LINEAR EQUATIONS

MATANIA BEN-ARTZI

December 2018

Notation

- The scalar product in $\mathbb{R}^m$ is denoted by $(\cdot,\cdot)$.
- Euclidean norm $|x|^2 = \sum_{i=1}^{m} x_i^2$ in $\mathbb{R}^m$.
- For every $A \in Hom(\mathbb{R}^m, \mathbb{R}^m)$ we denote by $\|A\|$ its (operator) norm with respect to $|\cdot|$.
- Notation: $B(x, r)$ for the OPEN ball of radius r center x. The CLOSED ball is denoted by $\overline{B}(x, r)$.
- (a) If $D \subseteq \mathbb{R}^n$ we denote by $C(D, \mathbb{R}^m)$ the set of continuous (vector) functions on D into $\mathbb{R}^m$.

  (b) We denote by $C_b(D, \mathbb{R}^m) \subseteq C(D, \mathbb{R}^m)$ the set of BOUNDED continuous functions on D.

  (c) We denote by $C^k(D, \mathbb{R}^m)$ the subset of functions in $C(D, \mathbb{R}^m)$ which are continuously differentiable up to (including) order k.

  (d) If $m = 1$ we simplify to $C(D), \quad C_b(D), \quad C^k(D)$.

****

- **BASIC DEFINITIONS AND FACTS**
- **
- We FIX $I = (\alpha, \beta) \subseteq \mathbb{R}$ to be an open (finite or infinite) interval.
- We denote $D = I \times \mathbb{R}^m$.
- DEFINITION: A matrix function $t \hookrightarrow A(t) \in Hom(\mathbb{R}^m, \mathbb{R}^m)$, $t \in I$, is **continuous** if all its entries $(a_{i,j}(t))$, $1 \leq i, j \leq m$ are continuous.
- DEFINITION (**A LINEAR DIFFERENTIAL EQUATION**): Given a continuous matrix function $A(t)$ on I and a vector function $b(t) \in C(I, \mathbb{R}^m)$, find a function $y(t) \in C^1(I, \mathbb{R}^m)$ such that:

$$(L) \qquad y'(t) = A(t)y(t) + b(t), \quad t \in I.$$

- The equation (L) refers to **ALL POSSIBLE** solutions .
- **INITIAL VALUE PROBLEM**: Suppose that $(t_0, y^0) \in D$.

  Find a solution of (L), in some open interval $I \subseteq \mathbb{R}$, such that

$$(IV) \qquad t_0 \in I, \quad y(t_0) = y^0.$$

- REMARK: The unknown vector function $y(t)$ consists of m scalar functions (its components). Therefore, sometimes (L) is referred to as a **LINEAR SYSTEM OF EQUATIONS**.
- DEFINITION ($m-$**th order scalar equation**): Let $a_0(t), a_1(t), ..., a_{m-1}(t) \in C(I)$. The equation

$$(SCm) \qquad z^{(m)}(t)+a_{m-1}(t)z^{(m-1)}(t)+...+a_1(t)z'(t)+a_0(t)z(t) = r(t), \quad t \in I,$$

is called an $m-$**th order scalar equation** for the unknown function $z(t)$.

- When the initial conditions (at $t_0 \in I$)

$$(IVSCm) \qquad z(t_0) = z_0, z'(t_0) = z_1, ..., z^{(m-1)}(t_0) = z_{m-1},$$

are added, we get the **initial value problem** for the equation.

- **BASIC OBSERVATION**: Defining a vector function

$$y(t) = (z(t), z'(t), ..., z^{(m-1)}(t))^T \in C(I, \mathbb{R}^m)$$

and $b(t) = (0, 0, ..., 0, r(t))^T \in C(I, \mathbb{R}^m)$, we see that the $m-$th order scalar equation can be written as:

$$y'(t) = A(t)y(t) + b(t),$$

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & 1 \\ -a_0(t) & -a_1(t) & ... & -a_{m-2}(t) & -a_{m-1}(t) \end{pmatrix}.$$

- The initial conditions can be written as:

$$y^0 = (z_0, z_1, ..., z_{m-1})^T.$$

- Thus, theorems formulated for (L) (with initial condition (IV)) hold for (SCm) (with initial conditions (IVSCm)).

****

- **FUNDAMENTAL EXISTENCE AND UNIQUENESS THEOREM**
-

Theorem. *The initial value problem (L)-(IV) has a unique (continuously differentiable) solution in $I = (a, b)$. In particular, for any point $t_0 \in I$ and initial vector $y^0 \in \mathbb{R}^m$, the (unique) solution exists over the full interval I.*

- *Proof.* (a) The equation certainly satisfies the local Lipschitz condition in D, so there exists a unique solution $y(t)$, satisfying the initial condition (IV), in some maximal interval $(t_0-\theta, t_0+\eta)$.

  (b) Suppose $t_0+\eta<\beta$. Take the scalar product of (L) with $2y(t)$ to get
$$\frac{d}{dt}|y(t)|^2 = 2(A(t)y(t), y(t)) + (b(t), y(t)).$$
  Using $|(\xi,\zeta)| \le \frac{1}{2}(|\xi|^2+|\zeta|^2)$, we get
$$\frac{d}{dt}|y(t)|^2 \le (2\|A(t)\|+1)|y(t)|^2 + |b(t)|^2 \le \gamma_1|y(t)|^2+\gamma_2,$$
  where $\gamma_1 = \max\limits_{[t_0,t_0+\eta]}(2\|A(t)\|+1), \quad \gamma_2 = \max\limits_{[t_0,t_0+\eta]}|b(t)|^2$. It follows that
$$\frac{d}{dt}(e^{-\gamma_1 t}|y(t)|^2) \le e^{-\gamma_1 t}\gamma_2, \quad t\in[t_0,t_0+\eta],$$
  so that $\sup\limits_{[t_0,t_0+\eta]}|y(t)|<\infty$.

  (c) The Maximal Interval Theorem now implies that the solution $y(t)$ can be extended beyond $t_0+\eta$, contradicting the assumption about the maximality of $(t_0-\theta, t_0+\eta)$. We conclude that $t_0+\eta=\beta$ and similarly $t_0-\theta=\alpha$. Note this argument holds also if $\beta=+\infty$ or $\alpha=-\infty$.

  □

****

- **THE SCALAR $m-$th ORDER EQUATION**
- COROLLARY: The IVP for Equation (SCm), with initial conditions (IVSCm), has a unique solution in $C^m(I)$.

****

- **COMPLEX VALUED SOLUTIONS**
- We always assume that $A(t)=(a_{i,j}(t))_{1\le i,j\le m}, \quad t\in I$, is a *real matrix*. Similarly, the coefficients $a_i(t)$ in (SCm) are always real.
- (i) If $g(t), \quad t\in I$, is a complex valued function, its derivative is defined simply by
$$\frac{d}{dt}g(t) = \frac{d}{dt}\Re g(t) + i\frac{d}{dt}\Im g(t),$$
  provided the real and imaginary parts $\Re g(t), \Im g(t)$ are differentiable (as real functions).

  (ii) The same definition applies to any order of differentiation.

(iii) The same definition applies to *vector functions* whose components are complex valued—simply differentiate componentwise.

- NOTE: If $\lambda \in \mathbb{C}$ then $\frac{d}{dt}(\lambda g(t)) = \lambda \frac{d}{dt} g(t)$.
- **HENCEFORTH WE ASSUME THAT** $y(t)$ **IN (L) IS COMPLEX-VALUED, i.e** $y(t) \in \mathbb{C}^m$.

****

- **THE HOMOGENEOUS EQUATION,** $b(t) \equiv 0$
-

$$(LH) \qquad y'(t) = A(t)y(t), \quad t \in I, y(t) \in \mathbb{C}^m.$$

- The initial data $y(t_0) = y^0 \in \mathbb{C}^m, \quad t_0 \in I$.
- OBSERVATION: The function $y(t)$ is a solution to (LH), with initial data y^0, if and only if $\Re y(t), \Im y(t)$ are solutions to (LH), with initial data $\Re y^0, \Im y^0$, respectively.
- Note that the fact that $A(t)$ is real is *crucial* in this case.
- COROLLARY(Existence and Uniqueness for Complex-Valued Solutions):

  The initial value problem for (LH) has a unique (continuously differentiable) solution in $I = (a, b)$. In particular, for any point $t_0 \in I$ and initial vector $y^0 \in \mathbb{C}^m$, the (unique) solution exists over the full interval I.
- **THE SUPERPOSITION PRINCIPLE**.
-

  Theorem. *The set of* all *solutions of (LH) is an $m-$dimensional linear space over $\mathbb{C}$.*
- COROLLARY: The set of *all* solutions of the scalar homogeneous equation

$$(SCmH) \ z^{(m)}(t) + a_{m-1}(t)z^{(m-1)}(t) + ... + a_1(t)z'(t) + a_0(t)z(t) = 0, \quad t \in I,$$

  is an $m-$dimensional linear space over $\mathbb{C}$.
- DEFINITION (**FUNDAMENTAL MATRIX**): Let $\{y^1(t), ..., y^m(t)\}$ be a basis for the set of solutions of (LH). The matrix $\Phi(t) \in HOM(\mathbb{C}^m, \mathbb{C}^m)$, whose columns are $\{y^1(t), ..., y^m(t)\}$, is called a **fundamental matrix** for (LH).
- CLAIM: If $\Phi(t) \in HOM(\mathbb{C}^m, \mathbb{C}^m)$ is a fundamental matrix for (LH) then the set of all fundamental matrices of (LH) is given by

$$\{\Phi(t)C, \quad C \in HOM(\mathbb{C}^m, \mathbb{C}^m), \quad det(C) \neq 0\}.$$

- Let $\Phi(t)$ be a fundamental matrix for (LH) and let $\Phi'(t)$ be the matrix obtained by differentiating all elements of Φ. Then:

$$\Phi'(t) = A(t)\Phi(t), \quad t \in I.$$

(consider separately each column).

- REMINDER: The *trace* of $A(t)$: $tr(A(t)) = \sum_{i=1}^{m} a_{i,i}(t)$.
- LEMMA: Let $\Psi(t) \in HOM(\mathbb{C}^m, \mathbb{C}^m)$ be a continuously differentiable matrix function satisfying

$$(*) \quad \Psi'(t) = A(t)\Psi(t), \quad t \in I.$$

Then:

$$\frac{d}{dt} det(\Psi(t)) = tr(A(t)) det(\Psi(t)).$$

- COROLLARY:

$$(a) \quad det(\Psi(t)) = det(\Psi(t_0)) e^{\int_{t_0}^{t} tr(A(s))ds}.$$

(b) Let $\Psi(t)$ be a matrix solution of (*).

Then a necessary and sufficient condition for Ψ to be a fundamental matrix is

$$det(\Psi(\tau)) \neq 0, \quad \text{for some} \quad \tau \in I.$$

- Part (b) can be proved by a direct argument as follows.
- **CLAIM:** If the columns $\{\Psi^1(t), ..., \Psi^m(t)\}$ are *linearly dependent* at *some point* in I, then there exist constants $\{\lambda_1, ..., \lambda_m\} \subseteq \mathbb{C}$, not all zero, such that

$$\sum_{i=1}^{m} \lambda_i \Psi^i(t) \equiv 0, \quad t \in I.$$

Proof. Suppose that

$$\sum_{i=1}^{m} \lambda_i \Psi^i(t_0) \equiv 0, \quad t_0 \in I.$$

Then $y(t) = \sum_{i=1}^{m} \lambda_i \Psi^i(t)$ is a solution of (LH) with $y(t_0) = 0$, so it must vanish identically by uniqueness (of the "trivial" zero solution). □

****

- **THE SCALAR m–th ORDER EQUATION—AGAIN**

- THE FUNDAMENTAL MATRIX:

$$\Phi(t) = \begin{pmatrix} z_1(t) & z_2(t) & z_3(t) & ... & z_m(t) \\ z_1'(t) & z_2'(t) & z_3'(t) & ... & z_m'(t) \\ z_1''(t) & z_2''(t) & z_3''(t) & ... & z_m''(t) \\ ... & ... & ... & ... & ... \\ z_1^{(m-1)}(t) & z_2^{(m-1)}(t) & z_3^{(m-1)}(t) & ... & z_m^{(m-1)}(t) \end{pmatrix}$$

  where $\{z_1(t), ..., z_m(t)\}$ are solutions of the homogeneous equation (SCmH) .
- COROLLARY: $\Phi(t)$ is a fundamental matrix if and only if $\{z_1(t), ..., z_m(t)\}$ are **linearly independent** over $\mathbb{C}$.

  Proof. In order for the columns of $\Phi(t)$ to satisfy

$$\sum_{i=1}^{m} \lambda_i \Phi_i(t) \equiv 0, \quad t \in I$$

  it is necessary and sufficient that

$$\sum_{i=1}^{m} \lambda_i z_i(t) \equiv 0, \quad t \in I.$$

  □

- DEFINITION(**BASIS FOR SOLUTIONS OF (SCmH)**): A set of solutions $\{z_1(t), ..., z_m(t)\}$ for which $\Phi(t)$ is nonsingular (i.e., which are linearly independent over $\mathbb{C}$) is called **a basis for the solutions** of (SCmH).
- In this case $tr(A(t)) = -a_{m-1}(t)$ so

$$\frac{d}{dt} det(\Phi(t)) = -a_{m-1}(t) det(\Phi(t)).$$

- We conclude:

$$W(t) = W(t_0) e^{-\int_{t_0}^{t} a_{m-1}(s) ds},$$

  where $W(t) = det(\Phi(t))$.
- DEFINITION (**WRONSKIAN**): $W(t)$ is called the **Wronskian** of the set $\{z_1(t), ..., z_m(t)\}$ (of solutions to (SCmH)).
- CONCLUSION: A necessary and sufficient condition for the linear independence of the set of solutions $\{z_1(t), ..., z_m(t)\}$ (to (SCmH)) is that at *some point* $t_0 \in I, \quad W(t_0) \neq 0$. It then follows that $W(t) \neq 0$ for *all* points $t \in I$.

****

- **THE CONSTANT COEFFICIENT CASE**—$A(t) \equiv A \in Hom(\mathbb{R}^m, \mathbb{R}^m)$**, A CONSTANT MATRIX**

- We still work with the *HOMOGENEOUS EQUATION*

$$y' = Ay.$$

- OBSERVE: Every solution is extended to the whole line.
- REMINDER(**EXPONENTIAL OF A MATRIX**):

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

- **CLAIM**: $\frac{d}{dt} e^{tA} = Ae^{tA}, \quad t \in \mathbb{R}$.

  Thus, e^{tA} is a *fundamental matrix* for $y' = Ay$.

  Proof. The differentiation formula is obtained by differentiating the series.

  Now, at $t = 0$ the exponential is I, the identity matrix. Since it is regular at $t = 0$, it is regular for all t.
- NOTE: The columns of e^{tA} are the m solutions of $y' = Ay$ with initial data at $t = 0$ equal to (respectively) the standard basis (of $\mathbb{C}^m$) vectors (over the complex numbers).
- **COROLLARY**: $e^{(t+s)A} = e^{tA}e^{sA}, \quad t, s \in \mathbb{R}$.

  Proof. Fix s and consider $e^{(t+s)A}$, $e^{tA}e^{sA}$ as functions of t. Then they are both fundamental matrices for the equation $y' = Ay$, with initial data e^{sA} at $t = 0$.
- **COROLLARY**: $det(e^A) = e^{tr(A)}$.

  Proof. Recall that by the differentiation formula for the determinant of a fundamental matrix

$$\frac{d}{dt} det(e^{tA}) = tr(A) det(e^{tA}),$$

  so that

$$det(e^{tA}) = e^{tr(A)t},$$

  and take $t = 1$.

****

- **THE SCALAR $m-$th ORDER EQUATION—CONSTANT COEFFICIENTS**
- **OPERATOR NOTATION**: For a function $z(t) \in C^m(\mathbb{R})$ we define

$$L(D)z(t) = z^{(m)} + a_{m-1}z^{(m-1)}(t) + ... + a_1 z'(t) + a_0 z(t), \quad t \in \mathbb{R}.$$

  The coefficients are now *real constants.*
- The homogeneous equation is $L(D)z = 0$.
- A basis $\{z_1(t), ..., z_m(t)\}$ consists of m linearly independent solutions.

- "ALGEBRAIZATION" of the problem:
  If $z(t) = e^{\lambda t}$, where $\lambda \in \mathbb{C}$, then
  $$L(D)z(t) = L(\lambda)z(t),$$
  where
  $$L(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + ... + a_1\lambda + a_0,$$
  is a polynomial with real coefficients (which is called the **characteristic polynomial** of the equation).
- DEFINITION: The (complex, in general) zeros of the characteristic polynomial $L(\lambda)$ are called the **characteristic values** of the equation.
- CLAIM: If $\mu \in \mathbb{C}$ is a characteristic value then $e^{\mu t}$ is a solution of the homogeneous equation.
- **COROLLARY**: If there are m different characteristic values (i.e., zeros of the characteristic polynomial) $\{\lambda_1, ..., \lambda_m\}$ then $\left\{z_1(t) = e^{\lambda_1 t}, ..., z_m(t) = e^{\lambda_m t}\right\}$ is a basis for the homogeneous solutions.
  PROOF: To show that it is a basis, consider the matrix
  $$\Phi(t) = \begin{pmatrix} z_1(t) & z_2(t) & z_3(t) & ... & z_m(t) \\ z_1'(t) & z_2'(t) & z_3'(t) & ... & z_m'(t) \\ z_1''(t) & z_2''(t) & z_3''(t) & ... & z_m''(t) \\ ... & ... & ... & ... & ... \\ z_1^{(m-1)}(t) & z_2^{(m-1)}(t) & z_3^{(m-1)}(t) & ... & z_m^{(m-1)}(t) \end{pmatrix},$$
  and note that at $t = 0$ it is regular (think of the *Vandermonde*).
- **AND WHAT IF THERE ARE ZEROS OF THE CHARACTERISTIC POLYNOMIAL WHICH ARE NOT SIMPLE?**
- *Use the operator formulation.*

$$L(D)\frac{\partial}{\partial\lambda}e^{\lambda t} = \frac{\partial}{\partial\lambda}L(D)e^{\lambda t} = \frac{\partial}{\partial\lambda}(L(\lambda)e^{\lambda t}) = (\lambda L(\lambda) + L'(\lambda))e^{\lambda t}.$$

- CONCLUSION: If μ is a double root of $L(\lambda) = 0$ then $e^{\mu t}$ and $te^{\mu t}$ are both solutions of the homogeneous differential equation.
- **THE GENERAL CASE.**
-

  Theorem. *Let* $\{\lambda_1, ..., \lambda_J\}$ *be all the (complex) zeros of* $L(\lambda) = 0$, *with (respective) multiplicities* $\{k_1, ..., k_J\}$*(so that* $k_1 + ... + k_J = m$*). Then the set of functions*
  $$t^l e^{\lambda_j t}, \quad 0 \le l \le k_j - 1,\ 1 \le j \le J,$$

is a basis for the solutions of the homogeneous differential equation.

****

- **THE NONHOMOGENEOUS EQUATION (L)-(IV)**

****

- DEFINITION(**A Particular Solution**): A solution to (L), *not necessarily satisfying* the initial condition (IV), is called a **particular solution**.
- CLAIM: The difference of *any two particular solutions* of (L) is a solution of the HOMOGENEOUS equation (LH).
- COROLLARY: Let $y^{part}(t)$ be a particular solution to (L). Then the set of ALL solutions to (L) is given by

$$y(t) = y^{part}(t) + \Phi(t)c, \quad c \in \mathbb{C}^m,$$

where $\Phi(t)$ is a fundamental matrix of (LH) and c is any arbitrary (column) vector in $\mathbb{C}^m$.

In particular, the (unique!) solution of (L)-(IV) is obtained by taking:

$$c = \Phi(t_0)^{-1}(y^0 - y^{part}(t_0)).$$

- **A GENERAL METHOD FOR FINDING THE SOLUTION TO (L)-(IV)**
- **VARIATION OF THE PARAMETERS**
- LEMMA: Let $\Phi(t)$ be a fundamental matrix of (LH). Then the unique solution to (L)-(IV) is given by:

$$(VP) \quad y(t) = \Phi(t)\Phi(t_0)^{-1}y^0 + \Phi(t)\int_{t_0}^{t} \Phi(s)^{-1}b(s)ds.$$

This formula is also known as **DUHAMEL'S PRINCIPLE.**

PROOF: Define $y(t)$ as in the lemma. Using the equality $\Phi'(t) = A(t)\Phi(t)$ we obtain by simple differentiation (and the fundamental theorem of calculus):

$$y'(t) = A(t)\Phi(t)\Phi(t_0)^{-1}y^0 + A(t)\Phi(t)\int_{t_0}^{t} \Phi(s)^{-1}b(s)ds + b(t),$$

so indeed $y'(t) = A(t)y(t) + b(t)$. Thus it satisfies (L) and from the definition it is clear that $y(t_0) = y^0$.

- **WHY IS IT CALLED "VARIATION OF THE PARAMETERS"?**

- **ANSWER** Consider the second term in the right-hand side of (VP), namely

$$\Phi(t)\int_{t_0}^{t}\Phi(s)^{-1}b(s)ds.$$

It is a particular solution (with initial value 0 at t_0). Denoting by $c(t)$ the column vector

$$c(t)=\int_{t_0}^{t}\Phi(s)^{-1}b(s)ds,$$

It can be written as

$$\Phi(t)\int_{t_0}^{t}\Phi(s)^{-1}b(s)ds=\sum_{j=1}^{m}c_j(t)\Phi_j(t),$$

where $\Phi_j(t)$ is the $j-$th column of $\Phi(t)$.

Note that if $c_j(t)$ were constants the sum would be a solution of the *homogeneous equation.* By "varying" these constants ("parameters") we get $c(t)$, that is determined by the fact that

$$\Phi(t)c'(t)=b(t),\quad c(t_0)=0.$$

- **THE CONSTANT COEFFICIENT CASE,** $A(t)\equiv A$.

We have $A(t)=e^{tA}$, so

$$y(t)=e^{(t-t_0)A}y^0+\int_{t_0}^{t}e^{(t-s)A}b(s)ds.$$

Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel

E-mail address: mbartzi@math.huji.ac.il