr EqpE,, = { 1 ifr = p, s-q\ 1.9 If y - Xx then clearly |(x | y)\ = \X\ \\x\\2 = ||x|| ||>||. Conversely, suppose that equality holds. We may assume that y f 0. Let a - (x\y)/ ||y|j2. Let z = x - ay. Then, by the hypothesis, (y|x) = {y\ay) + (y\z) = (x|y) + (yjz). It follows that (y|z) = 0. Since M = 1*1 I h\\ we then have 11^1^=11^ + ^1^ = 1^^^+11212 whence ||z||2 = 0 and consequently z — 0. Hence x = ay. 1.10 a-\. 1.11 This is a standard exercise in calculus. 1.12 Yes.

13. Solutions to the Exercises 199 1.13 Let Xl = (1,1,1), x2 = (0,l,l),x3 = (0,0,1). Takey, = x,/|M = ^(1,1,1). Now X2-Wyi)yi = 4(-2,1,1) so take y2 =^(-2,1,1). Next, Xi-(xi\y2)y2 -{*i\y\)y\ = 5(0,1, -l) so lake>3 = -^-(0,1,-1). An orthonormal basis is then {yi,>2,>3}. 1.14 Letxi = (1,1,0,1), Jt2 = (1,-2,0,0), jc3 = (1,0,-1,2). Then {xt ,x2txj} is linearly independent. Takey, = *,/!*, I = ^(1,1,0,1). Now ^2-(^21 >!>>.= 4(4,-5,0,1) so take y2 = ^(4, -5,0,1). Next, *3 -{xi\y2)y2 -{xi\yi)yi = £(-4,-2,-1,6) so lake y3 = ^(-4. ~2, -1,6). An orthonormal basis is then {^1,^2^3}. 1.15 Let jci = (1, /, 0), x2 = (1,2,1 - i). Recall the standard inner product of Example 1.2. Take>, = V INI = ^(U0). Now x2 - (x21>i)y, = (J + i, 1 - |i, 1 -1) so take y2 = ^ (5 +1,1 - ji, 1 - i). An orthonormal basis is then {yi, y2}. 1.16 We have/i : f»-»1 and/2: t»->t. Then ^1/.)=/^=1 so we can take gi = f\ as the first vector in the Gram-Schmidt process. Now (f2\8i)= [tdt=\ Jo and so/2 - (/2 \g\)g\ is the mapping h: 11-» f - \. Since (*I h) = /'[MOf A = [V -1 + j)</f = £, we can take g2 to be the

mapping t >-* 2\/3(f -1). Then {gi,g2} is an orthonormal basis. 1.17 (1) Observe that (X|X2) = f X3 = 0 and so {X,X2} is orthogonal. It therefore suffices to normalise both vectors. Now {X\X) = f X2 = § and (X2 |X2) = / X4 = f and so an orthonormal basis is {J^X, ^fX2}. (2) Let xi = 1, x2 = 2X - 1, jc3 = 12X2. Then we have (*i |xj) = 2 so we take y, = jc./|x,|| = ^5. Next, <*2|yi) = ^(2X-l) = -v6

 and so x2 - (x2 |>i)>, = 2X - 1 + 1 = 2X. So we take y2 = 2X/ ||2X|| = $X. Finally, (•*31 >2> = 0 and (x31 >i) = 4\/2~ and so *3 - <*31 y2>>2 - (xy | yi>>, = 4(3X2 -1) and so we take y3 = ^(3X2 -1). 1.18 Let {ei, e2) be the standard orthonormal basis of IR2. Let x and y have polar coordinates n.iJi and r2,i?2 respectively. Then cos(i?2-i?i)=^^ and similarly cos fr = x ' , sin i?i = x ' '. Parseval's identity, namely (x|y> = (x|e,)<ei |>> + <x |e2>(e2 |y), is then equivalent to r\ncos(i?2 ~$\) = rir2cos i?i cos <d2 + nr2 sin i?i sin tf2 from which we may remove the factor r\r2. 1.19 Bessel's inequality becomes *=-i» The rest of the question is a standard exercise in calculus. 1.20 That B is an orthonormal basis is clear from the fact that B is linearly independent and orthonormal. If (2,-3,1) = a(j-t0t fy + /3(-^,0, ^) + 7(0,1,0) then the Fourier coefficients a, /3,7 are given by a « ((2,-3,1)1(^,0,£))=£; /3= ((2,-3,1)1(-^,0,^)) = -^: 7 = ((2,-3,1)1(0,1,0)) = -3. 1.21 If M = [m;y]nxn then for all ij we have/(*y) = mye,-. Consequently (/"(*>) I ei) = (mc/«i I «i) = mu(ei I e,) = m/y. 1.22 Suppose first that a\,..., a, are linearly dependent. Then some fly is a linear combination of the ak with A ^ ;. Consequently, for every i, (a,-1 Ay) is a linear combination of the (a-, \ at) with k f j. Thus the;-th column of G* is a linear combination of the other columns and therefore det CA = 0. Suppose now that a\,..., a, are linearly independent. Then they form a basis for the sub- space that ihey generate. By Gram-Schmidt, let \b\,..., b,} be an orthonormal basis for this subspace. Let AT be the transition matrix from the ordered basis (a,), to the ordered basis (6,),. Then we have (a, I fly) = (£ mikbk Y^mj,b,) = mntnfl + ml2mj2 + ■ ■ ■ + mirmjr = [A/Af];>• *=i »=1 I Hence G* = MM' and so det G* = (det A/)2. Since transition matrices are invertible, we have det M f 0 whence det GA f 0.

 1.23 (l)=> (2): If/ is an isomorphism then \\f(x)\\2 = {f{x)\f(x)) = (x|x) = ||x||2 (2) =*(l):If(2) holds then l* + yU2 = ll/U+y)H2 = V(*+y)\A*+y)) = \\x\\2 + {f(x)\f{y)) + {f{y)\f(x)) + IMI2. Since ||*+y||2=||*|2 + <*|y)+<y|x)+IM|2 it follows that (f(x) \f(y)) + {f(y) |/(*)> = (x | y)

+ <y | x), whence the result if F = IR. Doing the same with |ix + yfl2 we see similarly that {fix) |/(>)> - {f(y) |/(x)) = (x |y) - (y | x). It follows that {f{x) |/(y)> = (x | y). (2) =► (3): If (2) holds then |/W -/(y)‖ = |lflx -y)‖ = flx -y‖. (3) =» (2): If (3) holds then \\f(x) -/(y)| = ‖x -y\\ and so, on taking y = 0, we have 1/WI = M- 1.24 We have [p(af+ bg)](x) = x(a/ + bg)(x) = x[a/(x) + **(*)] = xaf(x) + xfctfx) = Mfl/)l(jO + M^)](*) and so <p(af + bg) = p(fl/) + <p(bg). Thus p is linear. It preserves inner products since M/)M«)>= [[<pV))(x) [<p(8)](x)dx= f x2f(x)g(x)dx={f\g). Jo Jo However, it is not an isomorphism since it fails to be surjective. To see this, consider the constant function g : x i-» 1 in V. There is no/e V such that xf{x) = 1 for every x £ [0,1]. 2.1 Lett/, = {{Otatb)\ atb£ IR} and(/2 = {(a.O.fc); a,fc€ IR}. Then Vn(/, = {0} = VnU2. Since we then have IR3 = V© U\ = V© U2. 22 We have Ker t\ - {(0,0,0)} and so, by the dimension theorem, Imfi = IR3. Kerf2 = {(atata); a£ IR} and Imf2 = {(a,bt0); atb£R). Ker f3 = {(0,0,0)}, Imr3 = IR3. Kerf4={(0,0,fl); a € IR} and Im f4 = {{atbtb); atb£\R). For i = 1,2,3,4 we see that Im f.nKer r, = {(0,0,0)}. It follows by the dimension theorem that dim(Im f, + Ker f,) = dim IR3 and therefore Im r, + Ker f, = IR3. Hence IR3 = Im f, © Ker r,-. 2.3 Let dim V = /i and let {a j,..., ar} be a basis of A. Extend this to a basis {ait...tanblt...tb,,-r} of V. Then the subspace B that is spanned by {b\,..., b,,-r) is such that A © Z? = V. 2.4 Every line L of non-zero gradient that passes through the origin is a subspace of IR2 such thatX©L=IR2. 2.5 The cartesian product A x B is a vector space with dim A x B = dim A + dim B = dim V. The mapping f : A x B -* V given by f(at b) = a - b is linear with Im/ = A + B and Ker/ = {(x, x) ; x £ A n /?} ~ A n /?. By the dimension theorem we have dim A x B = dim Im/ + dim Ker /. Thus dim V = dim (A + B) + dim (A n B). Since A + B = V, it follows that dim(A n B) = 0 whence A n fl = {0v}. Hence V = A © B.

 2.6 Suppose that V = Im/© Ker/. Qearly, Im/2 C Im/. If now x € Im/ then x = /(a) for some a £ V, and a = p + 0 wherepe Im/and 0 € Ker/. Then x = /(p + q) = /(p) + Ov = /(p) € Im/2. Hence Im/C Im/2 and we have equality. Conversely, suppose that Im/ = Im/2. Then for every x € V we have/(x) = /2(>) for some > € V. This gives x -f(y) € Ker/and so x = /(>) + [x - f(y)) € Im/ + Ker/. Hence V — Im/ + Ker/. Observe now from the dimension theorem that dim Ker/= dim V-dim Im/= dim V-dim Im/2 = dim Ker/2. Since Ker/ C Ker/2 we deduce that Ker/ = Ker/2. Suppose now that x € Im/n Ker/. Then x = /(>) and/(x) = 0V. Consequently/2(y) = Ov so y £ Ker/2 = Ker/ whence Ov = /(>) = x. Hence Im/n Kerf- {Ov} and therefore V = Im/© Ker/.

2.7 It is clear that we have the chains VD Im/D Im/2 D ■•■ D Im/" D Im/m+I D ■••; {Ov} C Ker/C Ker/2 C ■ • C Ker/" C Ker/n+I C ■ • ■. Now we cannot have an infinite sumber of strict inclusions in the first chain since X c ^ implies that dim X < dim Y, and the dimension of V is finite. It follows that there is a smallest positive integer p such that Im/P = Im/**1 = • ■ ■. Since dim V - dim Im/P + dim Ker/r a corresponding result holds for the kernel chain. Writing g = /' we see that Img = Img2. It then follows by the previous exercise that V = Im/' © Ketf. 2.8 (1) is routine. (2) If x € Vi then x = x\{b\ + fc4) + x2(b2 + 63); so Vi is spanned by {b\ + fc4, k + &j}. Also, if xi(fci + fc4) + X2{bi + 63) = Ov then since B is a basis of V we have xi = X2 = 0. Thus Bi = {fci + 64,62 + 63} is a basis of Vi. Similarly, B2 = {b\ -64,62 -63} is a basis of V2. (3) It is clear from the definitions that Vj n V2 = {Ov}. Consequently, the sum Vj + V2 is direct. Since V\, V2 are each of dimension 2 and V is of dimension 4, it follows that V = Vj © V2. (4) As for the transition matrix from the basis B to the basis B\ u B2t we observe that 6, = £(6,+64) + 0(62 + 63) + 0(62-63) + J(6,-64) 62 = 0(6, + 64) + £(62 + 63) + £(62-63) + 0(6,-64) 63 = 0(6, + 64) + £(62 + 63) - £(62 -63) + 0(6, -64) 64 = £(6, + 64) + 0(62 + 63) + 0(62 -63) - £(6, -64). The matrix in question is therefore P- \ 0 0|-i i 0 0 £■ 0 0 (S) It is readily seen that 2P2 = /4 and so /»-' = 2/» = "10 0 1" 0 110 0 1-10 10 0-1 Suppose now that A/ is centro-symmetric; i.e. is of the form abed efgh hgfe d c b a

 Let tp be the linear mapping represented by M relative to the ordered basis B. Then the matrix of ip relative to the ordered basis C is given by PMP~\ which is readily seen to be of the form a p 0 0" „ 7 6 0 0 K~ oo£c 0 0 TJ i? Thus if M is centro-symmetric then it is similar to a matrix of the form K. 2.9 For every x € V we have *=(idv+/)(±x) + (idv-/)(±x) and so V = Im (idv + /) + Im (idv -/)• If now x £ Im (idv +/) n Im (idv -/) then for some y, z € V we have x = y +/(y) = z -/(z). Since by hypothesis/2 = idv, it follows that /W = /(y)+/2(y) = /(y) + y=x; /(*) = /(*)-/¾) = /(*)-*=-*. Hence x = 0v and therefore V = Im (idv +/) © Im (idv -/)• If d2 = /,, let/ be a linear mapping that is represented by A relative to some ordered basis of V. Then/2 = idv. Let {ai,...,ap} be a basis of Im (idv +/) and let {a^it...tan} be a basis of Im (idv -/). Then (fli,..., a,,) is a basis of V. Now since a\ = b + /(ft) for some b£ V wehave/(fl|) = f(b) + f2(b) = f(b) + b = A|, and similarly fora2,---,^- Likewise, a^i-c -f(c) for some c € V so/(fl^.|) = /(c) -/2(c) = /(c) - c = -flp+i, and similarly for flp+2, • ■ •, flu- Hence the matrix of/ relative to the ordered basis {a \,..., a,,} is *p = 0 -/,,-,, and A is then similar to this matrix. Conversely, if A is similar to a matrix of the form Ar then there is an invertible matrix Q such that Q~lAQ = Ap. Consequently, A2 = {Q\Q~1)2 = CAP2C"' = CM2"' = In- 2.10 Extend A to the basis {(1,0,1), (-1,1,2), (0,1,0)} of R3. Then, relative to this basis we have (1,2,1) = (1,0,1) + 2(0,1,0) and thereforepA(1,2,1) = (1,0,1). 2.11 If/ is the projection on A parallel to B then for x = a + b with a £ A and b£Bvtt have /(x) = a. Consequently (idv -/)(•*) = * -/(•*) = x -a = fc and so idv -/ is the projection on £ parallel to A. 2.12 Clearly, p\ +/¾ is a projection (=idempotent) if and only if, denoting composites by juxtspo- sition, p\pi + pip\ = 0. It is therefore clear that (2) =>• (1). (1) =j» (2): Suppose thatpi+p2 is a projection. Multiplying each side of p\Pi+pip\ — Oon the left by pi we obtain pipi+pifrpi = 0; and multiplying each side on the right by p\ we obtain P\PiP\ + PiPi = 0. It follows that p\P2 = pipi. But p\p2 + pipi = 0. Hence pifr = P2P1 = 0. Suppose now that p\ + /¾ is a projection. If x € Im (pi + /¾) then there exists y such that x = pi(y) + ^(y) whence x £ Impi + Imp2- Thus Im(p\ + /¾) C Impi + Imp2- The reverse inclusion holds since if x = pi (y)+p2(z) then applying p\ to this and using the fact that pip2 = 0 we obtainpi(x) = pi(y), and similarly p2(x) = ^(z). Hence x = pi(x)+p2(x) = (pi+#)(*)€ Im (pi + pi). It suffices now to show that the sum Impi + Imp2 is direct. For this purpose, let x £ Impi n Imp2. Then x = pi(y) = pi(z) for some ytz. Then x = pi(x) = p^d) = 0V. If x € Ker (pi + pi) then pi(x) + pi[x) = 0v. Applying pi and p^ and using the fact that p,pj = 0 = pip\, we obtain pi(x) = 0v = pi(x). Hence Ker(pi + pi) C Ker pi n Ker p2-

Conversely, if x £ Ker p\ n Ker pi then px{x) - Ov = pi(x) and so (pi + pi){x) = Ov, whence we have the reverse inclusion. 2.13 (1) ^> (2): Suppose that Impi = Imp2. Then for every x £ V there exists y £ V such that p\(x) = P2(y). Applying pi to this, we obtain P2Pi(x) = piiy) - p\(x), whence P2P1 = p\. Similarly, P1P2 = p2. (2) =>• (1): If (2) holds then for every x £ V we have p\{x) = pi\p\{x)\ £ Impi whence Impi C Imp2. Similarly, we have the reverse inclusion. 2.14 Since Impi = • • • = Imp* we have, by the previous exercise, p,p, = pj for all ij. Now if * P - £ ^tPt ^ p2 = Xjp2 + • • • + \\p\ + X1X2P1P2 + M^\PiP\ + • • - + >*>*-iP*p*-i = (i>i)AiPi + - + (l>)Atft i=l i-l = P and so p is also a projection. To show that Imp = Impi it suffices to prove that ppi = px and pip = p. As for the first of these, we have ppi = (*iPi + • • • + ^kPk)p\ = Xipi + • • • + \hP\ = Pi, and the second is obtained similarly. 2.15 (idv +/)(idv - 5/) = idv - \f+/ - 5/2 and since/ is idempotent the right hand side is idv- Hence idv +/ is invertible with inverse idv - \f- 2.16 The set {(1,0,1), (1,2, -2)} is linearly independent so can be extended to a basis of H"\ for example fl= {(1,0,1),(1,2,-2),(1,0,0)}. An application of the Gram-Schmidt process yields the orthonormal basis {^(1,0,0,^(3,4,-3), ^(2,-3,-2)}. Recall from the orthonormalisation process that the subspace spanned by {(1,0,1), (1,2, -2)} coincides with the subspace spanned by {^(1,0,1), ^j(3,4, -3)}; and if we denote this sub- space by W| then R3 = Wx © W2 where W2 = Span{(2,-3,-2)}. Then W2 is the desired orthogonal complement. 3.1/(jt,0,z,0)=(x + 2z,0,z,0)€W. 3.2 If/(x) = flx + fcthen [*>(/)](*) = x f {at + b)dt = x{\a + b) Jo so <p(f)£W. 33 (1) If Imp is /-invariant then for every x £ V there exists y £ V such that/p(x) = p(y). Then pfp{x) - p2(y) = p(y) - fp(x) and so pfp = fp. Conversely, if pfp = fp then for every x € V we have/p(x) = pfp(x) € Imp so that Imp is/-invariant. (2) By the Corollary to Theorem 2.4, Ker p is /-invariant if and only if Im (idv - p) is /-invariant. From the above, this is the case if and only if (idv -p)/(idv -p) = (idv ~p)f. Expanding each side, we see that this is equivalent to pfp — pf.

13. Solutions to the Exercises 205 3.4 Ufog = idv then/ is surjective, hence bijective since V is of finite dimension. Then g = /'" and so g of = idv. Suppose that W is a subspace of V that is /-invariant, so that/-(WO C W. Since/ is an isomorphism we have dim f~{W) = dim W and so/^(W) = W. Hence IV = g[f(W)] = g{W) and so IV is g-invariant. For the converse, interchange/ and g. The result is false for infinite-dimensional spaces. For example, consider the real vector space IR[X] of polynomials over IR. Let/ be the differentiation mapping and let g be the integration mapping. Each is linear and we have/o g = id but g of ^ id. n n 3.5 (a) Since V — 0 V{ , every x £ V can be expressed uniquely in the form x = £ *i r=I #=i n where xt £ V,- for each i. Then/(x) = £/(x,) where/(x,) £ V,- since by hypothesis each V,- is n /-invariant. Hence Im/ = 52 'm/i- Also, since each V,- is /-invariant, for each i we have i=i im/, n £ im/> c v, n E v> = {<M- Hence Im/ = 0 Im/,. i=I (ft) If now x € Ker/ then 0V = /(x) = £/(x,) = £/,(x,) € £ Im/,. Since the sum is i=i i=i i=i n direct it follows that each//(x,) = 0v whence x, £ Ker/j. Consequently, x £ £ Ker/, and so i=i Ker / = 51 Ker//. But Ker/,- C V/ for each i, and so i=i Ker/, n 5: Ker/, C V, n £ Vj = {0v}. m ifl It follows that Ker/ = 0 Ker/,. »=i 3.6 /(x, y, z) = (2x + y -z, -2x -y + 3z, z). Relative to the standard ordered basis, the matrix of/ is "2 1 -1" A= -2 -1 3 0 0 1 It is readily seen that cA(X) = X(X -1)2 = mA(X). Thus IR3=Ker/©Ker(/"-id)2 with Ker / of dimension 1 and Ker(/" - id)2 of

dimension 2. Now f(*,y,z) =(0,0,0) <=> 2x + y=0 = z and so a basis for Ker/ is {(-1,2,0)}. Also, (/•-id)(x,>,z)=(x + >-z,-2x-2y + 3z,0) and therefore (/-id)2(x,>,z)=(-x-> + 2z,2x + 2>-4z,0). Soa basis for Ker (f-id)2 is {(1,-1,0), (1,1,1)}. An ordered basis with respect to which the matrix of/ is in block diagonal form is then B= {(-1,2,0),(1,-1,0),(1,1,1)}.

206 Further Linear Algebra The transition matrix from B to the standard basis is -1 1 1 P= 2-11 0 0 1 and the block diagonal matrix that represents/ relative to B is "0 0 0" 3.7 f{xtytz) = (x + y + z,x + y + z,x + y + z). Relative to the standard ordered basis, the matrix of/ is M = It is readily seen that cM(X) - - X2(X - 3) and that mM{X) = X{X - 3). The matrix "3 0 0" N = "1 1 1 1 1 1 1" 1 1 also has these characteristic and minimum polynomials. So M and N reduce to the same block diagonal form and therefore are similar. 3.8 If/3 = / then/ satisfies the equation X{X + 1)(X - 1) = 0. The minimum polynomial of / must therefore be a product of distinct linear factors. Hence/ is diagonalisable. 3.9 f[xtytz) = {-2x -y + zt2x + y-3zt-z). Relative to the standard ordered basis of F3 the matrix of/ is ^-2 -1 1" 2 1 -3 0 0-1 A = It is readily seen that the eigenvalues are 0 and -1, the latter being of algebraic multiplicity 2. The minimum polynomial is X{X + 1)2. Consequently,/ is not diagonalisable. 3.10 {a) There are three distinct eigenvalues, namely 2,3,6 so/ is diagonalisable. (b) The minimum polynomial is (X - 1)(X - 2)2 so/is not diagonalisable. (c) The minimum polynomial is {X -2)3 so/is not diagonalisable. 4.1/(Jr,y,z)= (-jr-y-zt0tjr+y+z)so/2(jrtytZ) = /(-jr-y-z,0tjr+y+z)= (0,0,0). 4.2 The matrix that represents/ relative to the given basis is A = Since A3 = 0 we have that/3 = 0 so/ is nilpotent. 4.3 If \ is an eigenvalue of/ then X" is an eigenvalue off = 0. 4.4 fA(X) = AX-XA gives fl(X) = A2X - 2AXA + XA\ fi{X) = A3X - 3A2XA + 3AXA2 - XA3, and so on. If therefore A" = 0 we have/,"*1 = 0. 4.5 f(x, y, z) = (2x + y - z, -2x - y + 3z, z). A solution is provided as in Exercise 3.6, but here we proceed as in Examples 4.4 and 4.5. As before, we have IR3=Ker/ ©Ker(/"-id)2 r-5 -8 -5 1 4] 1 7 I 4

13. Solutions to the Exercises 207 with Ker / of dimension 1 and Ker{f - id)2 of dimension 2. Begin by choosing V| £ Ker/; as before we can take vi = (-1,2,0). Next, we have tf-id)(x,>tZ)=(x + >-Z|-2x-2>+3zt0) so we choose a non-zero v2 € Ker(/" - id), for example v2 = (1,-1,0). Next we choose v3 independent of v2 and such that (f - id)(v3) = av^. For example (taking a = 1) we choose v3 =(1,1,1). Then B = {vi,v2,V3} is an ordered basis with respect

to which the matrix of/ is upper triangular. 4.6 /(x, y, z) = (2x - 2y,x - >, -x + 3y + z). The matrix of/ relative to the standard basis of IR3is "2-2 0" A= 1-10 -1 3 1 It is readily seen that cA(X) = X[X - I)2 = mA(X). Thus IR3=Ker/©Ker(/'- id)2 with Ker / of dimension 1 and Ker(/"- id)2 of dimension 2. Begin by choosing V| £ Ker/. say vi = (1,1,-2). Now {f - id)(x, y, z) = (x - 2>, x - 2>, -x + 3>) so we choose a non-zero vj € Ker(/" - id), for example 1¾ = (0,0,1). Next we choose vj independent of v2 and such that (f — id)(v3) = av2. For example (taking a = 1) we choose v3 =(2,1,0). An ordered basis relative to which the matrix of/ is upper triangular is then *= {(1,1,-2),(0,0,1),(2,1,0)}. 4.7 In Exercise 4.5, relative to the ordered basis B the matrix of/ along with its Jordan decomposition is -D + N. ro o oi 0 1 1 0 0 1 = ro o oi 0 1 0 0 0 1 + ro o oi 0 0 1 0 0 0 The transition matrix P from B to the standard basis is -1 1 1 2 -1 1 0 0 1 /» = Simple computations give PDP~l = />"> = 1 1 2 1 0 0 -2" -3 1 2 1-2 -2 -1 4 0 0 1 PNP~X = 0 0 1 0 0-1 0 0 0 Thus the diagonal and nilpotent parts of/ are dj{x,y,z) - (2x + y-2z,-2x-y + 4z,z); n/(xtytz)= (z,-z,0).

[As a check, note that df + nf = f] Applying the same procedure to the linear mapping of Exercise 4.6, we obtain Mx. >. *) = (2* - 2y, x - y, ~2x + Ay + z); "/(*,>, z) = (0,0,x->). 4.8 The matrix of/ relative to the basis {b\,b%, b$) is -1 3 0 0 2 0 2 1-1 It is readily seen that the minimum polynomial is (X - 2){X + 1)2. Hence F3 = Ker (/" - 2id) © Ker(/" + id)2 where Ker(/" - 2id) is of dimension 1 and Ker(/" + id)2 is of dimension 2. First observe that/(61 + bi + by) = 2(ft| + 62 + ^3) so we can choose V| = b\ + bi + ft3 £ Ker(/"-2id). Next, observe that a+id)(M = 2*3; a+id)(fc2)= 3ft, + 3*2 + ^3; tf+id)(ft3)=0. We require a non-zero vi € Ker (/"+ id); so take vi = by. Next, we need to choose v3 independent of V2 and such that {f + id)(v3) = av2. Taking a = 1 we can choose v3 = b\. An ordered basis with respect to which the matrix of/ is upper triangular is then B={bi+ ^ + ^,^,^). The transition matrix from B to the standard basis is p = "1 0 1" 1 0 0 1 1 0 The matrix of/ relative to B is P~XAP = Simple calculations give PDP~l = -1 3 0 2 0 3 • t "2 0 0 0' 0 -1 0 -1 0 1 />-' = "0 1 0 -1 1 -1 0' 2 -1 = D + N. #W#»_I = "0 0 2 0" 1 0 0 0' 0 0 -2 0 Thus the diagonal and nilpotent parts of/are (relative to the basis {b\,b2M}) dfih) = -*,; nf{jbx) = 2by, d/ih) = 3bi + 2b2 + 3ft3; ji/(fc) = -2ft3; df(h) = -by nf(by) = 0.

4.9 For the given matrix A = "0 0 1 0 0 1 -1] 1 1 the minimum polynomial is {X + 1){X -1)2. If the linear mapping/ is represented by A relative to the standard basis then a basis with respect to which the matrix of/ is upper triangular is 5=((1,-2,1),(-1,0,1),(1,1,0)), the upper triangular form being = D + tf. -1 " 1 1 1 = r-i o o" 0 1 0 0 0 1 + ro o o] 0 0 1 0 0 0 The transition matrix P from B to the standard basis is I -1 1 />= -2 0 1 1 1 0 />"'=! Then p-*AnP = (p-'AP)n = (D + AO" = 1 -1 2 (-1)" 0 0 -1 1 2 0 1 0 1 3 2 0 n 1 from which A" can be obtained. 5.1 If x f 0v is such that/*-'(x) f 0V then for every k < p -1 we have/*(x) f 0V. To show that (x,/(x),... ,/7^1 (x)} is linearly independent, suppose that \0x + X|/(x) + ■ • ■ + \p-xT\x) - 0. Applying/7^1 to this we obtain \off>~l{x) = 0 whence \0 = 0. Thus we have \if(x) + --- + ^1/^^) = °- Applying f'2 to this we obtain similarly \ i = 0. Continuing in this way we see that every \; — 0 and consequently the set is linearly independent. 5.2 If/ is nilpotent of index n = dim V then (x,/(x),... ,f~l(x)} is a basis of V. The matrix of/ relative to this ordered basis is then that in the question. Conversely, if there is an ordered basis with respect to which the matrix of/ is of the given form then to see that/ is nilpotent of index n it suffices to observe that the matrix M in question is such that AT = 0 and A/"-1 ^ 0. 5.3 The eigenvalues are 2 of algebraic multiplicity 1. and 1 of algebraic multiplicity 3. The rank of the matrix A/ - /4 is 3 so for any linear mapping / represented by the matrix we have dim Ker(/"- id) = 4 -3 = 1. Thus there is a single Jordan block associated with the eigenvalue 1. The Jordan form is then 5.4 The eigenvalues are 2 of algebraic multiplicity 3, and 3 of algebraic multiplicity 2. The rank of the matrix A/ - 2/5 is 3 so for any linear mapping/ represented by the matrix we have dim Ker(/-2id) = 5-3 = 2. Thus there are two Jordan blocks associated with the eigenvalue

 2. Similarly there is only one Jordan block associated with the eigenvalue 3. Hence the Jordan form is '2 1 0 2 2 3 1 0 3 5.5 A basis is {1,X,X2,X3}, relatve to which the matrix of D is 0 10 0 0 0 2 0 0 0 0 3 0 0 0 0 The characteristic and minimum polynomials are X4 so the Jordan form is 0 10 0 0 0 10 0 0 0 1 0 0 0 0 5.6 We have v,(x2) = 2x>, y> (xy)=|y2. ¥>(r) = 0, ¥>(*) = >> ¥>(>) = 0, y>(l)=0 and therefore A/ = The characteristic polynomial of M is X6 and the minimum polynomial is X\ The Jordan form of M therefore has the eigenvalue 0 six times down the diagonal with at least one Jordan block of size 3x3. The number of Jordan blocks is dim Ker <p. Clearly, a basis for Ker ip is {1, y, y2}. Thus the Jordan form is y = 5.7 (1) Let A be the matrix of Exercise S.4 and let/4 : MatSx| IR -» Mat5xi IR be the linear mapping whose matrix relative to the standard basis of Mat Sx)

IR is A, so that/4(x) - Ax. We now proceed as in Example 5.8. Begin by choosing V| e Im {f - 2id) n Ker(/" - 2id), eg- V| = [1 0 1 0 0J. Next, choose V2 independent of V| such that (/ - 2id)(v2) = V|, eg. v2 = [0 I 0 I 0J. Next, choose v3 e Ker(/^ - 2id) such that {v1.v2.v3} is linearly independent, e.g. v3 = [2 1 0 0 1]. Now repeat the process with/ - 3id to obtain, for example, v4 = [-1 0 0 1 0] and v5 = [2 0 0 0 lj. A Jordan basis is then {11 0 1 0 0J,[0 1 0 1 0],[2 1 0 0 1J,[-1 0 0 1 0],[20 0 0 1]}.

13. Solutions to the Exercises 211 (2) A Jordan basis is {/, ,/2,/3,/4} where 0/i = 0, 0/2=/,, £)/3=/2, 0/4=/3. Choose/, = 1. Then/2 = X,/3 = \X2*U = \X%. So a Jordan basis is {6.6X, af2,**}. 5.8 (1) To find a Jordan basis, first choose v, e ImpflKer tp, e.g. v,(x,y) = jy2. Next choose v2 independent of v, and such that <p(v2) = v,, e.g. v2(x,y) = xy. Next choose v3 independent of v,, v2 with p(v3) = v2, e.g. v3(x,y) = jx2. Now choose v4 e Im tp n Ker p independent of v,, V2, V3, e.g. v4(x, y) = y. Next choose v5 independent of v,, V2, V3, v4 with p(v5) = v4, e.g. v5(x, y) = x. Finally, choose v6 e Ker ip independent of v,,..., vs, e.g. v6(x, y) = 1. Then a Jordan basis is B = {J^.xy, \x2tytx, 1}. (2) The transition matrix from B to the given basis is P- and is such that P~lAP = J. 5.9 The given matrix A has characteristic polynomial X\X -1)2, so the eigenvalues are 0 and 1. The transition matrix from the given basis of IR5 to the standard basis is 01 P = 11111 0 1111 0 0 111 0 0 0 11 0 0 0 0 1 />-' = 1-10 0 0 1-10 0 0 0 1-10 0 0 0 1-1 0 0 0 0 1 The matrix of/ relative to the standard basis is then 1 3 -1 -1 -2 0-4111 0-5 2 1 1 0-6122 0-5111 M = PAP'1 = -1 _ Now M -1 /5 has (column) rank 3 so the number of Jordan blocks associated with the eigenvalue 1 is 5-3 = 2. Also, the rank of the matrix A/-0/s is 4 so the number of Jordan blocks associated with the eigenvalue 0 is 5 -4 = 1. Hence the Jordan form is 1 1 0 1 0 1 0 To compute a Jordan basis we begin by choosing two independent vectors v,, v2 in the kernel of A/ - /5, say v, = [1 0 0 0 0] and v2 = [0 0 1 -1 0J. Next, we choose a vector vj in both the image and the kernel of M = M - 0/s, say v3 = [1 0 0 -1 1]; then a vector v4 independent of v3 and such that Mv4 = V3, say v4 = [0 - 1 - 1 - 5 2]; then a vector v5 independent of V3, v4 such that A/v5 = v4, say v5 = [1 - 3 - 3 -20 10]. The required Jordan basis is then {v,,...,v5}.

212 Further Linear Algebra 5.10 The characteristic polynomial of A is (X - 2)4 and the minimum polynomial is [X - 2)2. Since A - 2/4 has rank 1 the number of Jordan blocks associated with the (single) eigenvalue 2 is 4 -1 = 3.

Hence the Jordan form is J- 2 1 2 Now choose V| e Im (A -2/4)nKer(A -2/4), say V| = [-2 -2 -2 2], and then v2 independent of vi and such that (A -2/4)v2 = V|, say V2 = [1 0 0 0]. Finally, choose v3 and v4 such that {vi,..., v4} forms a basis, say v3 = [0 1 0 1] and v4 = [0 0 1 0]. Then we form the matrix P- -2 1 0 0 -2010 -2001 2 0 10 To solve the system x* = Ax we first solve the system y1 = /y, namely y, = 2y, + y2; y2 = 2y2; /3 = 2^ y4 = 2y4. The solution is clearly y* = c^e21; y3 = ^3«2'; yi = c2<?2'; yt = c2te1' + cxe2i'. Since now x - Py we deduce that xi = -2c2te2' -2c|tf2' + c2J*\ x2 - -2c2te1'-Icye21 + cjtf2'; xi = -Ic-ite1* -Ictf21 + C4*2'; x4 = Ic-ite* + 2^ + c3«2r. 5.11 The system is x' = Ax where Now A has Jordan form A- J- "1 0 0 3 7 9 -2] -4 -5 "1 0 0 1 1 0 01 0 1 and an invertible matrix P such that P lAP - J is ^3 0 1 P- 6 1 0 9 0 0 First solve y1 = /y to obtain yi = ae' + fcte'; y2 - be'; y3 = ce1. Then x* = Ax has the general solution x = py = fl[3 6 9]«' + fc([3 6 9]/«' + [0 1 0]*»') + c[1 0 0]<?'.

13. Solutions to the Exercises 213 5.12 The characteristic polynomial of A is (X + 2)(X - 2)2 which is also the minimum polynomial. The Jordan form is "2 1 0" / = 0 2 0 0 0 -2 A Jordan basis is {[12 4],[0 14],[1 -2 4]} and so an invertible matrix P such that P~lAP = / is P- 1 0 2 1 4 4 f -2 4 Now solve the system y* = /y to gel y, = c,e2' + c2r<r2r; - ,. -2i - ,. --21 >2 = C2<? ; >3 = Cye Then x = ^y gives x-xx- ci«2' + c2r<r2' + c3tf"2'. Now apply the initial conditions to get x - (At - 1)«2' + «~2r. 6.1 If Z C Ker/ then x + Z = y + Z gives x - y £ Z C Ker/ whence/(x) = /(y). Hence the assignment x + Zh /(x) defines a mapping #/ : V/Z -»IV. This mapping is linear since and tif[\(x + Z)] = tf,(Xx + Z) = /(Xx) = \f(x) = Xtf,(x + Z). Conversely, if the prescription x+Zm f(x) defines a (linear) mapping #/ : V/Z -»IV then x € Z gives x + Z = 0 + Z whence/(x) = /(0) = 0 and therefore x € Ker/. Hence ZC Ker/. 6.2 Taking Z = Ker / in the previous exercise we see that the assignment ti : x + Ker / »-» /(x) is a linear mapping from V/Ker/ to IV with Im# = Im/. Observe now that # is infective since #(x + Ker/) = #(y + Ker/) if and only if/(x) = /(y), i.e. if and only if x -y e Ker/, which is equivalent to x + Ker/ = y + Ker/. Hence # is an isomorphism from V/Ker/ onto Im/. 63 Ker/= {(0,y,z); y,z€ IR} (the y, z-plane). Now /f[(xIyIz) + Ker/l=/(xIyIz)+Ker/=(x,xIx) + Ker/=(xIyIz)+Ker/ since (x, x, x) - (x,y, z) € Ker/, and so/*" is the identity on IR3/Ker/. 6.4 /(fl,^c)= (0,*,*) so/(1,-1,3) = (0,1,-l)and/2(l,-1,3) = /(0,1,-1) =(0,0,1). Since then/3(l, -1,3) = (0,0,0) it follows that (a + bf + c/2)(l, -1,3) = (a, -a + fc, 3a - b + c) andsoZ(,_,3) = {(at-a + bt3a -b + c); fl,fc,c€iR}- 63 In the previous exercise the /-annihilator of (1, -1,3) is X3. 6.6/(x.y.z) = (x + z,y,-x-z). Then/(1,1,1) =

(2,1,-2),/2(1,1,1) = (0,1,0), /3(1,1,1) = (0,1,0) = /2(l, 1,1). Thusthe/-annihilatorof (1,1,1) is*3 -X2. 6.7 We have cA = mA - (X - a)3. As in Example 6.5, we have 3 = nt + • • • + nk with «| = 3. Thus k - 1 and the rational form is '0 0 fl3 T <V.)J= * 0 -3a2 0 1 3a 6.8 cA = (X + 1)2(X - 5) and m^ = (X + l)(X - 5). The rational canonical form is then diag{-l, -1,5} and coincides with the Jordan form.

6.9 cA - (X - l)(X - 2)2 and mA - {X - 1)(X - 2). The rational canonical form is then diag{ 1,2,2} and coincides with the Jordan form. 6.10 It is given that cf - {X - 1)3(X - 2)4 and mf - [X - l)2(X - 2)\ Then V = V, © V2 where dim V\ - 3 and dim V2 - 4. The induced mapping / on V\ has minimum polynomial (X - I)2. By Corollary 2 of Theorem 6.6 we have 3 = nx + • • • + nk with /i, = 2. So k = 2 and n2 = 1. Likewise the induced mapping /2 on V2 has minimum polynomial (X - 2)\ Then 4 = mi + • • • + mk with m\ = 3. So Ifc = 2 and /»2=1. Hence the rational canonical form is C{x-\y © C*-i © C(Jr_2)j © Cx-2 = 0 -1 1 2 0 0 1 0 0 1 8 -12 6 6.11 V has an /-cyclic vector a if and only if Za - V, which is the case if and only if dim Za - dim V, i.e. if and only if deg ma - deg cf. But deg ma ^ deg mf ^ deg C/ so if this holds then deg mf - deg cy, whence mf - cf. For the converse it suffices to consider the case where mf - p. As in the proof of Theorem 6.6, there is a non-zero vector x such that mx = p. Then if mf = c/ we have dim Zx - deg mx - deg C/ = dim V whence Zx = V and so x is an/-cyclic vector for V. 6.12 Follow the hint. 6.13 (1) cA = mA- Xi. Then as usual 3 = (rt| + form is then ~0 0 ■+n*)3withrt| = l,so)t= 1. The rational The classical form is the same as the Jordan form, namely 0 1 0 0 0 1 0 0 0 (2) cA - X\X - 5) and mA - X(X - 5). Forp, = X we have 2 = (n, + - • • + nk)\ with m - 1, so k - 2 and n2 - 1. For /¾ = X - 5 we have 1 = («i + • • • + nx)l with «| = 1, so k= \. The rational form is then Cx-5 = 0 0 0 0 0 0 0 0 5 which is the same as the classical form and the Jordan form. (3) cA = X2{X + 2)(X - 2) and mA = X(X + 2){X - 2). For p, = X we have 2 = (/11 + --- + /u)l with /i| = 1. Hence k - 2 and /i2 = 1. Then for /¾ = X - 2 we have 1 = (/i| + • - - + nk)\ with «| = 1. Hence k = 1. Similarly for/>3 = X + 2. The rational form is then Cx®Cx® Cx-2 © Cx+2 = diag (0,0,2, -2} which is the same as the classical form and the Jordan form.

6.14 V = V, © V2 where dim Vi = 6 and dim V2 = 4. The restriction/! of/ to Vt has minimum polynomial {X2 - X + 1)2. Then, as usual. 6 = («i + ■ ■ • + nk)2 with «, = 2 and so it = 2 and n2 = 1. The restriction /2 of/ to V2 has minimum polynomial (X + 1)2. Then, as usual. 4 = («i + • • • + nk)l with n, = 2 and so either k - 2 and n2 = 2, or it = 3 and «2 = «3 = 1. The rational form is then one of C{&-x+ip ©Ctf-x+i ©C(X+i)i ©C(X+,)2; C(x2-x+i)2 ©CX2_jm © Qx+iy © Cjr+i ©Cx+i. The classical form can be obtained from this as in Example 6.6. 6.15 In the complex case we have X2 -X + 1 = (X -a)(X -p) where a = j(l + iy/3) and J3 - \{l - iy/5). The rational form is the same as before, with C(Xi-x+i)» © Cx*-x+i replaced by <V*)i © Cx-a © C(x-4P © Cx-fi. In this case all the eigenvalues belong to the ground field C and so the classical form is the same as the Jordan form, which is either a 1 0 a a P 1 0 p P -1 1 0 -1 -1 1 0 -1 or the same matrix with the 1 in the (9,10) position replaced by 0. 7.1 Consider £?, : MatmxnF -> F defined by Efj[A) = a^ for each A = [fly]mXi». Clearly, £fj is linear and »./ #./<*«>-{' ifp=ii<? = j; otherwise. Hence{£fj ; i = 1,...,m and; = 1,...,«} is the dual basis. 7.2 That ipe is a linear form is routine. Suppose now that ip € Vd. Then relative to the dual basis of the previous exercise, we have V = ttvlEij)Elj. 7=1/=1 Let B = [by] € Malmxn be defined by by - ip{EitJ). Then 7=1i=l 7=li=l 7=1 7J The transition matrix from the given basis to the standard basis is 1— 10 0 1 1 -1 0 1 p = Hence the dual basis is /»-» = " 1 1 _M J ? J f ! ! L 2 2 2 {[jiji lli l~jili~jl, [5121 21}-

 7.4 {[2, -1,1,0], [7, -3,1, -1], 1-10,5,-2,1], [-8,3, -3,1]}. 7.5 The transition matrix from the basis B to the basis A is P- 1 2 3 4 /»-» = -2 3 2 1 4. Consequently, B4* = {-2pi + fp2, *fii - \h>i}- 7.6((4,5,-2,11),(3,4,-2,6),(2,3,-1,4),(0,0,0,1)}. 7.7 Proceed as in Example 7.9. If p - a0 + a\X + a2X2 then (\m + \2<p2 + \m)(P) =8 (*i + ^)°o + (i*i + >2)fli + (5^1 + 2X2)fl2- Then Xiy)| + X2P2 + *3¥>3 = 0 if and only if \i - \2 - h = 0, so that (pi, <p2t ^3} is linearly independent and so is a basis of (IR2[X])d. Since IR2[X] ~ IR3 under the correspondence ao + a\X+ a2X2 *-* (flo, <*\, o2) we have that R2[XY ~ (IR3)4. Now H>\{p) = flo+ |fli + 5^2 ¥>2(p) = fli + 2a2 ¥>3(p) = flo and so we can associate with {pi, v>2i <Pi) the basis ^- {H. i, il, 10.1,21, [1,0,0]} of (IR3)**. By considering the matrix whose inverse is we see that ((0,3,-1),(0,-1,2),(1,-3,1)} is the basis of IR3 = IR3 that is dual to B. Hence {3X - |X2, -\X + \X2,1 - 3X + \X2) is the basis of IR2[X] that is dual to the basis {<pi, <p2t <py}. 7.8 Properties (a) and (b) are immediate. If {a\,..., a,,) is the basis dual to (pi,..., <p,,) then the solution set C consists of those linear combinations tia\ + • • • + t,,a,, for which w('i«i+ --- + ^,,)^0. But p, = af and so p,(fiai + ■ ■ ■ + t,,a,,) = i<. 7.9/: V -» V is surjective [injective] if and only if there exists a linear mapping g : V -» V such that/og = idv [gof = idv]. The result is therefore an immediate consequence of Theorem 7.5(1)(3). p- r 2 '1 0 f j 1 0 .3 2 0. ro 2 -n Li -2 1 2 1

 7.10 That v is linear follows from Theorem 7.5. That \p is injective follows from the fact that f = tf gives, relative to some fixed ordered basis of V, Mat f = Mat g whence Mat / = Mat g and therefore/ = g. Since V is finite-dimensional, so is Lin(V, V). Hence tp is also surjective. 7.11 To find Y% we determine the dual of {[1,0,0], [1,1,0], [1,1,1]}. The transition matrix is i l r 0 1 1 0 0 1 , p-1- i 0 0 -1 1 0 0 -1 1 p- ThusK= {(1,-1,0),(0,1,-1),(0,0,1)}. The matrix of/ with respect to the standard basis is '2 1 1 0 1 1 0 -1 and the transition matrices relative to X, Y are respectively 1 0 -1 1 0 -1 By the change of basis theorem, the matrix of/ relative to X, Y is then 0] 0 1 • ri i n 0 1 1 0 0 1 ri o oi 1 1 0 1 1 1 "2 1 1 1 0 0 01 1 -1 "1 1 1] 0 1 1 0 0 1 = "2 3 31 3 5 6 3 5 5 The required matrix is then the transpose of this. 7.12 For every linear form # : \R„[X] -» R we have D'{ti) = tioD. Thus ti £ Ker D' if and only if ti o D - 0 which is the case if and only if ^(Hn-il*]) = 0. Hence a basis for Ker D' is {Dn}. 7.13 Observe that A C B implies B° C A° so for the first equality we need only show that P|V? C (£ V,)°. But if/ £ f|Vf then / annihilates each V, whence it annihilates £ V,. The second equality can be obtained from the first by applying the first equality to the family of dual spaces and identifying V with its bidual. 7.14 / - g £ W° if and only if (Vw £ W) f(w) - g{w), i.e. f]w = g\w- The linear mapping yd _» w* described by / •-» f\w therefore has kernel W°. Now apply the first isomorphism theorem. 7.15 5°° is a subspace of V with 5 C S°°. If now W - Span 5 then S C W gives S C 5°° C VT° = W. Hence S°°= W. 7.16 Every linear form/: R* -» R is given by f{x 1, X2t X3, X4) = fl1Xi + a2X2 + fl3*3 + «4*4 where ai = /(1,0,0,0), etc.. In order to have/e W° we require a{ + a2 = 0 = a3 + a4, so that / is given by /(X|,X2,X3,X4)= fliX| -fl|X2 +fl3*3~fl3*4. Now the dimension of W° is 4 - 2 = 2 and so we can obtain a basis for IV0 by first taking a, = l, a3 = 0 then fli = 0,03 = 1. A basis is therefore {/1,/2} where /i(*ii*2i*3i*4)= *i -*2, /2(-^1, *2,*3,*4) = *3-~*4. 7.17 7ri : Arf -»V* is given by 7^(/) = /irA, and similarly for 1¾. For every jfc £ Vd we have * = ifcidv = k(TtA + irB) = k\AirA + *|B7rs = ^(A:^) + ^(jfc|fl) € Im^ + Im^.

 ThusV'=Imiri + Im7r^ If now jfc £ Im k'a n Im Kg then there exist /, g with jfc = /ka and k = girB. Consequently, 0 = /0 = /kakb -

kitB - gKBKB = g-nB = k. Hence Im k*a n Im 71^ = {0} and therefore Vd = Im n^ © Im Kg. 7.18 Follow the hints. 7.19 That/^ is linear is routine. As for (a), we have /c.vLUO] = f*A(t I z>>) = ((/1 z)y I y)x = (/1 z) (y | y)x = |y| 2<f | z)x = |y| 2/,,(/). As for (b), we have ^.,(z)|/>=((z|y)x|/>=(z|y)(xU)=(z|^y)=(z|(r|x)y) = (z|/^(/)^ 7.20 (a) From the previous exercise we have that/,,, is normal if and only if it commutes with /ViJt, which is the case if and only if, for all z£V, Hv||2(z|x>x= M2(z\y)y- Taking z = x in this we obtain IMI2IMI2* = \42(*\y)y which gives x = Xy where X = (x | y)/ ||y 2 € C so that the condition is necessary. Conversely, if x = \y for some \ £ C then l|y||2(z|*>*= lly||2(z|y>>^=l>12lly|l2(z|y>y= IMI2(*|y>y so that the condition is also sufficient {b) From the previous exercise we have that/,^ is self-adjoint if and only if, for all z £ V, (z|y)x=(z|x)y. Taking z = y we obtain x = Xy where \ = (y | x)/ ||y||2. This gives (x | y) = X ||y||2 = (y | x), whence we see that \ £ IR. The converse is clear. 7.21 It is readily seen that (K(f) \g)and(f\ K(g)) are each equal to f yf(y)dy f xg(x)dx. Jo Jo 7.22 We have ((atb,c)\f(x,y,z)) = (fa, Ml (*,>,*)> = ((a + btb,a + b + c)\(xiyiz)) = x(fl + fc) + yb + z(fl + fc + c) = fl(x + z) + fc(x + y + z) + cz = ((fl,M|(x + z,x + y + z,z)). Hence/*(x, y, z) = (x + z, x + y + z, z). 7.23 For the 0 given in the hint, we have (a* 10) = /(a*) for each basis element ak. /9=(1,-2,4). 7.24 Follow the hint. 7.25 We have 0^)10= f Pq*= f qfr= (q\M) Jo Jo and so £=/p. For the rest of the question, follow the hint. 7.26 This is immediate from Theorem 7.11(1)(2)(4).

13. Solutions to the Exercises 219 7.27 If W is/-invariant then xeW1 =► {VweW) 0={f(w)\x) = {w\r(x)) =>f{x)£WL so W1 is/^invariant. If W1 is/^invariant then weW => (Vy£WL) 0={w\f(y)) = (f{w)\y) =>/(h>) € IV11 = »V so IV is /-invariant. 7.28 A* = A' = A and therefore det A* = det A = det A. 8.1 An orthonormal basis for W is { £(0, 1,2,0), £(1,0,0,1)}. The Fourier coefficients of (1,2,0,0) are £ and -7- and so /^(1,2,0,0)=2(0,1,2,0)+1(1,0,0,1)=(1,2/5,1). 8.2 We have V = IV© W1. Let {e\,..., ek) be an orthonormal basis for W and extend this to an orthonormal basis {¢1,...,^,^+1,...,^} of V. Then {¢4+1,...,«,,} is an orthonormal basis of W1. For every jt^Vwe then have * n x = E(* I ei)x + E (* I *')*=ivW+p»*(x) i=\ 1=4+1 whence/va = '^v ~Pw- 8J Suppose that IV is of dimension n. Then, for every x £ V, IM*)H2 = (E<* I «)« 11(* I <>)«) = E K* I «i>la < M2 8.4 The matrix representing a rotation through an angle 1? is *# = cosi? —sin i? sin 1? cos 1? and the matrix that represents a projection onto the x-axis is 1 0" A = 0 0 The required matrix representing pL is then R4AR-4. 8.5 See Example

4.5. The eigenspaccs are £_, = Span {(1,-2, l)}and£, = Span {(-1,0,1)}. The ortho-projection on E-\ is given by <(*,y,z)| £(1,-2,1))^(1,-2,1)= l(x + 2y-Z) (l,-2,l) and similarly that on E\ is l(-x + z)(-l, 0,1). 8.6 A*A - AA*\ take inverses, noting that (A*)-1 = (A-1)*. 8.7 Under the given conditions, we have (A + iB)(A + iB)* = A2 + iBA - iAB + B2 {A + iB)*(A + iB) = A2 - iBA + iAB + fl2. Hence A + i£ is normal if and only if A, B commute. 8.8 Observe that A* = A and B* = -/?. Moreover, A and Z? do not commute. 8.9 f\ = i(T+r) = £(/*+/) so/i is self-adjoint. Also,/} = -£(T-.T) = -£(/*-/) = /2 so/2 is also self-adjoint.

220 Further Linear Algebra It is clear that/ = /, + if2. If also/ = £, + ig2 where gu g2 are self-adjoint then we have /i ~gi = «(ft -/2). Consequently,/* -g\ = -½ -/2)*, i.e./, -gl = -½ -/2). It follows that/, -£, = 0 = g2-f2. /* = /f -1/2 = /1 - «/2- Hence/ is normal if and only if/, + i/2 commutes with/, -1/2. This is the case if and only if/, commutes with/2. 8.10 Relative to the standard basis the matrix A of/ is A = "2 1 f 2 ; A*= " 2 —1 1 2 Clearly, A is not self-adjoinL Since AA* = 5 2 + 2i 2-2i 5 = A*A we see that A is normal. 8.11 This follows by a simple application of Theorem 7.11. 8.12 Either verify directly that AA* = A*A or observe that A-iB-¥B* where "1 0 B = 1 -2i 2 0 i 0 3 and use Exercise 8.11. 8.13 Suppose that A is normal and that B = g{A). There is a unitary matrix P and a diagonal matrix D such that P~lAP = D. Then B=g{A)=Pg(D)p-\ Consequently,Wb = Pg(D)g(D)P~xjnd similarly BB' = /,g(D)g(D)/,~l. Since g{D) and g(D) are diagonal matrices it follows that B B— BB and so £ is normal. 8.14 If A* = p{A) then clearly AA* - A*A, so that A is normal. Conversely, suppose that A is normal. Then there is a unitary matrix P and a diagonal matrix D such that A = PDP~X. Let Ai,..., A, be the distinct diagonal elements of D. Then there exist unique at,...,ar such that \i = fl0 + fli^i + fl2^? + • ■ ■ + flr-iM-' A2 = ^0 + 1|^2 + «2^2 + ■ • • + flr-l^r' Ar = flo + fllV + 02\2r + ■ ■ ■ + flr-,)^-1. [In fact, since A,,..., Ar are distinct, the (Vandermonde) coefficient matrix is non-singular and so the system has a unique solution.] We then have D = a0l + a,D + a2& + • • ■ + ar-Jf'x and consequently A*= /'D/'-1 = flo/ + a\A + • ■ • + a^xA'"1. 8.15 Consider matrices relative to the standard ordered basis. The matrices in (1), (2), (4) are self-adjoint, hence normal. That in (3) is not normal. 8.16 Taking adjoints in A*A = -A we obtain -A* = A*A** - A*A whence A* = A and so A is self-adjoint Let \,,..., \r be the distinct non-zero eigenvalues of A. From the above we have A2 = -A, from which it follows that the distinct non-zero eigenvalues of -A, namely -\ x,..., -\r are

13. Solutions to the Exercises 221 precisely the distinct non-zero eigenvalues of A2, namely \2,..., \2. Consequently, \ \,..., \r are all negative. Let ctj = -X,- > 0 for each i, and suppose that Then this chain must coincide with the chain <a, a]<a\<--<a2r. Consequently, a,- = a2 for every i and so each a, = 1. Since by hypothesis a/ / 0 we deduce that \i = -tti = -1. 8.17 Suppose that A is unitary, so that A* = A~l. Then A*A = / = AA* and so A is normal. Also (regarding A as a linear mapping) we have (Ax|Ax)=(x,A*Ax) = (x,x) whence \\Ax\\ = ||x||. If now \ is any eigenvalue of A we have ||Ax| = N |*|| • Thus we have|X|=l. Conversely, suppose that A is normal and that every eigenvalue of A is of modulus 1. Let /: V -> V be an oitho-diagonalisable mapping that is represented by A (cf. Theorem 8.8). If the * * spectral resolution of/ is/ = £ \,pt then we know that the spectral resolution of/* is £ A,p,- i=i #=i and that i2 / °r=EMVi=r°/ f=i If each eigenvalue of/ is of modulus 1 it follows that /o/*=/* °/=I> = iclv. i=i Hence/* = /'" and consequently A is unitary. 8.18 The eigenvalues of A are -1,1,8. Corresponding general eigenvectors are (*^0). 0 x x x , 0 , 0 0 x -x An orthogonal matrix P such that P~'AP is diagonal is then P = •n 1 1 ^ 0 75 75 1 0 0 0 75 v^J />'" = 0 1 * ° Lv3 0 - i Then P-IAP = diag{-l, 1,8}. The matrix B = Pdiag{-1,1,2}P~' is then such that B3 = A. A simple calculation gives fl = 1| 0 ~\ 0-10 8.19 The matrix is symmetric and its eigenvalues are 1 (of algebraic multiplicity 2) and 4. Since these are strictly postive, A is positive definite. 8.20 By definition,/ is self-adjoint and (f(x) \ x) > 0 for every x e V; and similarly for g. Then / + g is self-adjoint and ((/" + g) (x) | x) = (f(x) | x) + (g(x) \ x) > 0. Hence/ + g is positive definite.

222 Further Linear Algebra 8.21 If fg is positive definite then it is self-adjoint. Since/, g are self-adjoint it follows that Conversely, suppose that/, g are positive definite and commute. Then (gf(x)\x)=(fg{x)\x) where the left hand side is (f(x) | g(x)) and the right hand side is (g(x) \f(x)). It follows that (gf{x) | x) £ IR and so, by Theorem 8.12, gf is self-adjoinL Now by Theorem 8.15 we have g = h2 where h is self-adjoint and, by Theorem 8.14, /i = q(g) for some polynomial q. Then/ commutes with h. Consequently, for every non-zero x £ V, we have {8f(x) | x) = (h2f(x) | x) = (hf(x) | h(x)) = (f/.(x) I *W) > 0. Hence gf is positive definite. 8.22 Let ((-1 -)) be an inner product on V. Then, by Exercise 7.23, for every 0 £ V there exists a unique 0' £ V such that ((-10)) = (-101). Define/: V -» V by setting f(0) = 0'. Then we have (a |/(0)) = ((a\0)) = ((0|a» = (0|/(a)> = (f(a)|/?>. Thus/is self-adjoint. Moreover, (f(a)|a)=(a|/(a))=((a|a))>0 and so/is positive definite. As to the uniqueness of/, if g is also such that (Va,0)

(M/?))=(s(a)|0) then clearly g — f. 9.1 (1) is bilinear; (2) is not. 9.2 The required matrices are (1) 0 2-5; (2) -3 -5 0 93 -xiyi + -Jtiy2 - I lxiy3 - 5x2yi + 2x2y2 -1 Ix2y3 + 5x3yi + x3y2 - *3y3- 94 If A is symmetric and if there is an invertible matrix P such that B = P'AP then B' = P'A'P - P'AP = B. '•5 /(*, y) = pU, y) + ?(x, y) where p(x,y) = \\f(xty) +/(y, x)] is symmetric and q{x, y) = 2 LA*i >) -/(yi *)] is skew-symmetric. If also/ = a + b where a is symmetric and b is skew- symmetric then a—p—q-b where the left hand side is symmetric and the right hand side is skew-symmetric, whence each is 0 and then a — p,b —q. 0) P- j[2xiyi + 3xiy2 + 3x2y, + 2x2y2]and? = j[-xiy2 + x2yi]. (2) p = /and?=0. 9.6 The set of quadratic forms is closed under addition and multiplication by scalars. 9.7 G(*i,..,*n)=X>?. 9.8 (1) The form is xMx where A = 3 "I -| -7 (2) The form is xMx where A = 3 -2 I _• 5 2 -2 L 2 2 -3

 9.9 Since \ is an eigenvalue of A there exist au...,annoiall zero such that Ax= \\ where x = [fl! ••• a,,]'. Then Q=x'Ax=\x'x=\'Ea2. i-l 9.10 We have Q(x,yt z) = x2 + 2xy + y2 -2xz -z2 = (x + y)2 + x2 - (x + z)2 so the required matrix is 9.11 (1) We have [1 0 0 1 0 0 0] 0 -1 2f -z2 + xy + xz = 2(y + Jx)2 - \x2 + xz-z2 = 2(y+Jx)2-|(x-4z)2 + z2 and so the rank is 3 and the signature is 1. (2) Putx = X + r,y = X->\z = Ztoobtain 2xy-xz-yz = 2(^-^)-(^+^-(^-1^ = 2X2-2y2-2XZ = 2(X -IZf-^-lY1. Thus the rank is 3 and the signature is -1. (3) Putx = X + Y,y = X - Y, z = Z,t - Ftoobtain yz + xz + xy + x/ + yf+ zf = X2-Y2+2XZ + 2XT + ZT = (X + Z + T)2-Y2-Z2-T2-ZT = (X + Z + T)2 -(T+ ±Z)2 -%Z2 -Y2. Thus the rank is 4 and the signature is -2. 9.12 (1) The matrix in question is A = 1 -1 2 -1 2 -3 2-3 9 Now x2 + 2y2 + 9z2 -2xy + 4xz-6yz = (x -y + 2z)2 + y2 + 5z2 -2yz = (x-y + 2z)2 + (y-z)2+4z2 = (2 + t,2 + C2 where £ = x - y + 2z, tj = y - z, (= 2z. Then X y z = /» \(] n LcJ = 1 I - = 0 I 0 0 I ! 2 J n LCJ and/M/^diagO.l.l}.

 A = (2) Here the matrix is " 1 1 0-1 14 3-4 0 3 1-7 -1-4-7-4 The quadratic form is Q{*.>,z,0 = (* + y ~*? + 3>* + z2 -5t2 + 6yz -6yt - 14zr = {x + y-t)i + 3(y + z-t)2-2z2-St2-Szt = (x + y-02 + 3(y + z-02-2(z + 2f)2 where £=x + y-M = V3(y + z - f). C = V5(z + 2f)• Writing r = t, we then have x' y z t = p \t] n r = 1 -A A -2^ 0 A -A 3 0 0^-2 .0 0 0 u and P'AP = di*g{ 1,1 ,-1,0}. 9.13 The matrix is that of Exercise 9.10. Since the rank is 3 and the signature is 1 the quadratic form is not positive definite. 9.14 (1) x2 + 3xy + y2 = (x + %y)2 - ly2 so the rank is 2 and the signature is 0. Hence the form is not positive definite. (2) 2x2 - Axy + 3y* -z2 = 2{x - y)2 + y2 -z2 so the rank is 3 and the signature is 1. Hence the form is not positive definite. 9.15 Follow the hint. 9.16 We have £ {XrS + r + s)xrX, = \ £ («rM«f) + E ("frlXj + E M«f) r,r-| r,*=l r.j=l r,i=l = x(E^)2 + 2(e^)(e«,) Now let yi = Erx'. v2=X>r, y3 = x3, ... yn-xn. i=i r=i Then the form is \y\ + 2yiy2, which can be written as f Hyi + b2)2~{A >^^0; 5(y.+y2)2-J(yi-y2)2 if A = 0. Hence in either case the rank is 2 and the signature is 0. 10.1 Let/* be the adjoint of/. Then since we are dealing with a real inner product space we have, for all x e V, (f(x)*)=(x\r(x))= {r{x)\x). It follows that (rtx)|x)=(i(/+/*)(x)|x)

 where \{f + /*) is self-adjoint. If now (f(x) | x) = (g{x) | x) where g is self-adjoint then we have ((/" - g){x) \ x) = 0 whence, by Theorem 10.3, f-g\s skew-symmetric. Consequently, since g is self-adjoint, and from this we obtain g = {(f+f). 10.2*(-A) = s(A')=k(A)]' = 0' = 0. Let the minimum polynomial of A be mA - a0 + a,X + a2*2 + • • ■ + aH-iX"~l + X". Since, by the first part of the question, mA{-A) = 0 we have that a0-aiX + a2X2 - ■ ■ + (-l)nXn is also the minimum polynomial of A. Hence fl| = a3 = •• • = 0. 103 Compare with Example 10.2. The minimum polynomial of A is X(X2 + 9). 10.4 By Exercise 10.2, both A and A' = —A have the same minimum polynomial. Then by the Corollary to Theorem 10.S, and Theorem 10.8, both A and A' have the same canonical form under orthogonal similarity. 11.1 The MAPLE code is > m:=Matrix(4,4,(i,j)->4*(i-l)+j) ; 11.2 The MAPLE code is > a:=Matrix([[1,2,3,4,5,6], [9,8,7,7,8,9],[1,3,5,7,2,4], [5,4,5,6,5,6], [2,9,2,7,2,9], [3,6,4,5,6,7]]); followed by > ia:=aA(-l); 113 The MAPLE code is > a:=Matrix([[x,2,0,3],[1,2,3,3],[1,0,1,1],[1,1,1,3]]); followed by > solve(Determinant(a)=0,x); The answer is that the matrix is invertible except when x = -\. 11.4 Input A and B then: > C:=AA3-B"2+A.B.(BA2-A) ; followed by > CharacteristicPolynomial(C,X); 11.5 The MAPLE code involves a loop:

 > big for a for b for c for i for j for k for x for y for z =-100 from 1 from 1 from 1 from 1 from 1 from 1 from 1 from 1 from 1 to 2 do to 2 do to 2 do to 2 do to 2 do to 2 do to 2 do to 2 do to 2 do m:=Matrix(3,3,[[a,b,c],[i,j,k],[x,y,z]]); if Determinant(m)>big then big:=Determinant(m) fi; od;od;od;od;od;od;od;od;od: print(big); The output is S. For the second part, the code is > for a from for b from 1 for c from 1 for i from 1 for j from 1 for k from 1 for x from 1 for y from 1 for z from 1 1 to to 2 to 2 to 2 to 2 to 2 to 2 to 2 to 2 m:=Matrix(3,3,[[a if Determinant(m): od;od;od;od;od;od The answer is 2 do do do do do do do do do b,c], [i.j =big

then •od;od ;od: ,k]([x,y,z]]); print(m) fi; ("1 2 2" 2 1 2 2 2 1 i [2 1 2] 2 2 1 1 2 2 > [2 2 1] 1 2 2 2 1 2 11.6 The MAPLE codes are the following: > f: = (i,j)->i"2+:T2; > for i from 1 to 10 do A:=Matrix(i,i,f); Determinant(A); od; > g:=proc(i,j); if i<j then i*2+j else j*2 fi; end:

 > for i from 1 to 10 do A:=Matrix(i,i,g): B:=Matrix(i,i,f); C:=A-B; print(Determinant(A), Determinant(C), Determinant(C)/Determinant(A)); od: > la: = [seq(Determinant(Matrix(i,i,g)), i=l..20)]; > lb:=[seq(Determinant(Matrixd,i,g)-Matrix(i,i,f)),i=l..20)]; > lc:= [seq(i"2+lb[i]/la[i],i=1..20)]; 11.7 The conjecture is that if A = [ajjnxn is such that a,j is a polynomial in i and) then, except for some small values of n, the determinant of A is 0.

