CHAPTER 0

PRELIMINARIES

This chapter gives all necessary prerequisites for the following chapters.

Throughout this thesis we shall denote the set of all natural numbers, real numbers and complex numbers by $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$; respectively.

A series of the form :

$$\sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + \dots$$

is called a <u>power series</u> in x, and a series of the form :

$$\sum_{n=0}^{\infty} c_n (x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + \dots$$

is called a power series in x-a. In general, a power series in n variables is given in the form :

$$\sum_{m_1,m_2,\dots,m_n=0}^{\infty} c_{m_1 m_2 \dots m_n} (x_1 - a_1)^{m_1} (x_2 - a_2)^{m_2} \dots (x_n - a_n)^{m_n} .$$

In this thesis we shall use $\sum_{0}^{\infty} c_{m_1 m_2 \dots m_n} (x_1 - a_1)^{m_1} (x_2 - a_2)^{m_2} \dots (x_n - a_n)^{m_n}$ instead of $\sum_{m_1,m_2,\dots,m_n=0}^{\infty} c_{m_1 m_2 \dots m_n} (x_1 - a_1)^{m_1} (x_2 - a_2)^{m_2} \dots (x_n - a_n)^{m_n}$.

The operations of series are defined as in [1, p.3].

A function of n variables, $f(x_1,\dots,x_n)$, is said to be an <u>analytic function at a point</u> $(a_1,a_2,\dots,a_n)$ if it can be expanded into a power series which converges to the function in a neighborhood of $(a_1,a_2,\dots,a_n)$; i.e.,

$$f(x_1,x_2,\ldots,x_n) = \sum_0^\infty c_{m_1m_2\ldots m_n}(x_1-a_1)^{m_1}(x_2-a_2)^{m_2}\ldots(x_n-a_n)^{m_n}$$

in some neighborhood of $(a_1,a_2,\ldots,a_n)$. If f is analytic at $(0,0,\ldots.0)$, then we can write f in the form :

$$f(x_1,x_2,\ldots,x_n) = \sum_0^\infty c_{m_1m_2\ldots m_n}x_1^{m_1}x_2^{m_2}\ldots x_n^{m_n}$$

in some neighborhood of $(0,0,\ldots,0)$. If f is analytic at every point in a domain D, then it is said to be <u>analytic</u> on D. If a function $f(x_1,\ldots,x_n)$ is equal to the power series $\sum_0^\infty c_{m_1m_2\ldots m_n}(x_1-a_1)^{m_1}(x_2-a_2)^{m_2}\ldots(x_n-a_n)^{m_n}$ in a neighborhood of $(x_1,x_2,\ldots,x_n) = (a_1,a_2,\ldots,a_n)$ and if $f(x_1,\ldots,x_n)$ is equal to the power series $\sum_0^\infty b_{m_1m_2\ldots m_n}(x_1-a_1)^{m_1}(x_2-a_2)^{m_2}\ldots(x_n-a_n)^{m_n}$ in a neighborhood of $(x_1,x_2,\ldots,x_n) = (a_1,a_2,\ldots,a_n)$, then these two power series are identical, coefficient by coefficient; i.e., $c_{m_1m_2\ldots m_n} = b_{m_1m_2\ldots m_n}$ for all $m_1,\ldots,m_n \in \mathbb{N}\cup\{0\}$, see [1, theorem 3]. If f is analytic at a, then it is continuous at a [1, theorem 1].

Let $(S,*)$ be a semigroup. If there exists a point 0 belonging to S such that for all x in S, $x * 0 = 0 * x = 0$, then we call the point 0 a <u>zero</u> of S and we call the semigroup S to which 0 belongs a <u>semigroup S with zero</u>. Moreover, if there is a point 1 belonging to S such that for all x in S, $x * 1 = 1 * x = x$, then we call 1 an <u>identity</u> of S. If a zero and an identity exist, then they are unique. We write a semigroup with zero 0 and identity 1 as $(S,0,1)$.

A mapping Φ from a semigroup S into a semigroup S' is said to be a <u>homomorphism</u> if for all a, b $\in$ S, $\Phi(ab) = \Phi(a)\Phi(b)$. It is said to be an <u>isomorphism</u> if Φ is one-to-one and onto. S and S' are said to be <u>isomorphic</u> if there is an isomorphism of S onto S'. In this case we write $S \cong S'$.

Let F be a field. If $f(x) = a_0 + a_1x + \ldots + a_nx^n \neq 0$ is a polynomial in $F[x]$ and $a_n \neq 0$, then the degree of $f(x)$ is n. If $a_n = 1$, then f is said to be monic. A polynomial $p(x)$ in $F[x]$ is said to be irreducible over F if whenever $p(x) = a(x)b(x)$ with $a(x), b(x) \in F[x]$, then one of $a(x)$ or $b(x)$ has degree zero (i.e., is a constant). If $g(x) = p(x)^n$ where $p(x)$ is an irreducible polynomial, then the order of $p(x)$ is n. A scalar $c \in F$ such that $p(c)=0$, is called a zero (or root) of the polynomial $p(x) \in F[x]$. Let $p(x) = a_mx^m + a_{m-1}x^{m-1} + \ldots + a_1x + a_0$ be a polynomial of degree m of $\mathbb{C}[x]$. Then $p(x)$ has the unique factorization (unique up to the order of the factors)

$$p(x) = a_m(x - c_1)(x - c_2)\ldots(x - c_m),$$

where $c_1, c_2, \ldots, c_m$ are zeros of $p(x)$ [4, theorem 6.2 ,p.267]. That is, a polynomial with coefficients which are complex numbers has all its roots in the complex field. If $p(x) = a_mx^m + a_{m-1}x^{m-1} + \ldots + a_1x + a_0$ is a polynomial of $\mathbb{R}[x]$ of degree m, then, over $\mathbb{R}$, $p(x)$ may be factored uniquely (up to the order of factors) into irreducible factors :

$$p(x) = a_m(x^2 + r_1x + s_1)\ldots(x^2 + r_kx + s_k)(x-c_{2k+1})\ldots(x-c_m),$$

where $x^2 + r_ix + s_i \in \mathbb{R}[x]$ are irreducible over $\mathbb{R}$ and $c_{2k+1}, \ldots, c_m$ are real numbers [4,theorem 6.3, p.269]. This states that the only irreducible, non-constant, polynomials over the field of real numbers are either of degree 1 or of degree 2.

Let V and W be vector spaces over the field F. A linear transformation from V into W is a function T from V into W such that

$$T(c\alpha + d\beta) = c(T(\alpha)) + d(T(\beta)),$$

for all α and β in V and all scalars c, d in F.

The vectors $v_1, v_2, \ldots, v_m$ in V are said to be linearly dependent if there exist scalars $a_1, \ldots, a_m$, not all zero, such that $a_1 v_1 + a_2 v_2 + \ldots + a_m v_m = 0$.

If $A = (a_{ij})_{i,j=1,2,\ldots,n}$ is a matrix of order n with elements in a field F, then the determinant of A, written det A, is the element $\sum_{\sigma \in S_n} (-1)^{\sigma} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ in F where S_n is the symmetric group of degree n and the symbol $(-1)^{\sigma}$ will mean +1 if σ is an even permutation and -1 if σ is an odd permutation.

We shall use the notation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix}$$

for the matrix $(a_{ij})_{i,j=1,2,\ldots,n}$ and we shall denote the space of all $n \times n$ matrices with elements in a field F by M(n,F). A matrix A is called similar to the matrix B if there exists a nonsingular matrix X such that $A = X^{-1} B X$. Let A be a square matrix with elements a_{ij}; $i,j = 1,2,\ldots,n$. Then the matrix

$$xI - A = \begin{bmatrix} x-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x-a_{nn} \end{bmatrix},$$

where x is an indeterminate and I is the identity matrix, is called the characteristic matrix of A. Its determinant $\varphi(x) = \det(xI-A)$ is evidently a polynomial of degree n in x; it is called the characteristic polynomial of A. The equation $\det(xI-A) = 0$ is called the characteristic equation of A, and the n roots of this equation are called the characteristic roots (or eigenvalues) of A. If A is a block matrix of the form.

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_s \end{bmatrix},$$

where $A_1,\ldots,A_s$ are square matrices, then its characteristic matrix, as is easy to see, has the form :

$$xI - A = \begin{bmatrix} xI_1-A_1 & 0 & \cdots & 0 \\ 0 & xI_2-A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & xI_s-A_s \end{bmatrix}$$

where $I_1,\ldots,I_s$ are identity matrices of suitable orders. From the theory of determinants it is known that the determinant of A is the product of the determinants of its diagonal blocks. Consequently,

$$\det(xI-A) = \det(xI_1-A_1)\det(xI_2-A_2)\ldots\det(xI_s-A_s).$$

If $A_1,A_2,\ldots,A_s$ are square submatrices and

$$A = \begin{bmatrix} A_1 & & & * \\ 0 & A_2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & A_s \end{bmatrix},$$

then $\det A = (\det A_1)(\det A_2)\ldots(\det A_s)$ (*'s indicate parts in which we are not interested in the explicit entries) [3, p.293].

If A and B are similar matrices, then they have the same characteristic polynomials and hence the same eigenvalues [4,theorem 1.2,p.279].

The matrix

$$\begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & \lambda \end{bmatrix},$$

with λ's on the diagonal, 1's on the super diagonal, and 0's elsewhere is called a <u>basic</u> <u>Jordan</u> <u>block</u> <u>belonging</u> to λ. If $A \in M(n,F)$ has all its characteristic root, $\lambda_1,\ldots,\lambda_k$ in F, then an invertible matrix $C \in M(n,F)$ can be found so that CAC^{-1} is of the form:

(0.1) $$\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix}$$

where each

$$J_i = \begin{bmatrix} B_{i1} & 0 & \cdots & 0 \\ 0 & B_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{ir_i} \end{bmatrix}$$

is an $n_i \times n_i$ matrix (n_i is the multiplicity of λ_i), and where $B_{i1},\ldots,B_{ir_i}$ are Jordan blocks belonging to λ_i [3, p.259]. The matrix in (0.1) is unique except for the order of the submatrices J_i down the diagonal [4,p.328].

Let $V_n(\mathbb{R})$ denote an n dimensional vector space over $\mathbb{R}$ and T a linear transformation of $V_n(\mathbb{R})$ into itself. A subspace V of $V_n(\mathbb{R})$ is called an <u>invariant subspace</u> of $V_n(\mathbb{R})$ for T if $T(V) \subseteq V$; that is, for each vector $v \in V$, $T(v) \in V$.

Let v be an arbitrary nonzero vector. Since there are atmost n+1 vectors in the set $\{v, T(v), T^2(v),\ldots, T^n(v)\}$, the vectors are linearly dependent. Let $T^k(v)$ be the first vector dependent on the preceeding vectors so that

(0.2) $$T^k(v) = a_0 v + a_1 T(v) + \ldots + a_{k-1} T^{k-1}(v) .$$

The linear subspace generated by $v,T(v),\ldots,T^{k-1}(v)$, written $L\{T^k(v)\}$, is an invariant subspace for T.

By simply transposing the right-hand side of (0.2) to the left we obtain

$$(T^k - a_{k-1}T^{k-1} - \ldots - a_1 T - a_0 I)(v) = 0.$$

Let $m(x) = x^k - a_{k-1}x^{k-1} - \ldots - a_1 x - a_0$. Then m(x) is a monic polynomial of $\mathbb{R}[X]$ of minimum degree such that $[m(T)](v) = 0$ [4,p.301]. If p(x) is a monic polynomial of degree k such that $[p(T)](v) = 0$, then $p(x) = m(x)$. Hence for each nonzero vector $v \in V_n(\mathbb{R})$ and each linear transformation $T : V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$, there exists a unique monic polynomial m(x) of minimum degree such that $[m(T)](v) = 0$ [4, theorem 1.1, p.302]. Such a monic polynomial m(x) is called the <u>relative minimal polynomial</u> of v with respect to T.

Remark: The degree of the relative minimal polynomial $m(x) \geq 1$. Hence $m(x)$ is not a constant polynomial.

If T is a linear transformation of $V_n(\mathbb{R})$ into itself, then there are nonzero vectors $v_1, \ldots, v_\ell$ in $V_n(\mathbb{R})$ such that

(0.3) $$V_n(\mathbb{R}) = L\{T^{i_1}(v_1)\} \oplus \ldots \oplus L\{T^{i_\ell}(v_\ell)\}$$

where $L\{T^{i_j}(v_j)\}$ is an invariant subspace of maximal dimension in

$$L\{T^{i_j}(v_j)\} \oplus L\{T^{i_{j+1}}(v_{j+1})\} \oplus \ldots \oplus L\{T^{i_\ell}(v_\ell)\}$$

for $j = 1, 2, \ldots, \ell$ [4, cor.to theorem 3.2, p.312]. The relative minimal polynomials, $m_i(x)$ of the vector v_i occuring in the direct sum decomposition of $V_n(\mathbb{R})$ in (0.3) are called the invariant factors of the linear transformation T. Notice that all definitions in this paragraph can be applied to an $n \times n$ real matrix A by using a matrix A in place of T and letting $V_n(\mathbb{R}) = \mathbb{R}^n$. It can be shown that the invariant factors of an $n \times n$ matrix A are independent of the particular vectors selected in a decomposition of $V_n(\mathbb{R})$ and are the same for any matrix $B = PAP^{-1}$ where P is a nonsingular $n \times n$ matrix [4, cor.to theorem 3.3, p.315]. Two $n \times n$ matrices A and B are similar if and only if they have the same invariant factors [4, cor. to theorem 4.1, p.320].

For a monic polynomial $p(x) \in \mathbb{R}[x]$, $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$, the companion matrix $C(p(x))$ is the $n \times n$ matrix

$$C(p(x)) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix},$$

and if $p(x) = x + a$, then $C(p(x)) = [-a]$. We shall use the notation $D(p(x)^b)$ for the matrix:

$$\begin{bmatrix} C(p(x)) & P(n) & 0 & \cdots & 0 \\ 0 & C(p(x)) & P(n) & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ & & & & P(n) \\ 0 & 0 & 0 & \cdots & C(p(x)) \end{bmatrix}$$

where $C(p(x))$ is the companion matrix of the irreducible polynomial $p(x)$ and is repeated b times along the diagonal, and where $P(n)$ is the $n \times n$ matrix ($n = \deg p(x)$):

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

having a single 1 in the lower left-hand place with all other elements zero. If $n = 1$, $P(1) = [1]$.

For example, let $p_1(x) = x^2 + bx + c$, $p_2(x) = x-a$ be polynomials in $\mathbb{R}[x]$. Then

$$C(p_1(x)) = \begin{bmatrix} 0 & 1 \\ -c & -b \end{bmatrix}, \quad P(2) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

If $p_1(x)$ is irreducible over $\mathbb{R}$, then

$$D(p_1(x)^3) = \left[\begin{array}{cc|cc|cc} 0 & 1 & 0 & 0 & 0 & 0 \\ -c & -b & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -c & -b & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -c & -b \end{array}\right].$$

We also have that

$$D(p_2(x)^3) = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}.$$

Let $m_1(x),\ldots, m_t(x)$ be the invariant factors of an $n \times n$ matrix A and suppose that the invariant factors $m_1(x),\ldots,m_t(x)$ are factored in $\mathbb{R}[x]$ into irreducible polynomials,

$$m_j(x) = p_{j1}(x)^{\ell_{j1}} p_{j2}(x)^{\ell_{j2}} \cdots p_{jk_j}(x)^{\ell_{jk_j}}, \tag{0.4}$$

for all $j = 1,2,\ldots,t$. Then there exists an invertible matrix C such that

$$CAC^{-1} = \begin{bmatrix} D(p_{11}(x)^{\ell_{11}}) & 0 & \cdots & 0 \\ 0 & D(p_{12}(x)^{\ell_{12}}) & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D(p_{tk_t}(x)^{\ell_{tk_t}}) \end{bmatrix}, \tag{0.5}$$

see [4, p.328]. A matrix A with invariant factors $m_i(x)$ of (0.4) is similar to a matrix of the form (0.5) and this form is unique except for the order of the submatrices $D(p_{ji}(x)^{\ell_{ji}})$ down the diagonal [4, p.328].

Let $A = (a_{ij})_{i,j=1,2,\ldots,n}$ be an element in $M(n,\mathbb{R})$ and define

(0.6) $$d(A,B) = \|A-B\| = \sum_{i,j=1}^{n} |a_{ij}-b_{ij}| \quad , \; B = (b_{ij}).$$

Then $M(n,\mathbb{R})$ with the metric induced by $\|\ \|$ is a complete metric space [2, theorem 1-11.1].

A sequence of matrices in $M(n,\mathbb{R})$,

(0.7) $$A_1, A_2, \ldots, A_m, A_{m+1}, \ldots ,$$

is said to <u>converge</u> to the matrix A if it converges to A with respect to $\|\ \|$ in (0.6). A sequence of matrices in (0.7) converges to A if and only if the elements of the matrices (0.7) in a given row and column converge to to the corresponding element of the matrix A [2, cor. 1-11.1]. It is immediately clear from this that if T is a fixed matrix, and the matrices A_m converge to A, then $T^{-1}A_mT$ will have $T^{-1}AT$ as their limit. If we replace $\mathbb{C}$ in place of $\mathbb{R}$ in this paragraph, we also have the same results.

Let $(x_n)_{n\in\mathbb{N}}$ be a convergent sequence in a metric space (M,d). If $f : (M,d) \to \mathbb{R}$ (or $\mathbb{C}$) is continuous, then $\lim\limits_{n\to\infty} f(x_n) = f(\lim\limits_{n\to\infty} x_n)$.

<u>Remark</u>: $\det : M(n,\mathbb{C}) \to \mathbb{C}$ (or $\det : M(n,\mathbb{R}) \to \mathbb{R}$) is a continuous function since it is a polynomial function.